



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2020 – 04:59 PM BST

PDB ID : 5X2Q  
Title : Crystal structure of the medaka fish taste receptor T1r2a-T1r3 ligand binding domains in complex with glycine  
Authors : Nuemket, N.; Yasui, N.; Atsumi, N.; Yamashita, A.  
Deposited on : 2017-02-02  
Resolution : 2.60 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 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	:	2.13.1
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)	:	2.13.1

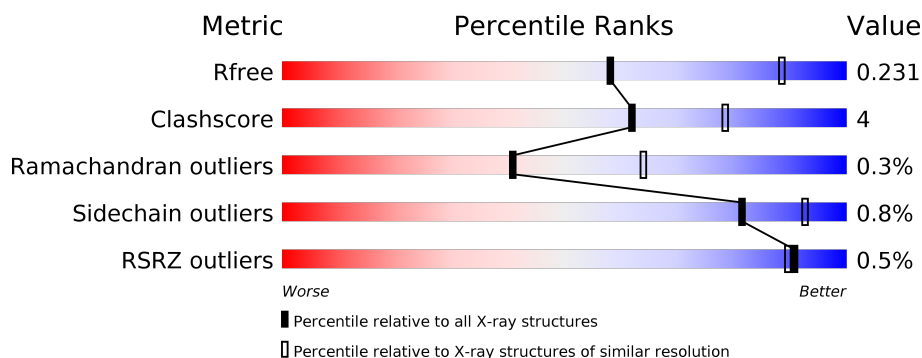
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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.





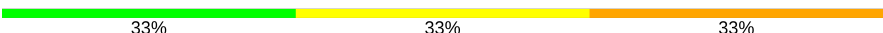

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 on 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	461	<div> <div>%</div> <div> <div></div> <div>82%</div> <div>9%</div> <div>9%</div> </div> </div>
1	C	461	<div> <div></div> <div>82%</div> <div>11%</div> <div>7%</div> </div>
2	B	478	<div> <div>%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>6%</div> </div> </div>
2	D	478	<div> <div>%</div> <div> <div></div> <div>84%</div> <div>10%</div> <div>6%</div> </div> </div>
3	H	225	<div> <div></div> <div>88%</div> <div>9%</div> <div>• •</div> </div>
3	J	225	<div> <div></div> <div>85%</div> <div>12%</div> <div>•</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	K	217	 89% 10%
4	L	217	 88% 12%
5	E	3	 33% 33% 33%
6	F	2	 50% 50%

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
7	NAG	A	902	-	-	-	X

## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 21264 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 Taste receptor, type 1, member 2a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	421	Total	C	N	O	S	0	0	0
			3342	2161	553	612	16			
1	C	431	Total	C	N	O	S	0	0	0
			3407	2197	566	628	16			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	475	SER	-	expression tag	UNP A0A173M0G2
A	476	GLY	-	expression tag	UNP A0A173M0G2
A	477	ILE	-	expression tag	UNP A0A173M0G2
A	478	GLU	-	expression tag	UNP A0A173M0G2
A	479	GLY	-	expression tag	UNP A0A173M0G2
A	480	ARG	-	expression tag	UNP A0A173M0G2
C	475	SER	-	expression tag	UNP A0A173M0G2
C	476	GLY	-	expression tag	UNP A0A173M0G2
C	477	ILE	-	expression tag	UNP A0A173M0G2
C	478	GLU	-	expression tag	UNP A0A173M0G2
C	479	GLY	-	expression tag	UNP A0A173M0G2
C	480	ARG	-	expression tag	UNP A0A173M0G2

- Molecule 2 is a protein called Taste receptor, type 1, member 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	449	Total	C	N	O	S	0	0	0
			3517	2260	567	675	15			
2	D	451	Total	C	N	O	S	0	0	0
			3536	2275	569	677	15			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	492	SER	-	expression tag	UNP A0A173M094
B	493	GLY	-	expression tag	UNP A0A173M094
B	494	ILE	-	expression tag	UNP A0A173M094
B	495	GLU	-	expression tag	UNP A0A173M094
B	496	GLY	-	expression tag	UNP A0A173M094
B	497	ARG	-	expression tag	UNP A0A173M094
D	492	SER	-	expression tag	UNP A0A173M094
D	493	GLY	-	expression tag	UNP A0A173M094
D	494	ILE	-	expression tag	UNP A0A173M094
D	495	GLU	-	expression tag	UNP A0A173M094
D	496	GLY	-	expression tag	UNP A0A173M094
D	497	ARG	-	expression tag	UNP A0A173M094

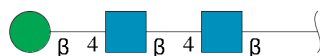
- Molecule 3 is a protein called Fab16A Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	219	Total	C	N	O	S	0	0	0
			1662	1056	274	324	8			
3	J	218	Total	C	N	O	S	0	0	0
			1654	1050	273	323	8			

- Molecule 4 is a protein called Fab16A Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	217	Total	C	N	O	S	0	0	0
			1679	1040	288	345	6			
4	K	216	Total	C	N	O	S	0	0	0
			1670	1035	287	342	6			

- Molecule 5 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



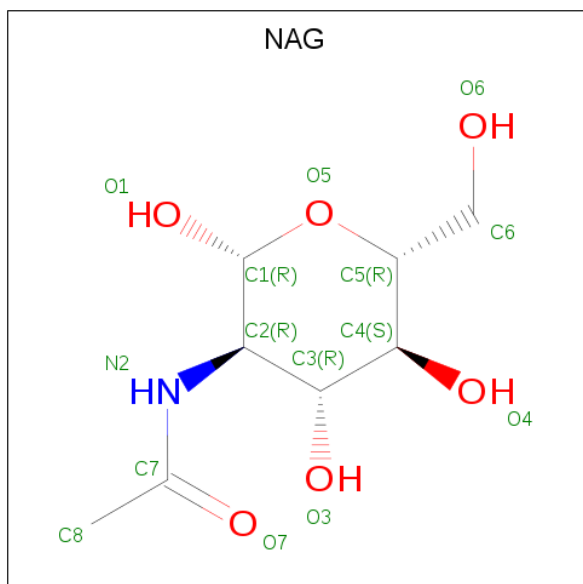
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	E	3	Total	C	N	O	0	0	0
			39	22	2	15			

- Molecule 6 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	F	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



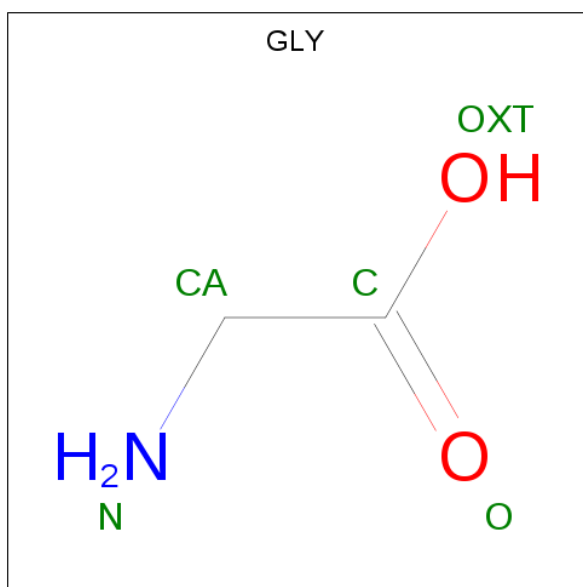
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	B	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	B	1	Total	C	N	O	0	0
			14	8	1	5		
7	B	1	Total	C	N	O	0	0
			14	8	1	5		
7	B	1	Total	C	N	O	0	0
			14	8	1	5		
7	B	1	Total	C	N	O	0	0
			14	8	1	5		
7	C	1	Total	C	N	O	0	0
			14	8	1	5		
7	C	1	Total	C	N	O	0	0
			14	8	1	5		
7	C	1	Total	C	N	O	0	0
			14	8	1	5		
7	C	1	Total	C	N	O	0	0
			14	8	1	5		
7	C	1	Total	C	N	O	0	0
			14	8	1	5		
7	D	1	Total	C	N	O	0	0
			14	8	1	5		
7	D	1	Total	C	N	O	0	0
			14	8	1	5		
7	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 8 is GLYCINE (three-letter code: GLY) (formula: C<sub>2</sub>H<sub>5</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	N	O	0	0
			5	2	1	2		
8	B	1	Total	C	N	O	0	0
			5	2	1	2		
8	C	1	Total	C	N	O	0	0
			5	2	1	2		
8	D	1	Total	C	N	O	0	0
			5	2	1	2		

- Molecule 9 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	2	Total	Na	0	0
			2	2		
9	D	1	Total	Na	0	0
			1	1		
9	C	2	Total	Na	0	0
			2	2		

- Molecule 10 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	B	1	Total	Cl	0	0
			1	1		
10	D	1	Total	Cl	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	C	1	Total	Cl	0	0
			1	1		

- Molecule 11 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	K	1	Total	Ca	0	0
			1	1		

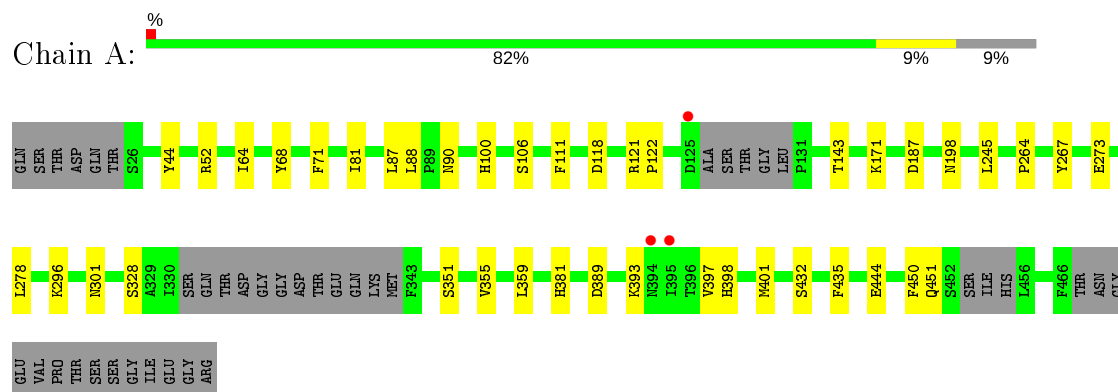
- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	79	Total	O	0	0
			79	79		
12	B	27	Total	O	0	0
			27	27		
12	C	84	Total	O	0	0
			84	84		
12	D	50	Total	O	0	0
			50	50		
12	H	38	Total	O	0	0
			38	38		
12	L	47	Total	O	0	0
			47	47		
12	J	42	Total	O	0	0
			42	42		
12	K	40	Total	O	0	0
			40	40		

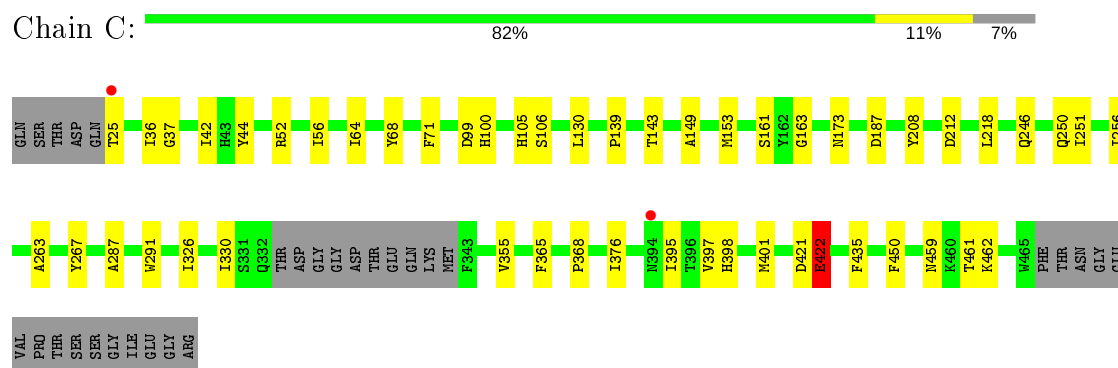
### 3 Residue-property plots

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.

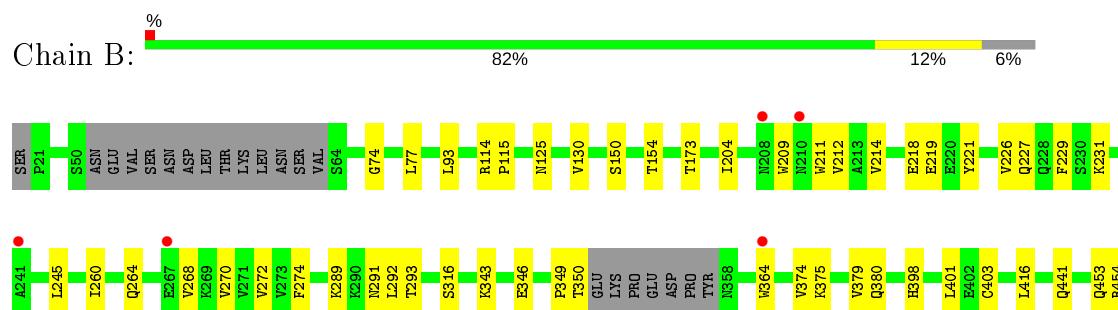
- Molecule 1: Taste receptor, type 1, member 2a



- Molecule 1: Taste receptor, type 1, member 2a

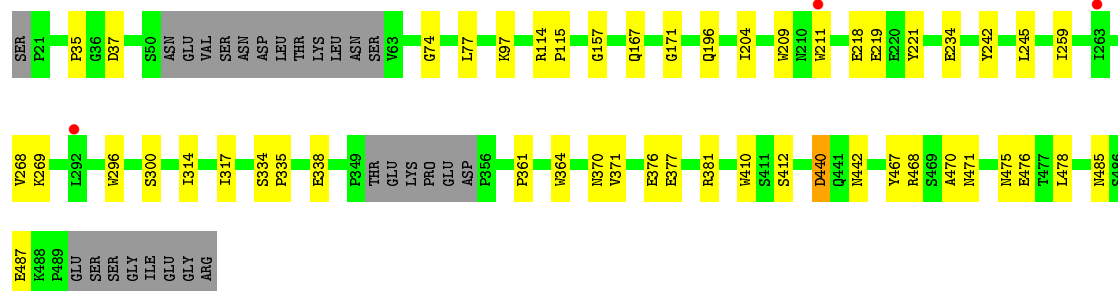
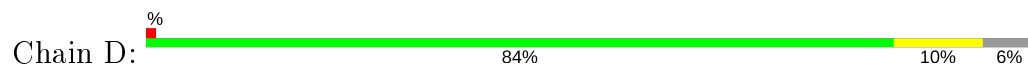


- Molecule 2: Taste receptor, type 1, member 3

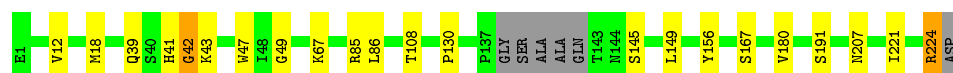
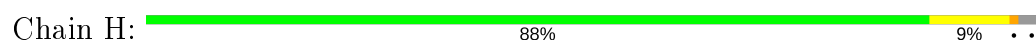




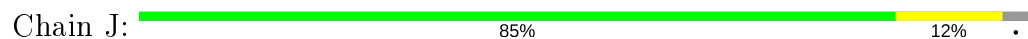
- Molecule 2: Taste receptor, type 1, member 3



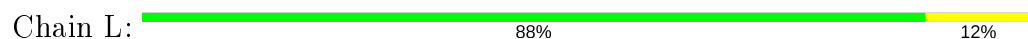
- Molecule 3: Fab16A Heavy chain



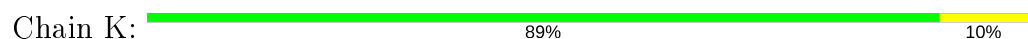
- Molecule 3: Fab16A Heavy chain



- Molecule 4: Fab16A Light chain



- Molecule 4: Fab16A Light chain



- Molecule 5: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose





- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:  50% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.52Å 112.72Å 128.69Å 90.00° 92.16° 90.00°	Depositor
Resolution (Å)	49.28 – 2.60 49.27 – 2.60	Depositor EDS
% Data completeness (in resolution range)	97.8 (49.28-2.60) 97.8 (49.27-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.02 (at 2.61Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.153 , 0.230 0.158 , 0.231	Depositor DCC
$R_{free}$ test set	4250 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.1	Xtriage
Anisotropy	0.439	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 49.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.032 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	21264	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% 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: NA, CA, BMA, NAG, CL

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	0.42	0/3428	0.56	0/4653
1	C	0.45	0/3495	0.57	0/4749
2	B	0.41	0/3598	0.54	0/4893
2	D	0.41	0/3619	0.54	0/4922
3	H	0.41	0/1706	0.59	0/2331
3	J	0.41	0/1696	0.58	1/2314 (0.0%)
4	K	0.47	0/1707	0.63	1/2320 (0.0%)
4	L	0.47	0/1716	0.59	0/2332
All	All	0.43	0/20965	0.57	2/28514 (0.0%)

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
3	H	0	1
3	J	0	1
All	All	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	K	215	ARG	NE-CZ-NH1	-6.68	116.96	120.30
3	J	50	LEU	CA-CB-CG	5.11	127.05	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	H	42	GLY	Peptide
3	J	42	GLY	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	A	3342	0	3262	26	0
1	C	3407	0	3329	30	0
2	B	3517	0	3476	34	0
2	D	3536	0	3497	31	0
3	H	1662	0	1642	18	0
3	J	1654	0	1629	14	0
4	K	1670	0	1596	13	0
4	L	1679	0	1602	22	0
5	E	39	0	34	2	0
6	F	28	0	25	2	0
7	A	98	0	91	4	0
7	B	70	0	65	2	0
7	C	84	0	78	1	0
7	D	42	0	39	1	0
8	A	5	0	2	0	0
8	B	5	0	2	1	0
8	C	5	0	2	0	0
8	D	5	0	2	0	0
9	A	2	0	0	0	0
9	C	2	0	0	0	0
9	D	1	0	0	0	0
10	B	1	0	0	0	0
10	C	1	0	0	0	0
10	D	1	0	0	0	0
11	K	1	0	0	0	0
12	A	79	0	0	0	0
12	B	27	0	0	0	0
12	C	84	0	0	3	0
12	D	50	0	0	3	1
12	H	38	0	0	0	0
12	J	42	0	0	0	1
12	K	40	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	L	47	0	0	1	0
All	All	21264	0	20373	180	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 180 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)
3:H:167:SER:H	3:H:207:ASN:HD21	1.09	0.94
4:K:193:HIS:O	4:K:215:ARG:NH1	2.08	0.87
1:A:88:LEU:H	1:A:381:HIS:HD2	1.28	0.79
4:K:199:GLU:HG2	4:K:210:VAL:HG22	1.65	0.79
2:D:440:ASP:HB3	2:D:442:ASN:H	1.46	0.79

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)
12:D:645:HOH:O	12:J:342:HOH:O[1_455]	2.17	0.03

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	413/461 (90%)	399 (97%)	14 (3%)	0	100	100
1	C	427/461 (93%)	411 (96%)	14 (3%)	2 (0%)	29	52
2	B	443/478 (93%)	426 (96%)	15 (3%)	2 (0%)	29	52
2	D	445/478 (93%)	430 (97%)	13 (3%)	2 (0%)	34	57
3	H	215/225 (96%)	207 (96%)	8 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	J	212/225 (94%)	208 (98%)	4 (2%)	0	100	100
4	K	214/217 (99%)	207 (97%)	6 (3%)	1 (0%)	29	52
4	L	215/217 (99%)	209 (97%)	5 (2%)	1 (0%)	29	52
All	All	2584/2762 (94%)	2497 (97%)	79 (3%)	8 (0%)	41	64

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	291	ASN
2	B	349	PRO
1	C	422	GLU
4	L	72	ARG
1	C	163	GLY

### 5.3.2 Protein sidechains ⓘ

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	374/407 (92%)	367 (98%)	7 (2%)	57	79
1	C	382/407 (94%)	378 (99%)	4 (1%)	76	90
2	B	401/428 (94%)	399 (100%)	2 (0%)	88	96
2	D	403/428 (94%)	401 (100%)	2 (0%)	88	96
3	H	189/192 (98%)	187 (99%)	2 (1%)	73	88
3	J	187/192 (97%)	185 (99%)	2 (1%)	73	88
4	K	190/191 (100%)	190 (100%)	0	100	100
4	L	191/191 (100%)	191 (100%)	0	100	100
All	All	2317/2436 (95%)	2298 (99%)	19 (1%)	81	92

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	463	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	395	ILE
3	H	145	SER
2	B	274	PHE
3	H	224	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	441	ASN
2	D	224	GLN
3	J	210	HIS
2	D	196	GLN
2	D	227	GLN

### 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 ⓘ

5 monosaccharides 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$
5	NAG	E	1	2,5	14,14,15	1.00	1 (7%)	17,19,21	1.04	1 (5%)
5	NAG	E	2	5	14,14,15	1.08	1 (7%)	17,19,21	2.18	6 (35%)
5	BMA	E	3	5	11,11,12	0.78	0	15,15,17	0.62	0
6	NAG	F	1	2,6	14,14,15	0.69	0	17,19,21	1.23	2 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	F	2	6	14,14,15	0.48	0	17,19,21	1.40	1 (5%)

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
5	NAG	E	1	2,5	-	2/6/23/26	0/1/1/1
5	NAG	E	2	5	-	4/6/23/26	0/1/1/1
5	BMA	E	3	5	-	0/2/19/22	0/1/1/1
6	NAG	F	1	2,6	-	2/6/23/26	0/1/1/1
6	NAG	F	2	6	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	1	NAG	O5-C1	-2.76	1.39	1.43
5	E	2	NAG	O5-C1	-2.57	1.39	1.43

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	2	NAG	C2-N2-C7	-4.80	116.06	122.90
5	E	2	NAG	C3-C4-C5	3.96	117.31	110.24
5	E	2	NAG	O5-C1-C2	-3.45	105.83	111.29
6	F	2	NAG	C1-O5-C5	3.30	116.66	112.19
5	E	2	NAG	C4-C3-C2	-3.04	106.57	111.02

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	2	NAG	C4-C5-C6-O6
5	E	2	NAG	C8-C7-N2-C2
5	E	2	NAG	O7-C7-N2-C2
5	E	1	NAG	C8-C7-N2-C2
5	E	1	NAG	O7-C7-N2-C2

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	2	NAG	2	0
6	F	2	NAG	2	0

## 5.6 Ligand geometry

Of 34 ligands modelled in this entry, 9 are monoatomic - leaving 25 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$
7	NAG	B	906	2	14,14,15	0.45	0	17,19,21	1.24	1 (5%)
7	NAG	A	906	1	14,14,15	0.64	0	17,19,21	0.83	0
7	NAG	A	902	1	14,14,15	0.32	0	17,19,21	1.29	1 (5%)
7	NAG	C	905	1	14,14,15	0.50	0	17,19,21	1.73	3 (17%)
7	NAG	A	905	1	14,14,15	0.76	0	17,19,21	1.09	1 (5%)
7	NAG	D	903	2	14,14,15	0.66	0	17,19,21	2.19	4 (23%)
7	NAG	C	901	1	14,14,15	0.82	1 (7%)	17,19,21	1.59	3 (17%)
7	NAG	C	906	1	14,14,15	0.65	0	17,19,21	1.06	1 (5%)
7	NAG	A	901	1	14,14,15	0.84	0	17,19,21	1.37	2 (11%)
7	NAG	B	901	2	14,14,15	1.80	3 (21%)	17,19,21	3.54	11 (64%)
7	NAG	C	904	1	14,14,15	0.59	0	17,19,21	1.11	2 (11%)
7	NAG	B	907	2	14,14,15	0.55	0	17,19,21	1.86	2 (11%)
7	NAG	D	904	2	14,14,15	0.76	0	17,19,21	1.05	1 (5%)
7	NAG	A	907	1	14,14,15	0.67	0	17,19,21	0.97	0
7	NAG	D	905	2	14,14,15	0.67	0	17,19,21	0.86	1 (5%)
7	NAG	B	905	2	14,14,15	0.54	0	17,19,21	1.46	3 (17%)
7	NAG	A	903	1	14,14,15	0.89	1 (7%)	17,19,21	1.71	5 (29%)
7	NAG	A	904	1	14,14,15	0.48	0	17,19,21	1.12	1 (5%)
7	NAG	C	903	1	14,14,15	1.19	2 (14%)	17,19,21	2.43	6 (35%)
7	NAG	B	908	2	14,14,15	1.21	1 (7%)	17,19,21	2.63	6 (35%)
7	NAG	C	902	1	14,14,15	0.46	0	17,19,21	1.58	3 (17%)

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
7	NAG	B	906	2	-	4/6/23/26	0/1/1/1
7	NAG	A	906	1	-	4/6/23/26	0/1/1/1
7	NAG	A	902	1	-	2/6/23/26	0/1/1/1
7	NAG	C	905	1	-	2/6/23/26	0/1/1/1
7	NAG	A	905	1	-	2/6/23/26	0/1/1/1
7	NAG	D	903	2	-	2/6/23/26	0/1/1/1
7	NAG	C	901	1	-	2/6/23/26	0/1/1/1
7	NAG	C	906	1	-	4/6/23/26	0/1/1/1
7	NAG	A	901	1	-	2/6/23/26	0/1/1/1
7	NAG	B	901	2	-	4/6/23/26	0/1/1/1
7	NAG	C	904	1	-	3/6/23/26	0/1/1/1
7	NAG	B	907	2	-	4/6/23/26	0/1/1/1
7	NAG	D	904	2	-	3/6/23/26	0/1/1/1
7	NAG	A	907	1	-	4/6/23/26	0/1/1/1
7	NAG	D	905	2	-	4/6/23/26	0/1/1/1
7	NAG	B	905	2	-	4/6/23/26	0/1/1/1
7	NAG	A	903	1	-	4/6/23/26	0/1/1/1
7	NAG	A	904	1	-	3/6/23/26	0/1/1/1
7	NAG	C	903	1	-	2/6/23/26	0/1/1/1
7	NAG	B	908	2	-	3/6/23/26	0/1/1/1
7	NAG	C	902	1	-	2/6/23/26	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	901	NAG	O5-C5	-4.46	1.34	1.43
7	B	908	NAG	O5-C1	-3.24	1.38	1.43
7	C	903	NAG	O5-C1	-2.83	1.39	1.43
7	A	903	NAG	C1-C2	2.69	1.56	1.52
7	C	903	NAG	O5-C5	-2.60	1.38	1.43

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	908	NAG	C1-O5-C5	-8.49	100.69	112.19
7	B	901	NAG	C4-C3-C2	7.92	122.62	111.02
7	B	901	NAG	O5-C1-C2	-6.94	100.33	111.29
7	D	903	NAG	C1-O5-C5	6.88	121.52	112.19
7	B	907	NAG	C1-O5-C5	5.71	119.93	112.19

There are no chirality outliers.

5 of 64 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	903	NAG	O5-C5-C6-O6
7	B	901	NAG	C4-C5-C6-O6
7	C	901	NAG	C8-C7-N2-C2
7	C	901	NAG	O7-C7-N2-C2
7	C	906	NAG	C8-C7-N2-C2

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	902	NAG	3	0
7	A	905	NAG	1	0
7	D	903	NAG	1	0
7	C	901	NAG	1	0
7	B	905	NAG	1	0
7	B	908	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 ⓘ

### 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	421/461 (91%)	-0.45	3 (0%) 87 86	20, 32, 55, 72	0
1	C	431/461 (93%)	-0.38	2 (0%) 91 89	17, 30, 59, 74	0
2	B	449/478 (93%)	-0.24	5 (1%) 80 78	23, 46, 74, 85	0
2	D	451/478 (94%)	-0.41	3 (0%) 87 86	17, 38, 69, 86	0
3	H	219/225 (97%)	-0.27	0 100 100	20, 36, 58, 76	0
3	J	218/225 (96%)	-0.43	1 (0%) 91 89	21, 32, 53, 76	0
4	K	216/217 (99%)	-0.49	0 100 100	20, 28, 45, 56	0
4	L	217/217 (100%)	-0.54	0 100 100	19, 30, 48, 73	0
All	All	2622/2762 (94%)	-0.39	14 (0%) 91 89	17, 34, 64, 86	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	25	THR	3.5
2	D	211	TRP	3.4
3	J	144	ASN	3.1
2	B	208	ASN	2.9
1	A	125	ASP	2.7

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

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

### 6.3 Carbohydrates ⓘ

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	NAG	F	2	14/15	0.85	0.28	34,55,61,61	0
5	BMA	E	3	11/12	0.85	0.22	89,92,94,96	0
5	NAG	E	2	14/15	0.88	0.20	49,73,78,87	0
5	NAG	E	1	14/15	0.95	0.12	28,45,49,58	0
6	NAG	F	1	14/15	0.96	0.16	33,43,48,54	0

## 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<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
7	NAG	A	902	14/15	0.66	0.47	87,106,111,112	0
7	NAG	A	905	14/15	0.69	0.22	73,86,93,94	0
7	NAG	B	901	14/15	0.76	0.40	108,122,127,128	0
7	NAG	A	903	14/15	0.77	0.26	79,92,103,103	0
7	NAG	B	908	14/15	0.77	0.24	68,93,103,107	0
7	NAG	C	906	14/15	0.78	0.32	82,89,91,92	0
7	NAG	C	903	14/15	0.82	0.33	74,81,87,92	0
7	NAG	B	906	14/15	0.84	0.33	74,91,96,101	0
7	NAG	B	905	14/15	0.85	0.21	74,88,97,100	0
7	NAG	A	907	14/15	0.86	0.19	72,84,93,96	0
7	NAG	D	905	14/15	0.87	0.33	74,92,99,105	0
7	NAG	B	907	14/15	0.88	0.22	55,67,81,82	0
7	NAG	D	904	14/15	0.88	0.20	93,104,110,115	0
9	NA	A	962	1/1	0.89	0.13	54,54,54,54	0
7	NAG	C	905	14/15	0.90	0.21	42,65,74,78	0
9	NA	C	962	1/1	0.91	0.14	35,35,35,35	0
7	NAG	D	903	14/15	0.92	0.17	46,52,64,70	0
7	NAG	C	902	14/15	0.92	0.18	43,60,66,72	0
7	NAG	A	906	14/15	0.93	0.14	42,47,53,55	0
7	NAG	A	904	14/15	0.93	0.22	53,66,84,85	0
7	NAG	C	904	14/15	0.93	0.14	42,46,51,56	0
9	NA	D	961	1/1	0.96	0.10	34,34,34,34	0
7	NAG	C	901	14/15	0.97	0.12	21,28,38,40	0
8	GLY	C	951	5/5	0.98	0.18	31,31,35,38	0
9	NA	C	961	1/1	0.98	0.10	37,37,37,37	0
8	GLY	A	951	5/5	0.98	0.15	26,26,27,30	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	NAG	A	901	14/15	0.98	0.12	18,29,36,37	0
10	CL	C	971	1/1	0.98	0.14	41,41,41,41	0
8	GLY	B	951	5/5	0.98	0.14	34,38,40,42	0
9	NA	A	961	1/1	0.98	0.08	35,35,35,35	0
10	CL	D	971	1/1	0.99	0.10	43,43,43,43	0
8	GLY	D	951	5/5	0.99	0.11	26,30,30,33	0
10	CL	B	971	1/1	0.99	0.11	46,46,46,46	0
11	CA	K	301	1/1	1.00	0.20	29,29,29,29	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.