



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 12:38 pm BST

PDB ID : 6J31  
Title : Crystal Structure Analysis of the Glycotransferase of kitacinnamycin  
Authors : Shi, J.; Liu, C.L.; Zhang, B.; Guo, W.J.; Zhu, J.P.; Xu, X.; Xu, Q.; Jiao, R.H.;  
Tan, R.X.; Ge, H.M.  
Deposited on : 2019-01-03  
Resolution : 2.24 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	:	2.11
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)	:	2.11

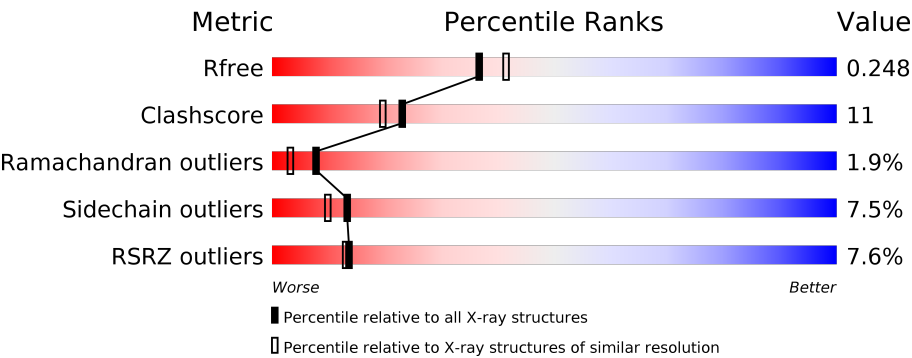
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.24 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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	396	<div><div>8%</div><div><div></div><div>71%</div><div>22%</div><div>• • •</div></div></div>
1	B	396	<div><div>8%</div><div><div></div><div>72%</div><div>21%</div><div>5%</div><div>• •</div></div></div>
1	C	396	<div><div>6%</div><div><div></div><div>77%</div><div>17%</div><div>• •</div></div></div>
1	D	396	<div><div>6%</div><div><div></div><div>77%</div><div>17%</div><div>• • •</div></div></div>
2	E	9	<div><div></div><div><div>44%</div><div>44%</div><div>11%</div></div></div>
2	F	9	<div><div>33%</div><div><div></div><div>44%</div><div>44%</div><div>11%</div></div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	9	 67% 33%
2	H	9	 78% 22%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12788 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.

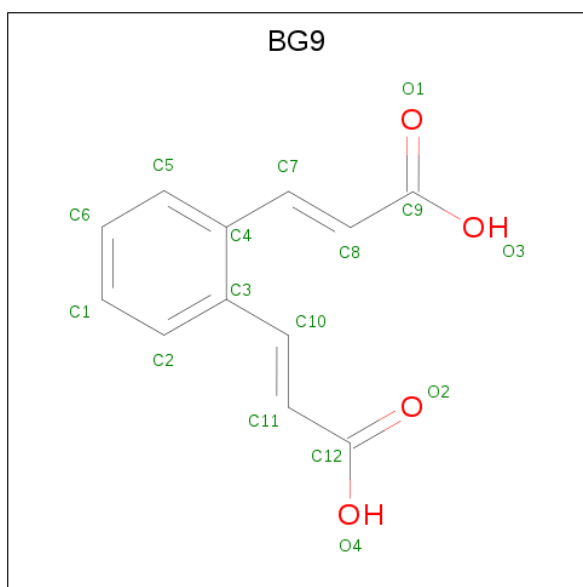
- Molecule 1 is a protein called kcn28.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	385	Total	C	N	O	S	0	0	0
			2970	1878	534	547	11			
1	B	387	Total	C	N	O	S	0	0	0
			2985	1888	536	549	12			
1	C	387	Total	C	N	O	S	0	0	0
			2972	1878	535	548	11			
1	D	387	Total	C	N	O	S	0	0	0
			2983	1886	536	550	11			

- Molecule 2 is a protein called DBB-DSG-VAL-MEA-VAL-GLY-GLY-DVA-DLE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	9	Total	C	N	O	0	0	0
			64	43	10	11			
2	F	9	Total	C	N	O	0	0	0
			64	43	10	11			
2	G	9	Total	C	N	O	0	0	0
			64	43	10	11			
2	H	9	Total	C	N	O	0	0	0
			64	43	10	11			

- Molecule 3 is (2E,2'E)-3,3'-(1,2-phenylene)di(prop-2-enoic acid) (three-letter code: BG9) (formula: C<sub>12</sub>H<sub>10</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	C	O	0	0
			14	12	2		
3	F	1	Total	C	O	0	0
			14	12	2		
3	G	1	Total	C	O	0	0
			14	12	2		
3	H	1	Total	C	O	0	0
			14	12	2		

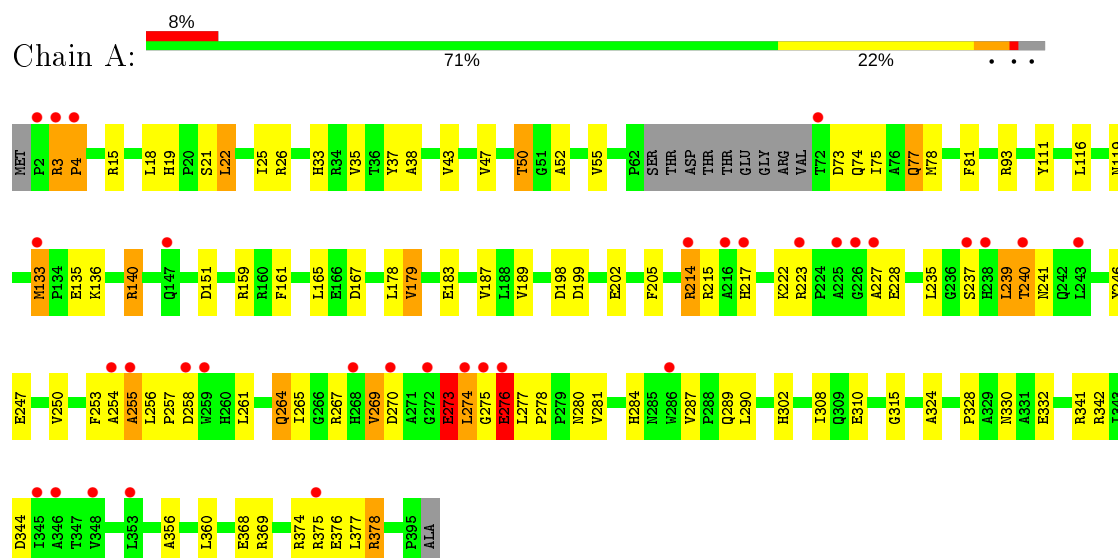
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	118	Total	O	0	0
			118	118		
4	B	140	Total	O	0	0
			140	140		
4	C	135	Total	O	0	0
			135	135		
4	D	160	Total	O	0	0
			160	160		
4	E	3	Total	O	0	0
			3	3		
4	G	3	Total	O	0	0
			3	3		
4	H	7	Total	O	0	0
			7	7		

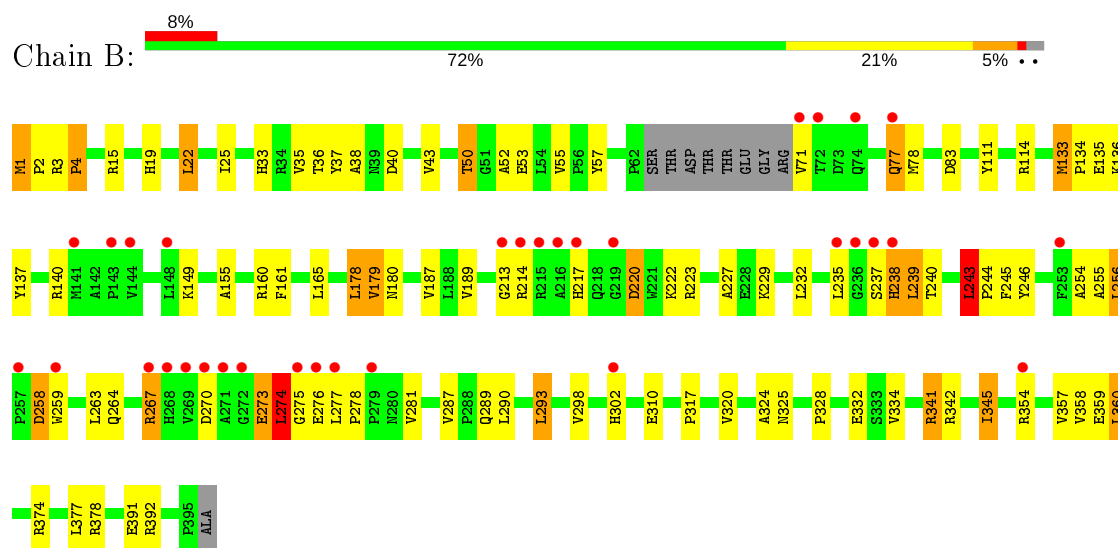
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

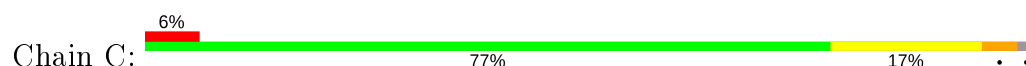
#### • Molecule 1: kcn28

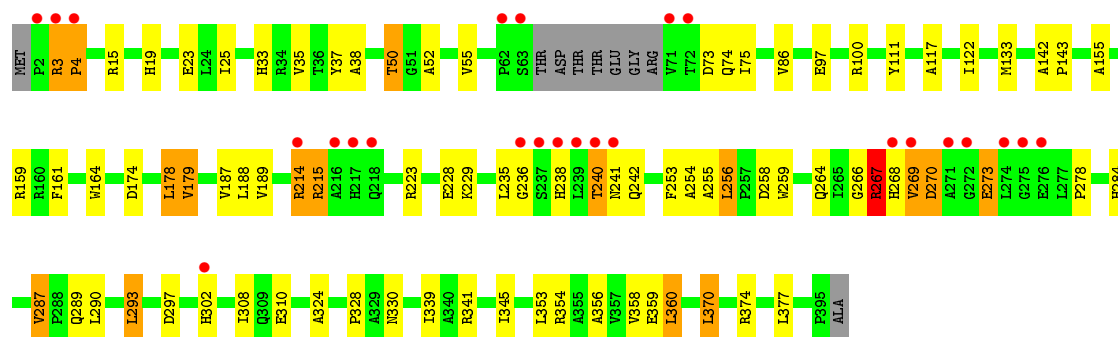


#### • Molecule 1: kcn28

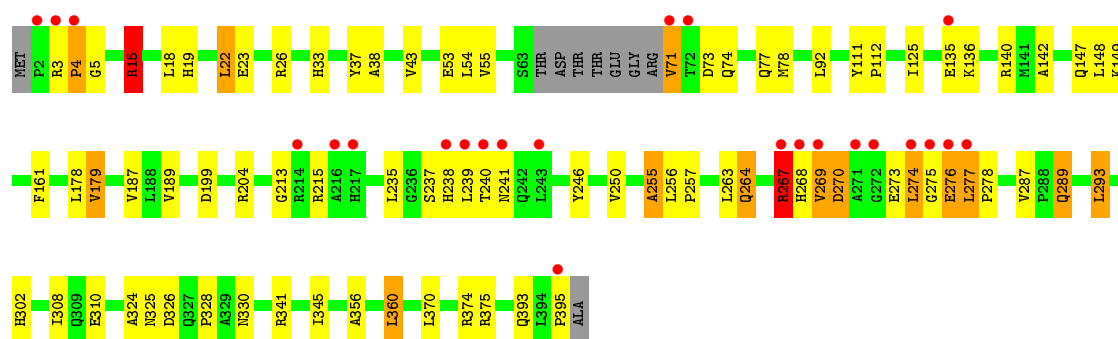


#### • Molecule 1: kcn28

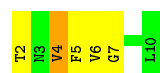




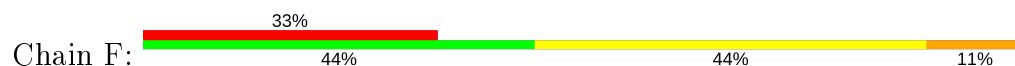
- Molecule 1: *kcn28*



- Molecule 2: DBB-DSG-VAL-MEA-VAL-GLY-GLY-DVA-DLE



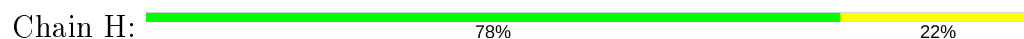
- Molecule 2: DBB-DSG-VAL-MEA-VAL-GLY-GLY-DVA-DLE

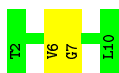


- Molecule 2: DBB-DSG-VAL-MEA-VAL-GLY-GLY-DVA-DLE



- Molecule 2: DBB-DSG-VAL-MEA-VAL-GLY-GLY-DVA-DLE







## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	243.16 Å   243.16 Å   243.16 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	39.45 – 2.24 39.45 – 2.24	Depositor EDS
% Data completeness (in resolution range)	100.0 (39.45-2.24) 100.0 (39.45-2.24)	Depositor EDS
$R_{merge}$	0.27	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.10 (at 2.24 Å)	Xtriage
Refinement program	PHENIX 1.14 _3260	Depositor
R, $R_{free}$	0.226 , 0.248 0.225 , 0.248	Depositor DCC
$R_{free}$ test set	5489 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.5	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 36.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.000 for -l,-k,-h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12788	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.92% 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: MEA, DLE, DBB, DVA, BG9, DSG

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	0.29	0/3042	0.53	0/4153
1	B	0.28	0/3057	0.56	0/4174
1	C	0.27	0/3044	0.51	0/4158
1	D	0.28	0/3055	0.59	2/4171 (0.0%)
2	E	3.11	2/20 (10.0%)	1.81	1/24 (4.2%)
2	F	3.14	2/20 (10.0%)	1.63	1/24 (4.2%)
2	G	3.10	2/20 (10.0%)	1.62	0/24
2	H	3.15	2/20 (10.0%)	1.71	0/24
All	All	0.37	8/12278 (0.1%)	0.56	4/16752 (0.0%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	6	VAL	C-N	9.01	1.49	1.33
2	F	6	VAL	C-N	8.99	1.49	1.33
2	G	6	VAL	C-N	8.90	1.49	1.33
2	E	6	VAL	C-N	8.83	1.49	1.33
2	F	7	GLY	C-N	8.18	1.47	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	267	ARG	NE-CZ-NH2	6.65	123.62	120.30
1	D	15	ARG	NE-CZ-NH2	-6.06	117.27	120.30
2	E	4	VAL	N-CA-C	-5.79	95.36	111.00
2	F	4	VAL	N-CA-C	-5.69	95.64	111.00

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	2970	0	2939	83	0
1	B	2985	0	2961	76	0
1	C	2972	0	2931	53	0
1	D	2983	0	2953	66	0
2	E	64	0	63	4	0
2	F	64	0	62	3	0
2	G	64	0	64	1	0
2	H	64	0	62	0	0
3	E	14	0	0	0	0
3	F	14	0	0	0	0
3	G	14	0	0	0	0
3	H	14	0	0	0	0
4	A	118	0	0	17	0
4	B	140	0	0	12	0
4	C	135	0	0	5	1
4	D	160	0	0	15	0
4	E	3	0	0	0	0
4	G	3	0	0	0	0
4	H	7	0	0	0	0
All	All	12788	0	12035	276	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 11.

The worst 5 of 276 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:223:ARG:NH1	1:A:227:ALA:O	1.92	1.01
1:A:214:ARG:HA	1:A:214:ARG:HH11	1.24	0.99
1:D:264:GLN:OE1	4:D:401:HOH:O	1.83	0.97
1:B:15:ARG:O	1:B:15:ARG:NH1	1.96	0.96
1:C:242:GLN:NE2	4:C:402:HOH:O	1.98	0.96

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)
4:C:487:HOH:O	4:C:527:HOH:O[6_566]	1.91	0.29

## 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	381/396 (96%)	356 (93%)	19 (5%)	6 (2%)	9	5
1	B	383/396 (97%)	359 (94%)	15 (4%)	9 (2%)	6	2
1	C	383/396 (97%)	363 (95%)	13 (3%)	7 (2%)	8	3
1	D	383/396 (97%)	358 (94%)	18 (5%)	7 (2%)	8	3
2	E	4/9 (44%)	3 (75%)	1 (25%)	0	100	100
2	F	4/9 (44%)	3 (75%)	1 (25%)	0	100	100
2	G	4/9 (44%)	3 (75%)	1 (25%)	0	100	100
2	H	4/9 (44%)	3 (75%)	1 (25%)	0	100	100
All	All	1546/1620 (95%)	1448 (94%)	69 (4%)	29 (2%)	8	3

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	179	VAL
1	A	276	GLU
1	B	4	PRO
1	B	179	VAL
1	B	239	LEU

### 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	312/321 (97%)	288 (92%)	24 (8%)	13	9
1	B	314/321 (98%)	288 (92%)	26 (8%)	11	7
1	C	311/321 (97%)	289 (93%)	22 (7%)	14	11
1	D	314/321 (98%)	291 (93%)	23 (7%)	14	10
2	E	2/2 (100%)	2 (100%)	0	100	100
2	F	2/2 (100%)	2 (100%)	0	100	100
2	G	2/2 (100%)	2 (100%)	0	100	100
2	H	2/2 (100%)	2 (100%)	0	100	100
All	All	1259/1292 (97%)	1164 (92%)	95 (8%)	13	10

5 of 95 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	273	GLU
1	C	161	PHE
1	D	274	LEU
1	B	293	LEU
1	B	360	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	119	ASN
1	B	119	ASN
1	C	268	HIS
1	D	289	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

20 non-standard protein/DNA/RNA residues 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	DBB	G	2	3,2	4,5,6	0.42	0	1,5,7	1.87	0
2	DBB	H	2	3,2	4,5,6	0.44	0	1,5,7	1.78	0
2	MEA	E	5	2	11,12,13	1.00	1 (9%)	13,14,16	1.14	1 (7%)
2	MEA	G	5	2	11,12,13	0.98	1 (9%)	13,14,16	1.17	1 (7%)
2	MEA	H	5	2	11,12,13	0.96	1 (9%)	13,14,16	1.06	1 (7%)
2	MEA	F	5	2	11,12,13	0.97	1 (9%)	13,14,16	1.24	2 (15%)
2	DBB	E	2	3,2	4,5,6	0.49	0	1,5,7	4.48	1 (100%)
2	DBB	F	2	3,2	4,5,6	0.54	0	1,5,7	3.21	1 (100%)

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	DBB	G	2	3,2	-	1/3/4/6	-
2	DBB	H	2	3,2	-	0/3/4/6	-
2	MEA	E	5	2	-	4/5/8/10	0/1/1/1
2	MEA	G	5	2	-	3/5/8/10	0/1/1/1
2	MEA	H	5	2	-	3/5/8/10	0/1/1/1
2	MEA	F	5	2	-	4/5/8/10	0/1/1/1
2	DBB	E	2	3,2	-	1/3/4/6	-
2	DBB	F	2	3,2	-	0/3/4/6	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	5	MEA	CB-CG	2.49	1.57	1.51
2	G	5	MEA	CB-CG	2.45	1.57	1.51
2	F	5	MEA	CB-CG	2.45	1.57	1.51
2	H	5	MEA	CB-CG	2.37	1.57	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	2	DBB	CG-CB-CA	-4.48	103.17	113.42
2	F	2	DBB	CG-CB-CA	-3.21	106.08	113.42
2	G	5	MEA	C1-N-CA	2.96	122.86	113.64
2	E	5	MEA	C1-N-CA	2.90	122.66	113.64
2	F	5	MEA	C1-N-CA	2.84	122.48	113.64

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	2	DBB	O-C-CA-CB
2	E	5	MEA	O-C-CA-CB
2	H	5	MEA	O-C-CA-CB
2	G	5	MEA	O-C-CA-CB
2	F	5	MEA	O-C-CA-CB

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	5	MEA	1	0
2	G	5	MEA	1	0
2	F	5	MEA	1	0
2	E	2	DBB	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

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
3	BG9	E	101	2	14,14,16	2.15	4 (28%)	16,16,20	1.46	3 (18%)
3	BG9	F	101	2	14,14,16	2.17	4 (28%)	16,16,20	1.70	3 (18%)
3	BG9	G	101	2	14,14,16	2.21	4 (28%)	16,16,20	1.70	4 (25%)
3	BG9	H	101	2	14,14,16	2.18	4 (28%)	16,16,20	1.76	3 (18%)

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
3	BG9	E	101	2	-	4/8/8/10	0/1/1/1
3	BG9	F	101	2	-	4/8/8/10	0/1/1/1
3	BG9	G	101	2	-	4/8/8/10	0/1/1/1
3	BG9	H	101	2	-	4/8/8/10	0/1/1/1

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	101	BG9	C4-C7	5.13	1.55	1.47
3	H	101	BG9	C4-C7	5.04	1.55	1.47
3	G	101	BG9	C3-C10	4.91	1.55	1.47
3	E	101	BG9	C4-C7	4.88	1.54	1.47
3	G	101	BG9	C4-C7	4.84	1.54	1.47

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	101	BG9	C3-C10-C11	-5.02	115.70	126.10
3	F	101	BG9	C3-C10-C11	-4.86	116.04	126.10
3	G	101	BG9	C3-C10-C11	-4.56	116.67	126.10
3	E	101	BG9	C3-C10-C11	-3.68	118.49	126.10
3	H	101	BG9	O2-C12-C11	-2.73	116.36	125.67

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	101	BG9	C10-C11-C12-O2
3	F	101	BG9	C10-C11-C12-O2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	G	101	BG9	C10-C11-C12-O2
3	H	101	BG9	C7-C8-C9-O3
3	H	101	BG9	C10-C11-C12-O2

There are no ring outliers.

No monomer is involved in short contacts.

## 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	385/396 (97%)	0.29	33 (8%) 10 10	26, 47, 106, 159	0
1	B	387/396 (97%)	0.29	33 (8%) 10 10	28, 50, 106, 175	0
1	C	387/396 (97%)	0.25	25 (6%) 18 18	29, 46, 112, 171	0
1	D	387/396 (97%)	0.11	24 (6%) 20 20	29, 45, 110, 170	0
2	E	4/9 (44%)	0.30	0 100 100	67, 69, 72, 72	0
2	F	4/9 (44%)	2.81	3 (75%) 0 0	93, 94, 97, 99	0
2	G	4/9 (44%)	0.70	0 100 100	63, 64, 69, 70	0
2	H	4/9 (44%)	-0.60	0 100 100	39, 40, 45, 49	0
All	All	1562/1620 (96%)	0.24	118 (7%) 13 13	26, 47, 110, 175	0

The worst 5 of 118 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	71	VAL	20.0
1	C	71	VAL	9.9
1	C	239	LEU	8.8
1	B	276	GLU	7.6
1	D	216	ALA	7.3

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

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
2	DSG	E	3	8/9	0.70	0.23	79,83,84,85	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	DLE	F	10	9/9	0.79	0.28	86,94,115,125	0
2	DVA	F	9	7/8	0.79	0.24	93,95,98,98	0
2	MEA	F	5	12/13	0.80	0.17	88,89,94,94	0
2	DSG	G	3	8/9	0.82	0.27	72,76,84,84	0
2	DBB	E	2	6/7	0.82	0.16	86,88,90,95	0
2	DSG	F	3	8/9	0.84	0.17	101,103,107,114	0
2	DBB	F	2	6/7	0.84	0.17	120,122,138,150	0
2	DLE	E	10	9/9	0.86	0.22	56,61,73,81	0
2	DSG	H	3	8/9	0.89	0.17	45,50,64,65	0
2	DLE	G	10	9/9	0.89	0.19	56,58,62,62	0
2	DVA	G	9	7/8	0.91	0.16	63,65,66,67	0
2	DVA	E	9	7/8	0.91	0.14	58,60,63,63	0
2	MEA	E	5	12/13	0.93	0.12	49,54,65,65	0
2	DBB	G	2	6/7	0.94	0.14	66,70,71,72	0
2	MEA	G	5	12/13	0.94	0.13	64,64,65,66	0
2	DBB	H	2	6/7	0.94	0.13	42,45,50,58	0
2	DLE	H	10	9/9	0.95	0.20	47,48,58,60	0
2	DVA	H	9	7/8	0.96	0.12	45,45,46,48	0
2	MEA	H	5	12/13	0.97	0.10	32,35,39,39	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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
3	BG9	F	101	14/16	0.76	0.25	98,100,117,117	0
3	BG9	G	101	14/16	0.82	0.20	73,79,82,83	0
3	BG9	E	101	14/16	0.83	0.28	93,94,95,95	0
3	BG9	H	101	14/16	0.93	0.12	38,39,40,41	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.