



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 25, 2020 – 03:23 PM BST

PDB ID : 3R8R  
Title : Transaldolase from Bacillus subtilis  
Authors : Schneider, G.; Sandalova, T.; Samland, A.  
Deposited on : 2011-03-24  
Resolution : 1.90 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  
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

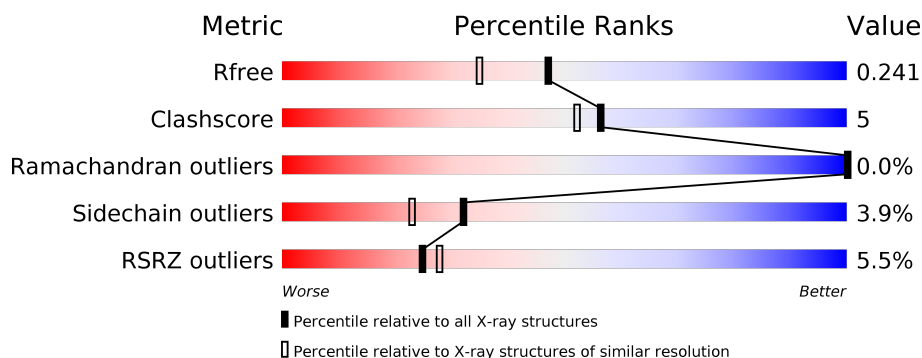
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	212	<div> <div>2%</div> <div> <div></div> <div>90%</div> <div>10%</div> </div> </div>
1	B	212	<div> <div>4%</div> <div> <div></div> <div>89%</div> <div>10%</div> <div>•</div> </div> </div>
1	C	212	<div> <div>12%</div> <div> <div></div> <div>86%</div> <div>10%</div> <div>•</div> </div> </div>
1	D	212	<div> <div>7%</div> <div> <div></div> <div>84%</div> <div>13%</div> <div>•</div> </div> </div>
1	E	212	<div> <div>3%</div> <div> <div></div> <div>88%</div> <div>11%</div> </div> </div>
1	F	212	<div> <div>5%</div> <div> <div></div> <div>89%</div> <div>9%</div> <div>•</div> </div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	212	
1	H	212	
1	I	212	
1	J	212	
1	K	212	
1	L	212	
1	M	212	
1	N	212	
1	P	212	
1	R	212	
1	T	212	
1	U	212	
1	V	212	
1	W	212	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 35462 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 Transaldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	B	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	C	204	Total	C	N	O	S	0	0	0
			1555	987	270	294	4			
1	D	207	Total	C	N	O	S	0	0	0
			1573	998	273	298	4			
1	E	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	F	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	G	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	H	211	Total	C	N	O	S	0	0	0
			1606	1018	279	305	4			
1	I	210	Total	C	N	O	S	0	0	0
			1598	1014	277	303	4			
1	J	206	Total	C	N	O	S	0	0	0
			1565	992	273	296	4			
1	V	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	K	211	Total	C	N	O	S	0	0	0
			1603	1017	278	304	4			
1	L	211	Total	C	N	O	S	0	0	0
			1603	1017	278	304	4			
1	M	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	N	202	Total	C	N	O	S	0	0	0
			1541	978	268	291	4			
1	W	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	P	202	Total	C	N	O	S	0	0	0
			1540	976	268	292	4			
1	R	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	T	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			
1	U	212	Total	C	N	O	S	0	0	0
			1611	1021	280	306	4			

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	E	1	Total	C	O	0	0
			6	3	3		
2	F	1	Total	C	O	0	0
			6	3	3		
2	G	1	Total	C	O	0	0
			6	3	3		
2	V	1	Total	C	O	0	0
			6	3	3		
2	L	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	W	1	Total	C	O	0	0
			6	3	3		
2	R	1	Total	C	O	0	0
			6	3	3		
2	U	1	Total	C	O	0	0
			6	3	3		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	G	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	I	1	Total	O	S	0	0
			5	4	1		
3	J	1	Total	O	S	0	0
			5	4	1		
3	V	1	Total	O	S	0	0
			5	4	1		
3	K	1	Total	O	S	0	0
			5	4	1		
3	L	1	Total	O	S	0	0
			5	4	1		
3	M	1	Total	O	S	0	0
			5	4	1		
3	N	1	Total	O	S	0	0
			5	4	1		
3	W	1	Total	O	S	0	0
			5	4	1		
3	P	1	Total	O	S	0	0
			5	4	1		
3	R	1	Total	O	S	0	0
			5	4	1		
3	T	1	Total	O	S	0	0
			5	4	1		
3	U	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	187	Total	O	0	0
			187	187		
4	B	203	Total	O	0	0
			203	203		
4	C	154	Total	O	0	0
			154	154		
4	D	176	Total	O	0	0
			176	176		
4	E	185	Total	O	0	0
			185	185		
4	F	179	Total	O	0	0
			179	179		
4	G	174	Total	O	0	0
			174	174		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	173	Total 173	O 173	0	0
4	I	155	Total 155	O 155	0	0
4	J	148	Total 148	O 148	0	0
4	V	173	Total 173	O 173	0	0
4	K	168	Total 168	O 168	0	0
4	L	166	Total 166	O 166	0	0
4	M	144	Total 144	O 144	0	0
4	N	175	Total 175	O 175	0	0
4	W	191	Total 191	O 191	0	0
4	P	166	Total 166	O 166	0	0
4	R	156	Total 156	O 156	0	0
4	T	147	Total 147	O 147	0	0
4	U	177	Total 177	O 177	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.

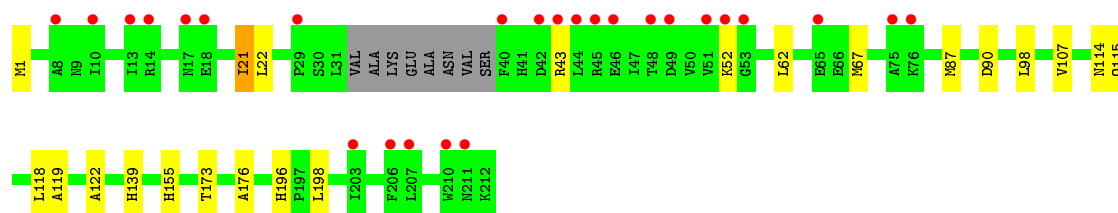
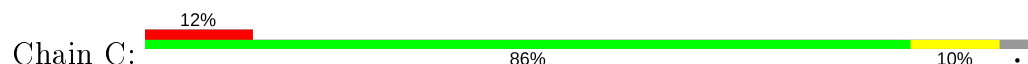
- Molecule 1: Transaldolase



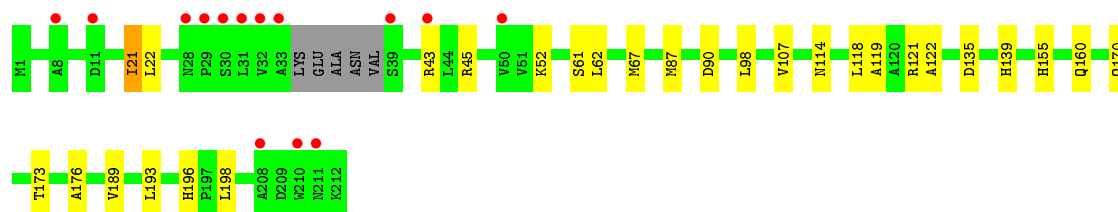
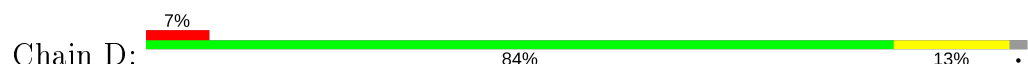
- Molecule 1: Transaldolase



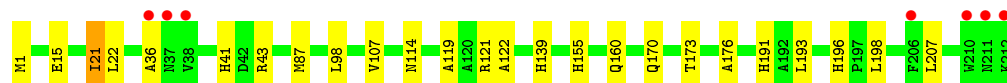
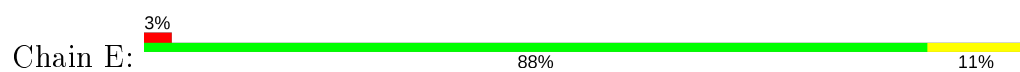
- Molecule 1: Transaldolase



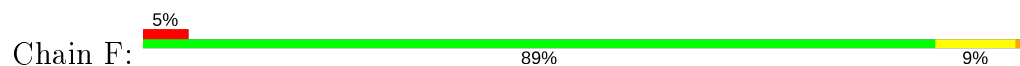
- Molecule 1: Transaldolase



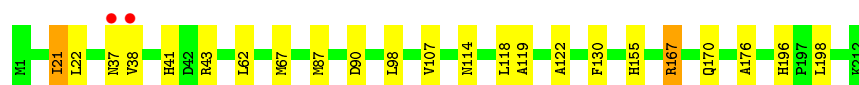
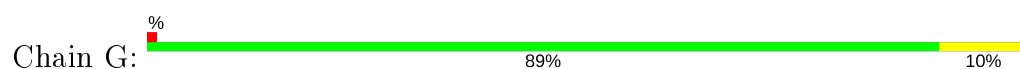
- Molecule 1: Transaldolase



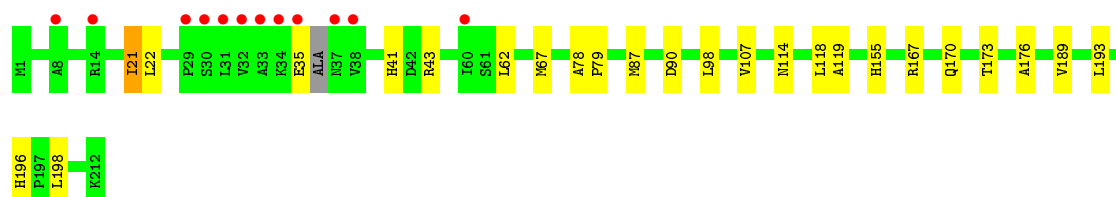
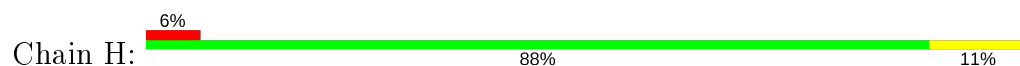
• Molecule 1: Transaldolase



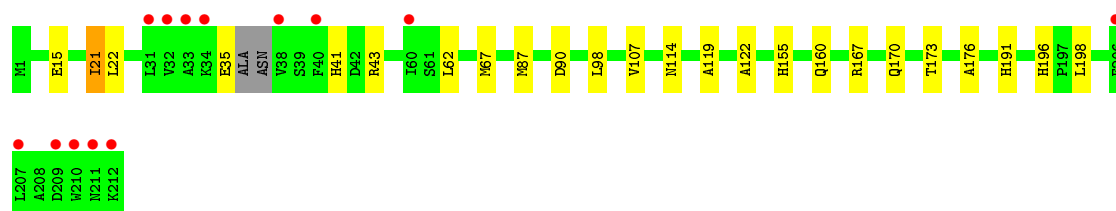
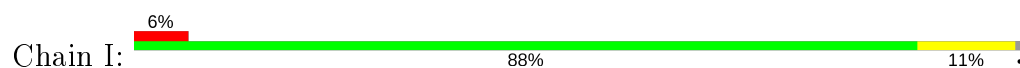
• Molecule 1: Transaldolase



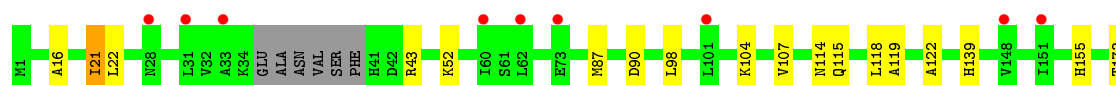
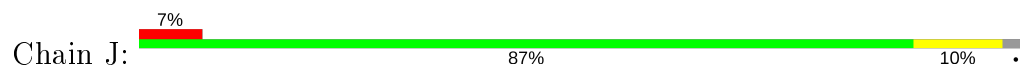
• Molecule 1: Transaldolase

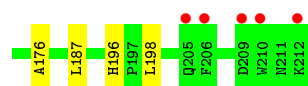


• Molecule 1: Transaldolase

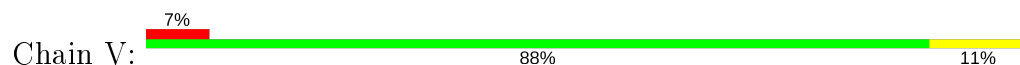


• Molecule 1: Transaldolase

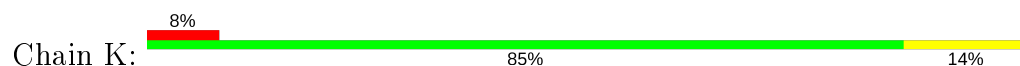




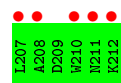
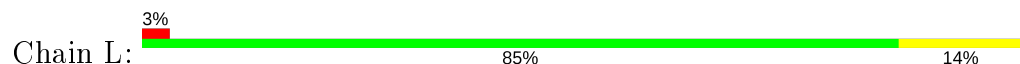
● Molecule 1: Transaldolase



● Molecule 1: Transaldolase



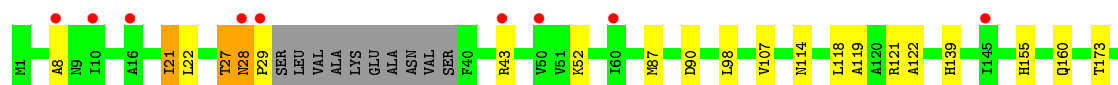
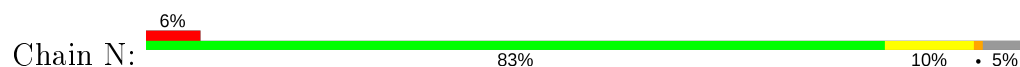
● Molecule 1: Transaldolase

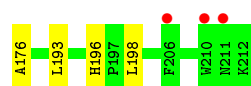


● Molecule 1: Transaldolase

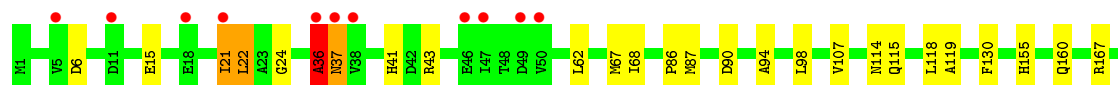
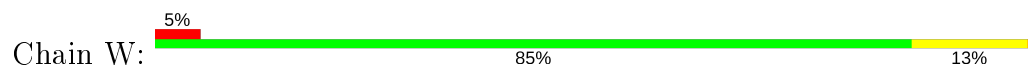


● Molecule 1: Transaldolase

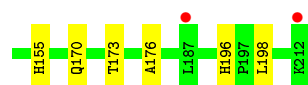
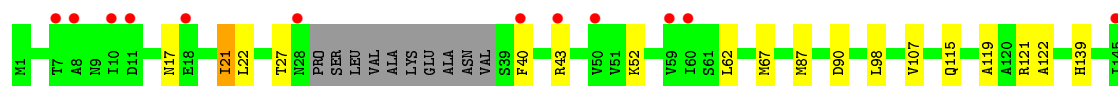
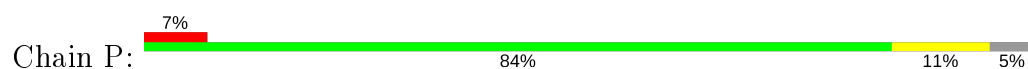




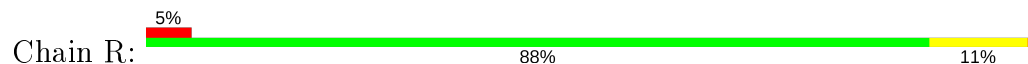
• Molecule 1: Transaldolase



• Molecule 1: Transaldolase



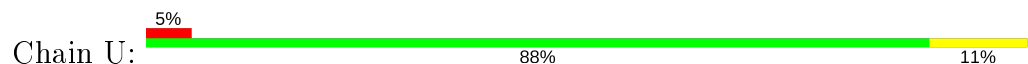
• Molecule 1: Transaldolase



• Molecule 1: Transaldolase



• Molecule 1: Transaldolase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	143.30Å 127.28Å 158.45Å 90.00° 98.10° 90.00°	Depositor
Resolution (Å)	58.93 – 1.90 58.07 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (58.93-1.90) 100.0 (58.07-1.90)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.11 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.200 , 0.229 0.212 , 0.241	Depositor DCC
$R_{free}$ test set	22163 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.4	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 52.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	35462	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 56.42 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.7436e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: GOL, SO4

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	0.48	0/1635	0.62	0/2218
1	B	0.49	0/1635	0.61	0/2218
1	C	0.52	0/1578	0.61	0/2139
1	D	0.47	0/1596	0.59	0/2164
1	E	0.46	0/1635	0.65	1/2218 (0.0%)
1	F	0.47	0/1635	0.74	4/2218 (0.2%)
1	G	0.46	0/1635	0.63	1/2218 (0.0%)
1	H	0.49	0/1629	0.61	0/2208
1	I	0.44	0/1621	0.59	0/2197
1	J	0.42	0/1587	0.57	0/2151
1	K	0.49	0/1626	0.59	0/2204
1	L	0.46	0/1626	0.61	0/2204
1	M	0.45	0/1635	0.58	0/2218
1	N	0.49	0/1564	0.63	1/2120 (0.0%)
1	P	0.47	0/1562	0.57	0/2116
1	R	0.43	0/1635	0.59	1/2218 (0.0%)
1	T	0.49	0/1635	0.58	0/2218
1	U	0.52	0/1635	0.62	0/2218
1	V	0.51	0/1635	0.61	0/2218
1	W	0.54	0/1635	0.82	4/2218 (0.2%)
All	All	0.48	0/32374	0.63	12/43901 (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
1	W	0	1

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	W	36	ALA	CB-CA-C	-19.59	80.72	110.10
1	F	36	ALA	CB-CA-C	-16.34	85.59	110.10
1	W	36	ALA	N-CA-C	12.08	143.61	111.00
1	N	27	THR	CB-CA-C	-10.60	82.99	111.60
1	W	36	ALA	C-N-CA	9.37	145.13	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	W	36	ALA	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	1611	0	1653	21	0
1	B	1611	0	1653	21	0
1	C	1555	0	1594	19	0
1	D	1573	0	1613	29	0
1	E	1611	0	1653	22	0
1	F	1611	0	1653	22	0
1	G	1611	0	1653	19	0
1	H	1606	0	1647	22	0
1	I	1598	0	1641	16	0
1	J	1565	0	1612	17	0
1	K	1603	0	1646	25	0
1	L	1603	0	1646	25	0
1	M	1611	0	1653	20	0
1	N	1541	0	1578	22	0
1	P	1540	0	1576	20	0
1	R	1611	0	1653	17	0
1	T	1611	0	1653	17	0
1	U	1611	0	1653	23	0
1	V	1611	0	1653	22	0
1	W	1611	0	1653	24	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	6	0	8	1	0
2	B	6	0	8	1	0
2	E	6	0	8	1	0
2	F	6	0	8	2	0
2	G	6	0	8	1	0
2	L	6	0	8	0	0
2	R	6	0	8	0	0
2	U	6	0	8	2	0
2	V	6	0	8	1	0
2	W	6	0	8	2	0
3	A	5	0	0	0	0
3	B	5	0	0	1	0
3	C	5	0	0	0	0
3	D	5	0	0	0	0
3	E	5	0	0	1	0
3	F	5	0	0	0	0
3	G	5	0	0	0	0
3	H	5	0	0	0	0
3	I	5	0	0	0	0
3	J	5	0	0	0	0
3	K	5	0	0	0	0
3	L	5	0	0	0	0
3	M	5	0	0	0	0
3	N	5	0	0	0	0
3	P	5	0	0	0	0
3	R	5	0	0	0	0
3	T	5	0	0	0	0
3	U	5	0	0	0	0
3	V	5	0	0	0	0
3	W	5	0	0	0	0
4	A	187	0	0	2	0
4	B	203	0	0	1	0
4	C	154	0	0	1	0
4	D	176	0	0	6	0
4	E	185	0	0	5	0
4	F	179	0	0	3	0
4	G	174	0	0	3	0
4	H	173	0	0	4	0
4	I	155	0	0	3	0
4	J	148	0	0	1	0
4	K	168	0	0	2	0
4	L	166	0	0	2	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	M	144	0	0	5	0
4	N	175	0	0	1	0
4	P	166	0	0	2	0
4	R	156	0	0	2	0
4	T	147	0	0	1	0
4	U	177	0	0	3	0
4	V	173	0	0	3	0
4	W	191	0	0	5	0
All	All	35462	0	32816	322	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 5.

The worst 5 of 322 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:V:87:MET:HE1	1:V:107:VAL:HG11	1.37	1.06
1:F:87:MET:HE1	1:F:107:VAL:HG11	1.46	0.97
1:V:87:MET:CE	1:V:107:VAL:HG11	1.94	0.97
1:M:87:MET:HE1	1:M:119:ALA:HB2	1.44	0.97
1:E:87:MET:HE1	1:E:107:VAL:HG11	1.48	0.92

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
1	B	210/212 (99%)	204 (97%)	6 (3%)	0	100	100
1	C	200/212 (94%)	195 (98%)	5 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	203/212 (96%)	200 (98%)	3 (2%)	0	100	100
1	E	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
1	F	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
1	G	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
1	H	207/212 (98%)	202 (98%)	5 (2%)	0	100	100
1	I	206/212 (97%)	204 (99%)	2 (1%)	0	100	100
1	J	202/212 (95%)	199 (98%)	3 (2%)	0	100	100
1	K	207/212 (98%)	206 (100%)	1 (0%)	0	100	100
1	L	207/212 (98%)	204 (99%)	3 (1%)	0	100	100
1	M	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
1	N	198/212 (93%)	195 (98%)	2 (1%)	1 (0%)	29	18
1	P	198/212 (93%)	196 (99%)	2 (1%)	0	100	100
1	R	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
1	T	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
1	U	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
1	V	210/212 (99%)	208 (99%)	2 (1%)	0	100	100
1	W	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
All	All	4138/4240 (98%)	4073 (98%)	64 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	N	28	ASN

### 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	173/174 (99%)	167 (96%)	6 (4%)	36	27

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	173/174 (99%)	167 (96%)	6 (4%)	36	27
1	C	167/174 (96%)	161 (96%)	6 (4%)	35	26
1	D	169/174 (97%)	163 (96%)	6 (4%)	35	26
1	E	173/174 (99%)	168 (97%)	5 (3%)	42	35
1	F	173/174 (99%)	165 (95%)	8 (5%)	27	17
1	G	173/174 (99%)	166 (96%)	7 (4%)	31	22
1	H	173/174 (99%)	165 (95%)	8 (5%)	27	17
1	I	172/174 (99%)	164 (95%)	8 (5%)	26	16
1	J	168/174 (97%)	162 (96%)	6 (4%)	35	26
1	K	172/174 (99%)	164 (95%)	8 (5%)	26	16
1	L	172/174 (99%)	165 (96%)	7 (4%)	30	21
1	M	173/174 (99%)	168 (97%)	5 (3%)	42	35
1	N	165/174 (95%)	159 (96%)	6 (4%)	35	26
1	P	165/174 (95%)	157 (95%)	8 (5%)	25	16
1	R	173/174 (99%)	165 (95%)	8 (5%)	27	17
1	T	173/174 (99%)	167 (96%)	6 (4%)	36	27
1	U	173/174 (99%)	166 (96%)	7 (4%)	31	22
1	V	173/174 (99%)	166 (96%)	7 (4%)	31	22
1	W	173/174 (99%)	167 (96%)	6 (4%)	36	27
All	All	3426/3480 (98%)	3292 (96%)	134 (4%)	32	23

5 of 134 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	21	ILE
1	K	35	GLU
1	T	35	GLU
1	J	43	ARG
1	V	41	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 145 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	160	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	K	191	HIS
1	T	160	GLN
1	V	41	HIS
1	V	170	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](#)

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](#)

30 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	GOL	W	221	-	5,5,5	0.52	0	5,5,5	0.83	0
2	GOL	U	221	-	5,5,5	0.47	0	5,5,5	0.69	0
2	GOL	F	221	-	5,5,5	0.39	0	5,5,5	0.42	0
3	SO4	D	301	-	4,4,4	0.22	0	6,6,6	0.34	0
2	GOL	L	221	-	5,5,5	0.47	0	5,5,5	0.36	0
2	GOL	G	221	-	5,5,5	0.44	0	5,5,5	0.68	0
2	GOL	E	221	-	5,5,5	0.41	0	5,5,5	0.78	0
3	SO4	H	301	-	4,4,4	0.20	0	6,6,6	0.24	0
3	SO4	G	301	-	4,4,4	0.20	0	6,6,6	0.20	0
3	SO4	V	301	-	4,4,4	0.14	0	6,6,6	0.45	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GOL	B	221	-	5,5,5	0.44	0	5,5,5	0.55	0
3	SO4	C	301	-	4,4,4	0.18	0	6,6,6	0.24	0
3	SO4	E	301	-	4,4,4	0.18	0	6,6,6	0.32	0
3	SO4	A	301	-	4,4,4	0.18	0	6,6,6	0.28	0
3	SO4	L	301	-	4,4,4	0.18	0	6,6,6	0.25	0
2	GOL	A	221	-	5,5,5	0.58	0	5,5,5	0.85	0
3	SO4	I	301	-	4,4,4	0.17	0	6,6,6	0.21	0
3	SO4	T	301	-	4,4,4	0.14	0	6,6,6	0.15	0
3	SO4	P	301	-	4,4,4	0.17	0	6,6,6	0.24	0
3	SO4	R	301	-	4,4,4	0.17	0	6,6,6	0.28	0
3	SO4	W	301	-	4,4,4	0.17	0	6,6,6	0.16	0
3	SO4	B	301	-	4,4,4	0.17	0	6,6,6	0.39	0
3	SO4	N	301	-	4,4,4	0.12	0	6,6,6	0.26	0
3	SO4	M	301	-	4,4,4	0.15	0	6,6,6	0.20	0
2	GOL	R	221	-	5,5,5	0.41	0	5,5,5	0.43	0
3	SO4	U	301	-	4,4,4	0.18	0	6,6,6	0.25	0
3	SO4	F	301	-	4,4,4	0.14	0	6,6,6	0.28	0
3	SO4	K	301	-	4,4,4	0.16	0	6,6,6	0.26	0
2	GOL	V	221	-	5,5,5	0.61	0	5,5,5	0.78	0
3	SO4	J	301	-	4,4,4	0.21	0	6,6,6	0.20	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	W	221	-	-	2/4/4/4	-
2	GOL	U	221	-	-	1/4/4/4	-
2	GOL	F	221	-	-	0/4/4/4	-
2	GOL	B	221	-	-	0/4/4/4	-
2	GOL	L	221	-	-	0/4/4/4	-
2	GOL	G	221	-	-	0/4/4/4	-
2	GOL	E	221	-	-	0/4/4/4	-
2	GOL	V	221	-	-	4/4/4/4	-
2	GOL	A	221	-	-	2/4/4/4	-
2	GOL	R	221	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	221	GOL	O1-C1-C2-C3
2	R	221	GOL	O1-C1-C2-C3
2	R	221	GOL	C1-C2-C3-O3
2	V	221	GOL	O1-C1-C2-O2
2	V	221	GOL	O1-C1-C2-C3

There are no ring outliers.

10 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	W	221	GOL	2	0
2	U	221	GOL	2	0
2	F	221	GOL	2	0
2	G	221	GOL	1	0
2	E	221	GOL	1	0
2	B	221	GOL	1	0
3	E	301	SO4	1	0
2	A	221	GOL	1	0
3	B	301	SO4	1	0
2	V	221	GOL	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<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	212/212 (100%)	0.26	4 (1%) 66 69	12, 21, 49, 62	0
1	B	212/212 (100%)	0.24	8 (3%) 40 43	14, 24, 50, 60	0
1	C	204/212 (96%)	0.41	26 (12%) 3 4	8, 16, 44, 50	0
1	D	207/212 (97%)	0.43	14 (6%) 17 19	11, 21, 48, 69	0
1	E	212/212 (100%)	0.09	7 (3%) 46 49	12, 21, 49, 62	0
1	F	212/212 (100%)	0.14	10 (4%) 31 34	12, 21, 50, 70	0
1	G	212/212 (100%)	0.02	2 (0%) 84 85	11, 22, 50, 65	0
1	H	211/212 (99%)	0.31	12 (5%) 23 26	12, 21, 47, 59	0
1	I	210/212 (99%)	0.33	13 (6%) 20 23	15, 29, 56, 68	0
1	J	206/212 (97%)	0.61	14 (6%) 17 19	14, 31, 64, 81	0
1	K	211/212 (99%)	0.32	17 (8%) 12 13	14, 24, 52, 63	0
1	L	211/212 (99%)	0.07	7 (3%) 46 49	14, 23, 49, 66	0
1	M	212/212 (100%)	0.23	9 (4%) 36 39	13, 25, 55, 70	0
1	N	202/212 (95%)	0.31	12 (5%) 22 25	13, 22, 48, 68	0
1	P	202/212 (95%)	0.43	14 (6%) 16 19	15, 25, 51, 64	0
1	R	212/212 (100%)	0.26	11 (5%) 27 30	16, 27, 56, 69	0
1	T	212/212 (100%)	0.36	16 (7%) 14 15	14, 25, 55, 70	0
1	U	212/212 (100%)	0.19	11 (5%) 27 30	13, 20, 48, 62	0
1	V	212/212 (100%)	0.26	14 (6%) 18 20	13, 22, 51, 65	0
1	W	212/212 (100%)	0.19	11 (5%) 27 30	13, 22, 49, 63	0
All	All	4196/4240 (98%)	0.27	232 (5%) 25 28	8, 23, 52, 81	0

The worst 5 of 232 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	W	36	ALA	14.8
1	R	210	TRP	13.2
1	R	206	PHE	9.7
1	H	38	VAL	9.5
1	V	210	TRP	8.6

## 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 ⓘ

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<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	GOL	F	221	6/6	0.65	0.23	34,39,41,41	0
2	GOL	A	221	6/6	0.73	0.24	31,33,34,35	0
2	GOL	U	221	6/6	0.74	0.20	24,29,30,32	0
2	GOL	B	221	6/6	0.76	0.14	38,39,39,40	0
2	GOL	W	221	6/6	0.77	0.21	35,36,37,39	0
2	GOL	V	221	6/6	0.77	0.28	34,35,36,36	0
2	GOL	R	221	6/6	0.81	0.14	36,40,40,41	0
2	GOL	G	221	6/6	0.83	0.21	33,35,36,39	0
2	GOL	E	221	6/6	0.93	0.12	24,28,29,30	0
2	GOL	L	221	6/6	0.94	0.11	20,25,26,27	0
3	SO4	R	301	5/5	0.96	0.10	39,39,41,41	0
3	SO4	B	301	5/5	0.97	0.10	38,39,39,40	0
3	SO4	N	301	5/5	0.97	0.09	34,35,36,37	0
3	SO4	I	301	5/5	0.97	0.10	46,46,47,47	0
3	SO4	C	301	5/5	0.97	0.10	40,41,42,43	0
3	SO4	J	301	5/5	0.97	0.11	42,43,43,43	0
3	SO4	D	301	5/5	0.98	0.08	31,32,33,33	0
3	SO4	V	301	5/5	0.98	0.09	35,36,36,38	0
3	SO4	M	301	5/5	0.98	0.09	35,35,36,36	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	T	301	5/5	0.98	0.09	36,36,36,37	0
3	SO4	U	301	5/5	0.98	0.08	29,30,31,31	0
3	SO4	F	301	5/5	0.98	0.11	35,36,36,37	0
3	SO4	K	301	5/5	0.98	0.08	35,36,36,36	0
3	SO4	P	301	5/5	0.98	0.07	33,33,34,34	0
3	SO4	E	301	5/5	0.98	0.10	33,34,35,36	0
3	SO4	W	301	5/5	0.99	0.07	36,38,38,38	0
3	SO4	G	301	5/5	0.99	0.11	36,36,38,38	0
3	SO4	A	301	5/5	0.99	0.09	35,35,36,36	0
3	SO4	L	301	5/5	0.99	0.08	35,35,36,37	0
3	SO4	H	301	5/5	0.99	0.08	32,33,33,34	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.