



# wwPDB X-ray Structure Validation Summary Report

Oct 10, 2024 – 04:07 PM EDT

PDB ID : 9AU1  
Title : SARS-CoV-2 XBB.1.5 RBD bound to the VIR-7229 and the S309 Fab fragments  
Authors : Rietz, T.; Park, Y.J.; Errico, J.; Czudnochowski, N.; Nix, J.C.; Corti, D.; Snell, G.; Marco, A.D.; Pinto, D.; Cameroni, E.; Seattle Structural Genomics Center for Infectious Disease (SSGCID); Veesler, D.; Structural Genomics Consortium (SGC)  
Deposited on : 2024-02-27  
Resolution : 2.41 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](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 : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

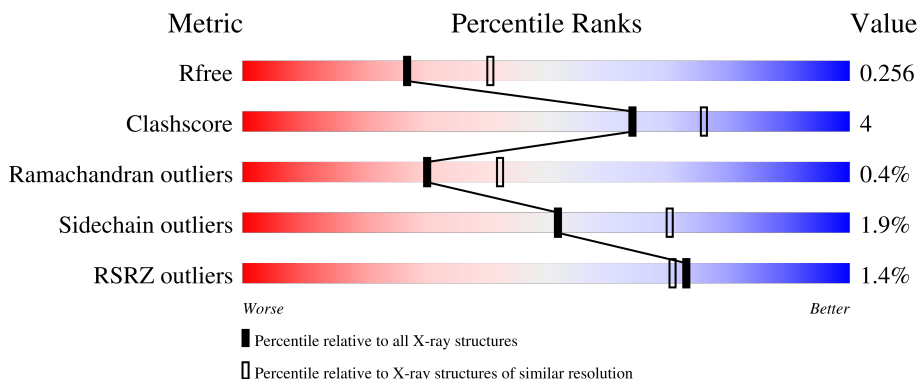
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*





The reported resolution of this entry is 2.41 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.



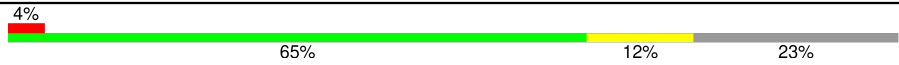
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	5670 (2.44-2.40)
Clashscore	180529	6299 (2.44-2.40)
Ramachandran outliers	177936	6232 (2.44-2.40)
Sidechain outliers	177891	6233 (2.44-2.40)
RSRZ outliers	164620	5670 (2.44-2.40)

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  $\geq 3$ , 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	M	230	 91% 9%
2	N	212	 89% 11%
3	H	225	 91% 7% .
4	L	216	 83% 14% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	R	250	

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
6	K	N	301	-	-	-	X

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 8190 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 S309 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	M	230	1714	1082	286	338	8	0	0	0

- Molecule 2 is a protein called S309 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	N	212	1606	998	271	332	5	0	0	0

- Molecule 3 is a protein called VIR-7229 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	H	219	1668	1071	271	320	6	0	0	0

- Molecule 4 is a protein called VIR-7229 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	L	212	1589	998	262	324	5	0	0	0

- Molecule 5 is a protein called SARS-CoV-2 XBB.1.5 RBD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	R	193	1501	960	254	279	8	0	1	0

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

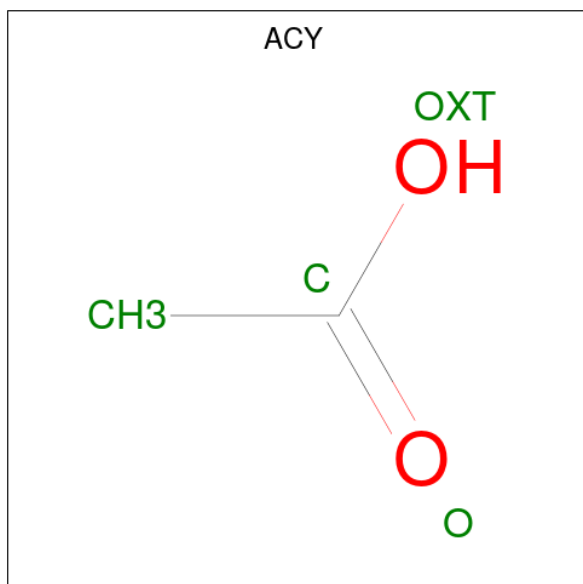
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	N	1	Total 1 K 1	0	0

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



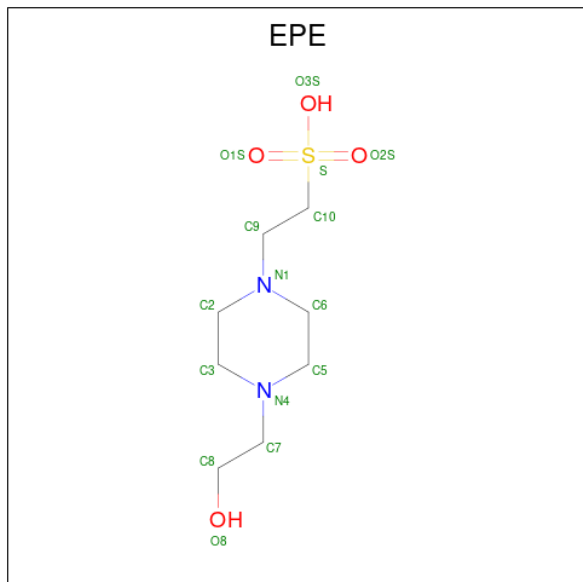
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	H	1	Total	C	O	0	0
			4	2	2		
7	R	1	Total	C	O	0	0
			4	2	2		

- Molecule 8 is ACETIC ACID (three-letter code: ACY) (formula:  $C_2H_4O_2$ ).



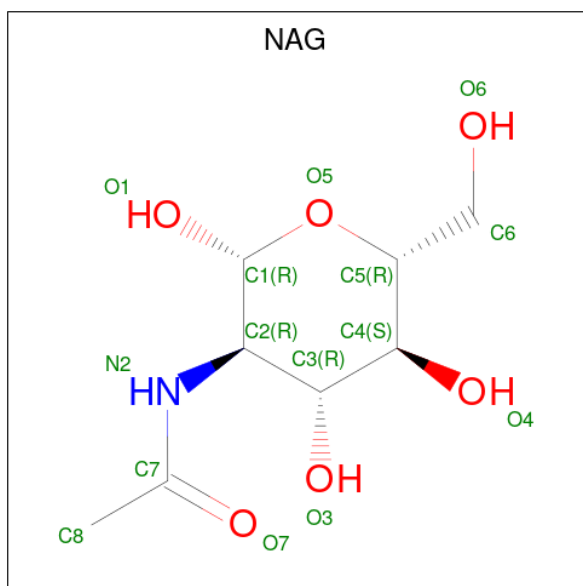
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	L	1	Total	C	O	0	0
			4	2	2		

- Molecule 9 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula:  $C_8H_{18}N_2O_4S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
9	L	1	15	8	2	4	1	0	0

- Molecule 10 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
10	R	1	14	8	1	5	0	0

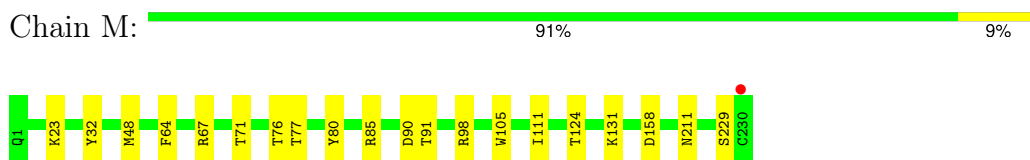
- Molecule 11 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
11	M	25	Total 25	O 25	0	0
11	N	15	Total 15	O 15	0	0
11	H	12	Total 12	O 12	0	0
11	L	9	Total 9	O 9	0	0
11	R	9	Total 9	O 9	0	0

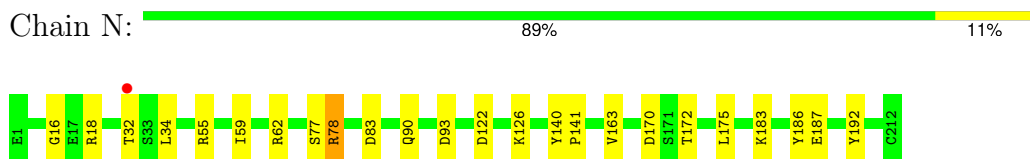
### 3 Residue-property plots [i](#)

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.

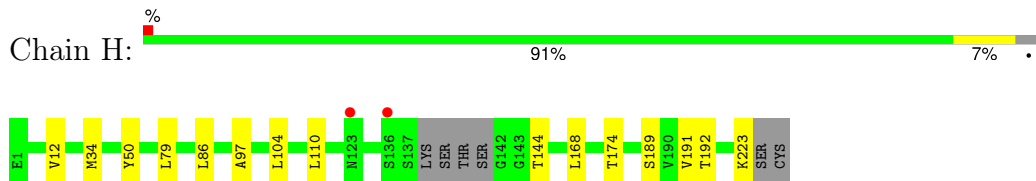
- Molecule 1: S309 Fab heavy chain



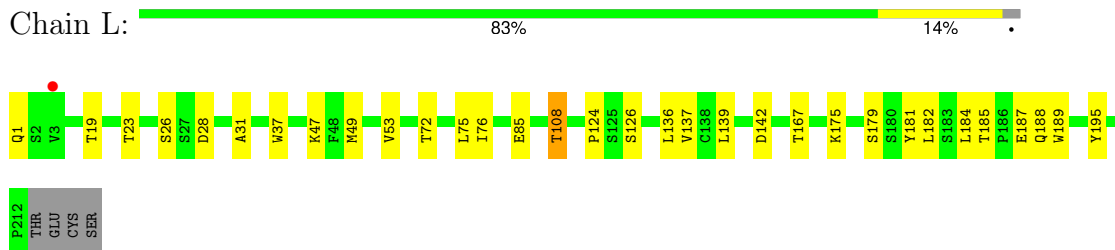
- Molecule 2: S309 Fab light chain



- Molecule 3: VIR-7229 Fab heavy chain



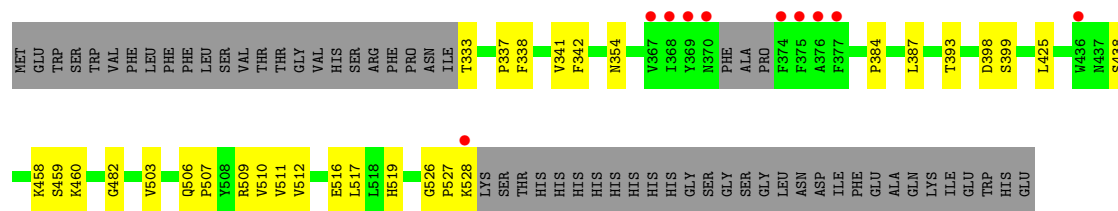
- Molecule 4: VIR-7229 Fab light chain



- Molecule 5: SARS-CoV-2 XBB.1.5 RBD







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.73Å 112.53Å 170.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.98 – 2.41 44.98 – 2.41	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.98-2.41) 99.9 (44.98-2.41)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.25 (at 2.42Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.227 , 0.257 0.225 , 0.256	Depositor DCC
$R_{free}$ test set	2760 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	60.6	Xtrriage
Anisotropy	0.303	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 29.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8190	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	76.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	M	0.45	0/1757	0.61	0/2395
2	N	0.47	0/1637	0.58	0/2224
3	H	0.44	0/1712	0.57	0/2336
4	L	0.42	0/1624	0.55	0/2220
5	R	0.41	0/1547	0.56	0/2110
All	All	0.44	0/8277	0.57	0/11285

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 [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	M	1714	0	1658	12	0
2	N	1606	0	1554	12	0
3	H	1668	0	1626	9	0
4	L	1589	0	1520	18	0
5	R	1501	0	1405	15	0
6	N	1	0	0	0	0
7	H	4	0	6	0	0
7	R	4	0	6	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	L	4	0	3	0	0
9	L	15	0	17	0	0
10	R	14	0	13	1	0
11	H	12	0	0	0	1
11	L	9	0	0	0	0
11	M	25	0	0	1	1
11	N	15	0	0	0	0
11	R	9	0	0	0	0
All	All	8190	0	7808	61	1

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.

The worst 5 of 61 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:L:19:THR:HG22	4:L:76:ILE:HG23	1.72	0.72
5:R:399:SER:HB3	5:R:511:VAL:HG12	1.75	0.68
5:R:384:PRO:HA	5:R:387:LEU:HD12	1.76	0.67
4:L:136:LEU:HD12	4:L:182:LEU:HD23	1.76	0.67
4:L:85:GLU:HG3	4:L:108:THR:HA	1.77	0.65

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
11:M:322:HOH:O	11:H:412:HOH:O[3_555]	2.06	0.14

## 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	M	228/230 (99%)	219 (96%)	9 (4%)	0	100	100
2	N	210/212 (99%)	204 (97%)	6 (3%)	0	100	100
3	H	215/225 (96%)	206 (96%)	9 (4%)	0	100	100
4	L	210/216 (97%)	202 (96%)	7 (3%)	1 (0%)	25	36
5	R	190/250 (76%)	177 (93%)	10 (5%)	3 (2%)	8	10
All	All	1053/1133 (93%)	1008 (96%)	41 (4%)	4 (0%)	30	42

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	R	482	GLY
5	R	519	HIS
4	L	53	VAL
5	R	527	PRO

### 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	M	188/192 (98%)	186 (99%)	2 (1%)	70	84
2	N	181/183 (99%)	177 (98%)	4 (2%)	47	65
3	H	183/192 (95%)	182 (100%)	1 (0%)	86	94
4	L	177/181 (98%)	173 (98%)	4 (2%)	45	64
5	R	160/218 (73%)	154 (96%)	6 (4%)	28	45
All	All	889/966 (92%)	872 (98%)	17 (2%)	52	70

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	R	510	VAL
5	R	517	LEU
4	L	28	ASP
4	L	108	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	L	167	THR

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

Mol	Chain	Res	Type
4	L	188	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](#)

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
4	PCA	L	1	4	7,8,9	2.50	1 (14%)	9,10,12	1.33	1 (11%)

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
4	PCA	L	1	4	-	0/0/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	L	1	PCA	CD-N	5.69	1.48	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	L	1	PCA	CB-CG-CD	2.89	108.88	104.41

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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
9	EPE	L	302	-	15,15,15	1.05	1 (6%)	19,20,20	1.85	6 (31%)
8	ACY	L	301	-	3,3,3	1.22	0	3,3,3	1.50	0
10	NAG	R	602	5	14,14,15	0.54	0	17,19,21	0.63	1 (5%)
7	EDO	R	601	-	3,3,3	0.46	0	2,2,2	0.30	0
7	EDO	H	301	-	3,3,3	0.62	0	2,2,2	0.06	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
10	NAG	R	602	5	-	0/6/23/26	0/1/1/1
9	EPE	L	302	-	-	6/9/19/19	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	EDO	H	301	-	-	1/1/1/1	-
7	EDO	R	601	-	-	1/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	L	302	EPE	C10-S	3.48	1.82	1.77

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	L	302	EPE	C5-N4-C3	3.99	117.43	108.84
9	L	302	EPE	C7-N4-C5	3.37	120.23	111.24
9	L	302	EPE	O3S-S-C10	3.27	112.40	106.00
9	L	302	EPE	O1S-S-C10	2.81	110.97	106.73
9	L	302	EPE	C7-N4-C3	2.23	117.19	111.24

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	L	302	EPE	C9-C10-S-O1S
9	L	302	EPE	C9-C10-S-O3S
9	L	302	EPE	C9-C10-S-O2S
9	L	302	EPE	S-C10-C9-N1
9	L	302	EPE	C10-C9-N1-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	R	602	NAG	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 [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, 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	M	230/230 (100%)	-0.10	1 (0%) 89 87	54, 69, 89, 138	0
2	N	212/212 (100%)	-0.11	1 (0%) 87 86	56, 72, 93, 123	0
3	H	219/225 (97%)	-0.01	2 (0%) 81 79	63, 73, 114, 136	0
4	L	211/216 (97%)	-0.01	1 (0%) 87 86	66, 79, 97, 125	0
5	R	193/250 (77%)	0.29	10 (5%) 34 31	51, 76, 120, 139	1 (0%)
All	All	1065/1133 (93%)	0.01	15 (1%) 73 71	51, 73, 104, 139	1 (0%)

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	R	374	PHE	4.3
5	R	369	TYR	3.9
5	R	370	ASN	3.8
5	R	368	ILE	3.4
5	R	528	LYS	3.2

### 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<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(Å <sup>2</sup> )	Q<0.9
4	PCA	L	1	8/9	0.69	0.19	105,111,116,117	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<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(Å <sup>2</sup> )	Q<0.9
6	K	N	301	1/1	0.37	0.46	205,205,205,205	0
8	ACY	L	301	4/4	0.79	0.15	88,90,90,91	0
10	NAG	R	602	14/15	0.79	0.17	85,90,98,105	14
7	EDO	R	601	4/4	0.80	0.23	93,93,94,96	0
7	EDO	H	301	4/4	0.85	0.15	63,64,65,66	0
9	EPE	L	302	15/15	0.93	0.12	72,75,82,90	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.