Full wwPDB X-ray Structure Validation Report (i)

Apr 30, 2024 – 05:40 pm BST

PDB ID	:	8OWT
Title	:	SARS-CoV-2 spike RBD with A8 and H3 nanobodies bound
Authors	:	Mikolajek, H.; Naismith, J.H.; Owens, R.J.
Deposited on	:	2023-04-28
Resolution	:	2.37 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 (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 (1)) were used in the production of this report:

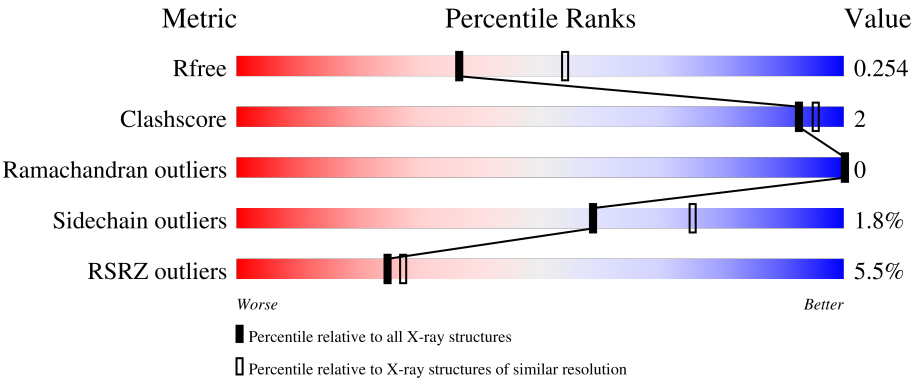
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.37 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.



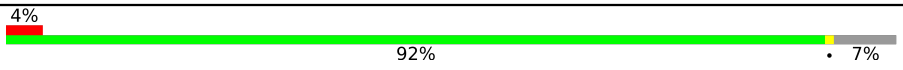
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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	AAA	132	6% 89% · 7%
1	DDD	132	15% 87% 6% 7%
2	BBB	210	% 86% 8% 5%
2	EEE	210	% 87% 6% 7%
3	CCC	136	7% 90% · 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	FFF	136	

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7194 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 Nanobody A8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	123	Total	C	N	O	S	0	1	0
			943	594	166	178	5			
1	DDD	123	Total	C	N	O	S	0	0	0
			940	592	166	177	5			

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	BBB	199	Total	C	N	O	S	0	2	0
			1590	1019	268	295	8			
2	EEE	196	Total	C	N	O	S	0	0	0
			1554	997	260	289	8			

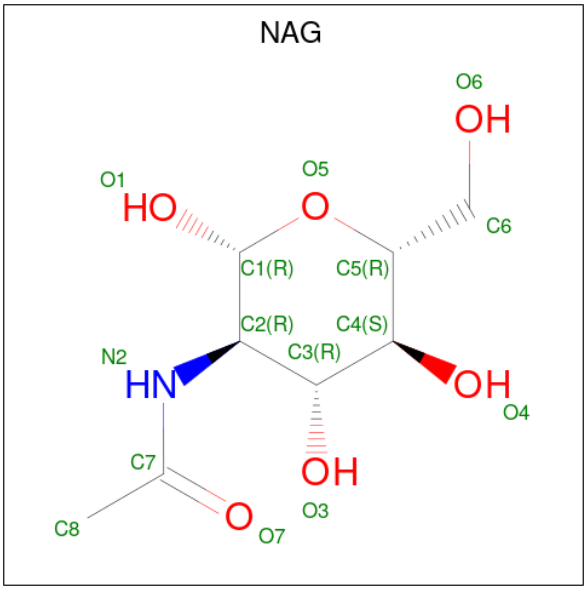
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	533	LYS	-	expression tag	UNP P0DTC2
BBB	534	HIS	-	expression tag	UNP P0DTC2
BBB	535	HIS	-	expression tag	UNP P0DTC2
BBB	536	HIS	-	expression tag	UNP P0DTC2
BBB	537	HIS	-	expression tag	UNP P0DTC2
BBB	538	HIS	-	expression tag	UNP P0DTC2
BBB	539	HIS	-	expression tag	UNP P0DTC2
EEE	533	LYS	-	expression tag	UNP P0DTC2
EEE	534	HIS	-	expression tag	UNP P0DTC2
EEE	535	HIS	-	expression tag	UNP P0DTC2
EEE	536	HIS	-	expression tag	UNP P0DTC2
EEE	537	HIS	-	expression tag	UNP P0DTC2
EEE	538	HIS	-	expression tag	UNP P0DTC2
EEE	539	HIS	-	expression tag	UNP P0DTC2

- Molecule 3 is a protein called Nanobody H3.

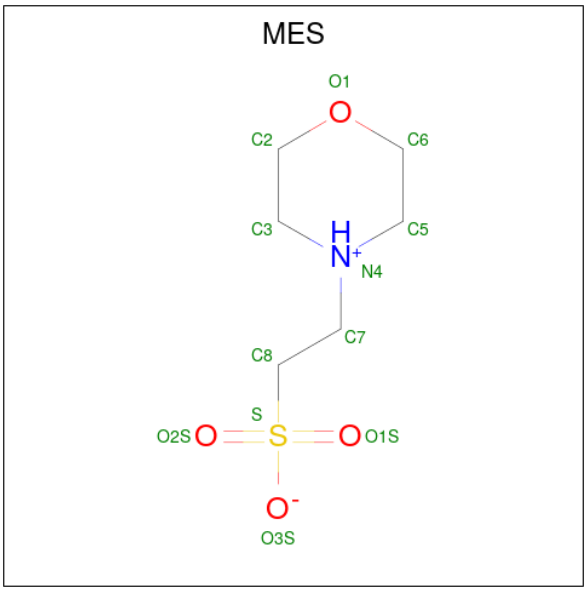
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	CCC	125	Total	C	N	O	S	0	1	0
			973	612	167	189	5			
3	FFF	127	Total	C	N	O	S	0	0	0
			982	616	169	192	5			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	BBB	1	Total	C	N	O	0	0
			14	8	1	5		
4	EEE	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



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

- Molecule 6 is water.

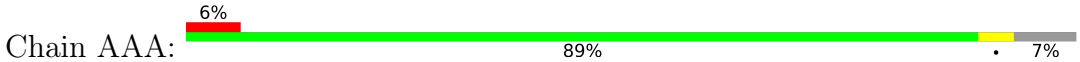
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AAA	16	Total O 16 16	0	0
6	BBB	65	Total O 65 65	0	0
6	CCC	26	Total O 26 26	0	0
6	DDD	5	Total O 5 5	0	0
6	EEE	29	Total O 29 29	0	0
6	FFF	31	Total O 31 31	0	0

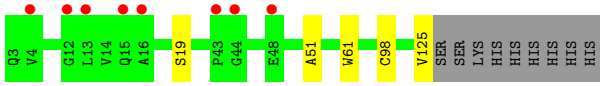
3 Residue-property plots (i)

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: Nanobody A8

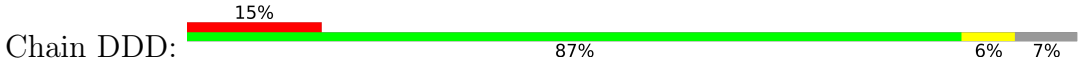
Chain AAA:

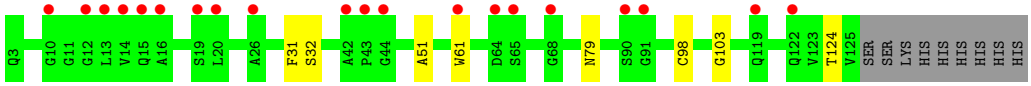




• Molecule 1: Nanobody A8

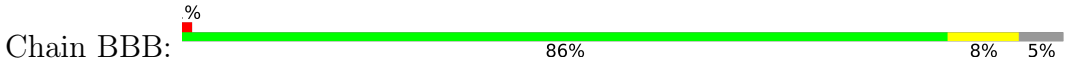
Chain DDD:





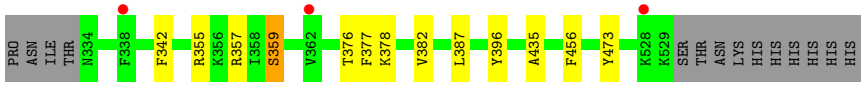
• Molecule 2: Spike protein S1

Chain BBB:



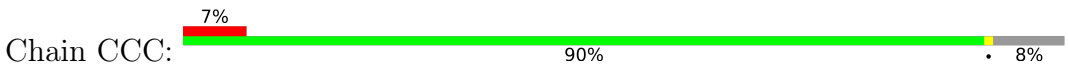
• Molecule 2: Spike protein S1

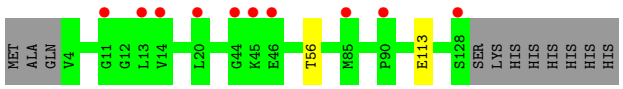
Chain EEE:



• Molecule 3: Nanobody H3

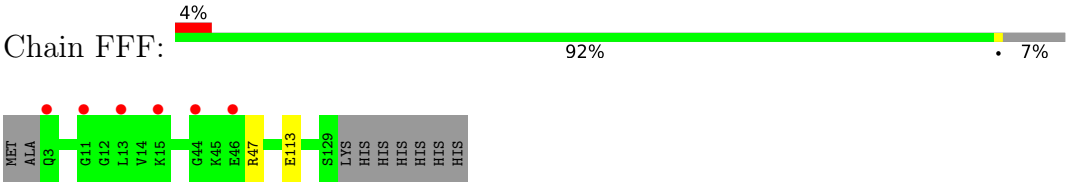
Chain CCC:





• Molecule 3: Nanobody H3

Chain FFF:


4%
92%
7%
MET
ALA
Q3
G11
G12
L13
V14
K15
G44
K45
E46
R47
E113
S129
LYS
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	90.61Å 97.20Å 117.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.78 – 2.37 58.79 – 2.37	Depositor EDS
% Data completeness (in resolution range)	100.0 (58.78-2.37) 100.0 (58.79-2.37)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.23 (at 2.37Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.213 , 0.249 0.218 , 0.254	Depositor DCC
R_{free} test set	2056 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	46.3	Xtriage
Anisotropy	0.433	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 46.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7194	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% 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: NAG, 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	AAA	0.65	0/970	0.74	0/1317
1	DDD	0.64	0/964	0.73	0/1309
2	BBB	0.62	0/1640	0.73	0/2230
2	EEE	0.63	0/1598	0.71	0/2173
3	CCC	0.65	0/998	0.74	0/1350
3	FFF	0.65	0/1004	0.75	0/1358
All	All	0.64	0/7174	0.73	0/9737

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
2	BBB	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	BBB	494	SER	Peptide

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	AAA	943	0	888	1	0
1	DDD	940	0	883	3	0
2	BBB	1590	0	1521	12	0
2	EEE	1554	0	1478	9	0
3	CCC	973	0	932	1	0
3	FFF	982	0	937	0	0
4	BBB	14	0	13	0	0
4	EEE	14	0	13	1	0
5	EEE	12	0	13	0	0
6	AAA	16	0	0	0	0
6	BBB	65	0	0	1	0
6	CCC	26	0	0	0	0
6	DDD	5	0	0	0	0
6	EEE	29	0	0	0	0
6	FFF	31	0	0	0	0
All	All	7194	0	6678	21	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 (21) 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)
2:BBB:458:LYS:NZ	2:BBB:474:GLN:O	2.30	0.51
2:BBB:430:THR:HG22	6:BBB:701:HOH:O	2.10	0.51
2:EEE:382:VAL:HG21	2:EEE:387:LEU:HD21	1.93	0.51
2:BBB:364:ASP:O	2:BBB:367:VAL:HG12	2.12	0.49
2:BBB:333:THR:HG21	2:EEE:355:ARG:H	1.78	0.48
1:DDD:103:GLY:O	2:EEE:378:LYS:HA	2.14	0.48
2:BBB:470:THR:HB	3:CCC:56:THR:HG21	1.96	0.46
2:BBB:394:ASN:OD1	2:EEE:357:ARG:NH2	2.50	0.45
2:BBB:376:THR:HB	2:BBB:435:ALA:HB3	1.99	0.44
2:BBB:333:THR:O	2:BBB:333:THR:HG22	2.18	0.43
2:EEE:342:PHE:HB2	4:EEE:601:NAG:H82	2.01	0.43
2:EEE:376:THR:HB	2:EEE:435:ALA:HB3	2.00	0.42
2:BBB:500:THR:O	2:BBB:500:THR:HG22	2.18	0.42
2:BBB:338:PHE:CE1	2:BBB:358:ILE:HD13	2.54	0.42
2:BBB:425:LEU:HD21	2:BBB:512:VAL:HG11	2.02	0.41
2:BBB:357:ARG:HH22	2:EEE:359:SER:HB2	1.86	0.41
2:EEE:355:ARG:HH11	2:EEE:396:TYR:HB3	1.85	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:DDD:31:PHE:CD2	1:DDD:79:ASN:HA	2.56	0.41
2:EEE:456:PHE:HB3	2:EEE:473:TYR:CD2	2.55	0.41
1:DDD:51:ALA:HA	1:DDD:61:TRP:O	2.21	0.40
1:AAA:51:ALA:HA	1:AAA:61:TRP:O	2.21	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	AAA	122/132 (92%)	119 (98%)	3 (2%)	0	100	100
1	DDD	121/132 (92%)	116 (96%)	5 (4%)	0	100	100
2	BBB	199/210 (95%)	190 (96%)	9 (4%)	0	100	100
2	EEE	194/210 (92%)	187 (96%)	7 (4%)	0	100	100
3	CCC	124/136 (91%)	118 (95%)	6 (5%)	0	100	100
3	FFF	125/136 (92%)	123 (98%)	2 (2%)	0	100	100
All	All	885/956 (93%)	853 (96%)	32 (4%)	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	AAA	95/103 (92%)	92 (97%)	3 (3%)	39	56
1	DDD	94/103 (91%)	91 (97%)	3 (3%)	39	56
2	BBB	174/183 (95%)	172 (99%)	2 (1%)	73	86
2	EEE	169/183 (92%)	167 (99%)	2 (1%)	71	84
3	CCC	103/112 (92%)	102 (99%)	1 (1%)	76	87
3	FFF	104/112 (93%)	102 (98%)	2 (2%)	57	73
All	All	739/796 (93%)	726 (98%)	13 (2%)	59	75

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	19	SER
1	AAA	98	CYS
1	AAA	125	VAL
2	BBB	333	THR
2	BBB	377	PHE
3	CCC	113	GLU
1	DDD	32	SER
1	DDD	98	CYS
1	DDD	124	THR
2	EEE	359	SER
2	EEE	377	PHE
3	FFF	47	ARG
3	FFF	113	GLU

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)

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)

3 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	$\#\|Z\| > 2$
4	NAG	EEE	601	2	14,14,15	0.25	0	17,19,21	0.92	1 (5%)
5	MES	EEE	602	-	12,12,12	0.75	0	14,16,16	0.36	0
4	NAG	BBB	601	2	14,14,15	0.29	0	17,19,21	0.70	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
4	NAG	EEE	601	2	-	1/6/23/26	0/1/1/1
5	MES	EEE	602	-	-	0/6/14/14	0/1/1/1
4	NAG	BBB	601	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	EEE	601	NAG	O5-C5-C6	2.49	111.11	107.20

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	EEE	601	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	EEE	601	NAG	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	AAA	123/132 (93%)	0.69	8 (6%) 18 20	43, 58, 95, 124	0
1	DDD	123/132 (93%)	1.02	20 (16%) 1 1	46, 69, 110, 121	0
2	BBB	199/210 (94%)	0.41	2 (1%) 82 83	35, 47, 66, 92	0
2	EEE	196/210 (93%)	0.47	3 (1%) 73 75	41, 55, 80, 100	0
3	CCC	125/136 (91%)	0.64	10 (8%) 12 13	38, 51, 100, 113	0
3	FFF	127/136 (93%)	0.42	6 (4%) 31 34	40, 49, 87, 100	0
All	All	893/956 (93%)	0.58	49 (5%) 25 27	35, 54, 97, 124	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	DDD	43	PRO	7.1
3	CCC	11	GLY	5.8
1	DDD	14	VAL	5.6
1	AAA	43	PRO	4.8
1	DDD	91	GLY	4.7
3	CCC	128	SER	4.5
1	DDD	90	SER	4.4
1	AAA	44	GLY	4.0
1	DDD	44	GLY	3.7
3	CCC	20	LEU	3.4
1	DDD	12	GLY	3.3
3	CCC	44	GLY	3.2
1	DDD	13	LEU	3.2
1	AAA	15	GLN	3.1
1	AAA	12	GLY	3.1
3	FFF	13	LEU	3.1
2	EEE	362	VAL	2.9
3	CCC	13	LEU	2.8
1	AAA	13	LEU	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	DDD	65	SER	2.7
1	DDD	68	GLY	2.6
1	AAA	16	ALA	2.6
3	FFF	15	LYS	2.6
2	EEE	528	LYS	2.5
2	BBB	333	THR	2.4
3	CCC	90	PRO	2.4
1	DDD	61	TRP	2.4
1	DDD	16	ALA	2.4
3	FFF	46	GLU	2.4
3	FFF	3	GLN	2.4
3	CCC	14	VAL	2.4
2	BBB	334	ASN	2.3
1	DDD	42	ALA	2.3
1	DDD	19	SER	2.2
1	DDD	15	GLN	2.2
3	CCC	45	LYS	2.1
2	EEE	338	PHE	2.1
1	DDD	20	LEU	2.1
3	CCC	85	MET	2.1
1	DDD	122	GLN	2.1
1	DDD	26	ALA	2.0
1	AAA	48	GLU	2.0
1	DDD	10	GLY	2.0
3	FFF	44	GLY	2.0
1	AAA	4	VAL	2.0
3	CCC	46	GLU	2.0
1	DDD	119	GLN	2.0
3	FFF	11	GLY	2.0
1	DDD	64	ASP	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

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	MES	EEE	602	12/12	0.81	0.18	80,85,96,106	0
4	NAG	EEE	601	14/15	0.91	0.18	73,81,90,97	0
4	NAG	BBB	601	14/15	0.94	0.11	42,48,51,53	0

6.5 Other polymers

There are no such residues in this entry.