



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 10:54 am BST

PDB ID : 4R9K  
Title : Structure of thermostable eightfold mutant of limonene epoxide hydrolase from *Rhodococcus erythropolis*  
Authors : Floor, R.J.; Wijma, H.J.; Jekel, P.A.; Terwisscha van Scheltinga, A.C.; Dijkstra, B.W.; Janssen, D.B.  
Deposited on : 2014-09-05  
Resolution : 1.50 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](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

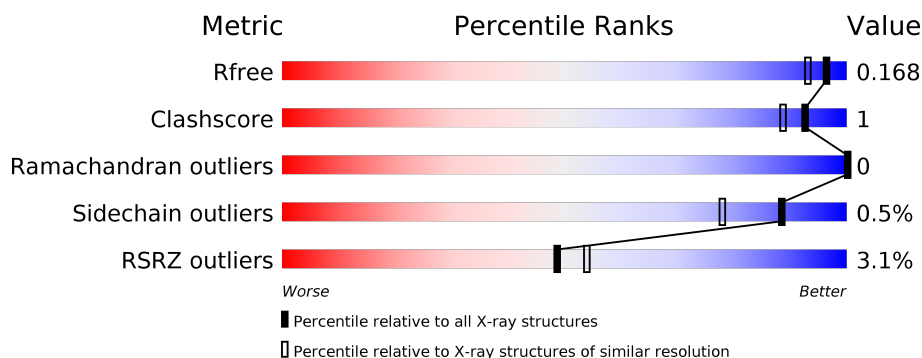
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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	174	<div> <div>2%</div> <div> <div></div> <div>74%</div> <div>9%</div> <div>17%</div> </div> </div>
1	B	174	<div> <div>4%</div> <div> <div></div> <div>83%</div> <div>6%</div> <div>11%</div> </div> </div>
1	C	174	<div> <div>2%</div> <div> <div></div> <div>81%</div> <div>10%</div> <div>8%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4037 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 Limonene-1,2-epoxide hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	144	Total	C	N	O	S	0	2	0
			1140	733	185	218	4			
1	B	155	Total	C	N	O	S	0	1	0
			1218	783	199	232	4			
1	C	160	Total	C	N	O	S	0	1	0
			1260	807	205	244	4			

There are 105 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP Q9ZAG3
A	2	ALA	-	EXPRESSION TAG	UNP Q9ZAG3
A	15	PRO	SER	ENGINEERED MUTATION	UNP Q9ZAG3
A	19	LYS	ALA	ENGINEERED MUTATION	UNP Q9ZAG3
A	45	LYS	GLU	ENGINEERED MUTATION	UNP Q9ZAG3
A	76	LYS	THR	ENGINEERED MUTATION	UNP Q9ZAG3
A	85	VAL	THR	ENGINEERED MUTATION	UNP Q9ZAG3
A	92	LYS	ASN	ENGINEERED MUTATION	UNP Q9ZAG3
A	96	PHE	TYR	ENGINEERED MUTATION	UNP Q9ZAG3
A	124	ASP	GLU	ENGINEERED MUTATION	UNP Q9ZAG3
A	150	LYS	-	EXPRESSION TAG	UNP Q9ZAG3
A	151	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
A	152	GLY	-	EXPRESSION TAG	UNP Q9ZAG3
A	153	PRO	-	EXPRESSION TAG	UNP Q9ZAG3
A	154	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
A	155	GLN	-	EXPRESSION TAG	UNP Q9ZAG3
A	156	LYS	-	EXPRESSION TAG	UNP Q9ZAG3
A	157	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
A	158	ILE	-	EXPRESSION TAG	UNP Q9ZAG3
A	159	SER	-	EXPRESSION TAG	UNP Q9ZAG3
A	160	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
A	161	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
A	162	ASP	-	EXPRESSION TAG	UNP Q9ZAG3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	163	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
A	164	ASN	-	EXPRESSION TAG	UNP Q9ZAG3
A	165	SER	-	EXPRESSION TAG	UNP Q9ZAG3
A	166	ALA	-	EXPRESSION TAG	UNP Q9ZAG3
A	167	VAL	-	EXPRESSION TAG	UNP Q9ZAG3
A	168	ASP	-	EXPRESSION TAG	UNP Q9ZAG3
A	169	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
A	170	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
A	171	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
A	172	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
A	173	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
A	174	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
B	1	MET	-	EXPRESSION TAG	UNP Q9ZAG3
B	2	ALA	-	EXPRESSION TAG	UNP Q9ZAG3
B	15	PRO	SER	ENGINEERED MUTATION	UNP Q9ZAG3
B	19	LYS	ALA	ENGINEERED MUTATION	UNP Q9ZAG3
B	45	LYS	GLU	ENGINEERED MUTATION	UNP Q9ZAG3
B	76	LYS	THR	ENGINEERED MUTATION	UNP Q9ZAG3
B	85	VAL	THR	ENGINEERED MUTATION	UNP Q9ZAG3
B	92	LYS	ASN	ENGINEERED MUTATION	UNP Q9ZAG3
B	96	PHE	TYR	ENGINEERED MUTATION	UNP Q9ZAG3
B	124	ASP	GLU	ENGINEERED MUTATION	UNP Q9ZAG3
B	150	LYS	-	EXPRESSION TAG	UNP Q9ZAG3
B	151	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
B	152	GLY	-	EXPRESSION TAG	UNP Q9ZAG3
B	153	PRO	-	EXPRESSION TAG	UNP Q9ZAG3
B	154	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
B	155	GLN	-	EXPRESSION TAG	UNP Q9ZAG3
B	156	LYS	-	EXPRESSION TAG	UNP Q9ZAG3
B	157	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
B	158	ILE	-	EXPRESSION TAG	UNP Q9ZAG3
B	159	SER	-	EXPRESSION TAG	UNP Q9ZAG3
B	160	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
B	161	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
B	162	ASP	-	EXPRESSION TAG	UNP Q9ZAG3
B	163	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
B	164	ASN	-	EXPRESSION TAG	UNP Q9ZAG3
B	165	SER	-	EXPRESSION TAG	UNP Q9ZAG3
B	166	ALA	-	EXPRESSION TAG	UNP Q9ZAG3
B	167	VAL	-	EXPRESSION TAG	UNP Q9ZAG3
B	168	ASP	-	EXPRESSION TAG	UNP Q9ZAG3
B	169	HIS	-	EXPRESSION TAG	UNP Q9ZAG3

*Continued on next page...*

*Continued from previous page...*

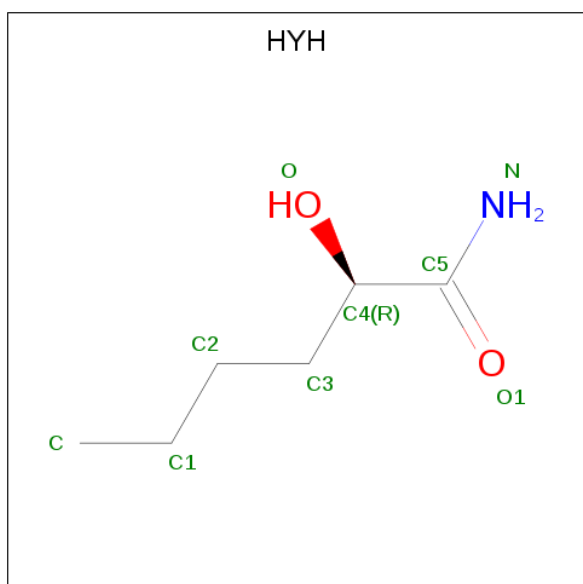
Chain	Residue	Modelled	Actual	Comment	Reference
B	170	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
B	171	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
B	172	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
B	173	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
B	174	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
C	1	MET	-	EXPRESSION TAG	UNP Q9ZAG3
C	2	ALA	-	EXPRESSION TAG	UNP Q9ZAG3
C	15	PRO	SER	ENGINEERED MUTATION	UNP Q9ZAG3
C	19	LYS	ALA	ENGINEERED MUTATION	UNP Q9ZAG3
C	45	LYS	GLU	ENGINEERED MUTATION	UNP Q9ZAG3
C	76	LYS	THR	ENGINEERED MUTATION	UNP Q9ZAG3
C	85	VAL	THR	ENGINEERED MUTATION	UNP Q9ZAG3
C	92	LYS	ASN	ENGINEERED MUTATION	UNP Q9ZAG3
C	96	PHE	TYR	ENGINEERED MUTATION	UNP Q9ZAG3
C	124	ASP	GLU	ENGINEERED MUTATION	UNP Q9ZAG3
C	150	LYS	-	EXPRESSION TAG	UNP Q9ZAG3
C	151	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
C	152	GLY	-	EXPRESSION TAG	UNP Q9ZAG3
C	153	PRO	-	EXPRESSION TAG	UNP Q9ZAG3
C	154	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
C	155	GLN	-	EXPRESSION TAG	UNP Q9ZAG3
C	156	LYS	-	EXPRESSION TAG	UNP Q9ZAG3
C	157	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
C	158	ILE	-	EXPRESSION TAG	UNP Q9ZAG3
C	159	SER	-	EXPRESSION TAG	UNP Q9ZAG3
C	160	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
C	161	GLU	-	EXPRESSION TAG	UNP Q9ZAG3
C	162	ASP	-	EXPRESSION TAG	UNP Q9ZAG3
C	163	LEU	-	EXPRESSION TAG	UNP Q9ZAG3
C	164	ASN	-	EXPRESSION TAG	UNP Q9ZAG3
C	165	SER	-	EXPRESSION TAG	UNP Q9ZAG3
C	166	ALA	-	EXPRESSION TAG	UNP Q9ZAG3
C	167	VAL	-	EXPRESSION TAG	UNP Q9ZAG3
C	168	ASP	-	EXPRESSION TAG	UNP Q9ZAG3
C	169	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
C	170	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
C	171	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
C	172	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
C	173	HIS	-	EXPRESSION TAG	UNP Q9ZAG3
C	174	HIS	-	EXPRESSION TAG	UNP Q9ZAG3

- 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	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is (2R)-2-hydroxyhexanamide (three-letter code: HYH) (formula:  $C_6H_{13}NO_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			9	6	1	2		
3	B	1	Total	C	N	O	0	0
			9	6	1	2		
3	C	1	Total	C	N	O	0	0
			9	6	1	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	103	Total	O	0	0
			103	103		
4	B	114	Total	O	0	0
			114	114		
4	C	145	Total	O	0	0
			145	145		

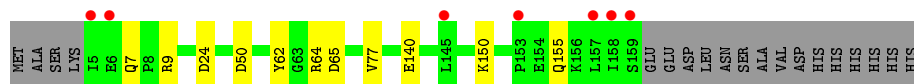
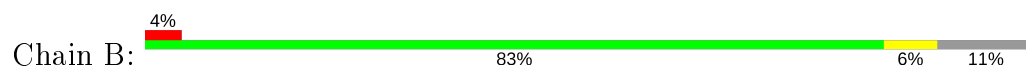
### 3 Residue-property plots [i](#)

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.

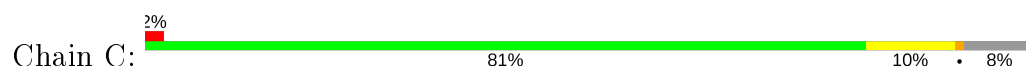
- Molecule 1: Limonene-1,2-epoxide hydrolase



- Molecule 1: Limonene-1,2-epoxide hydrolase



- Molecule 1: Limonene-1,2-epoxide hydrolase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.05Å 88.05Å 110.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.00 – 1.50 44.02 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.00-1.50) 99.9 (44.02-1.50)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.02	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.18 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, $R_{free}$	0.141 , 0.165 0.147 , 0.168	Depositor DCC
$R_{free}$ test set	4008 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.0	Xtriage
Anisotropy	0.021	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 46.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	4037	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.36% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HYH, GOL

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	1.13	4/1171 (0.3%)	1.15	7/1587 (0.4%)
1	B	1.16	1/1247 (0.1%)	1.16	6/1688 (0.4%)
1	C	1.15	2/1289 (0.2%)	1.23	13/1745 (0.7%)
All	All	1.15	7/3707 (0.2%)	1.18	26/5020 (0.5%)

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	21	SER	CB-OG	-5.79	1.34	1.42
1	B	140	GLU	CD-OE2	-5.76	1.19	1.25
1	A	77	VAL	CB-CG2	-5.54	1.41	1.52
1	A	21	SER	CB-OG	-5.29	1.35	1.42
1	A	9	ARG	CZ-NH2	5.14	1.39	1.33
1	A	86	PHE	CB-CG	-5.07	1.42	1.51
1	C	156	LYS	C-N	-5.02	1.22	1.34

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	64	ARG	NE-CZ-NH2	-9.45	115.58	120.30
1	C	14	ASP	CB-CG-OD2	-9.08	110.13	118.30
1	C	139	PHE	CB-CG-CD1	8.03	126.42	120.80
1	B	65	ASP	CB-CG-OD1	7.68	125.21	118.30
1	C	65	ASP	CB-CG-OD2	-7.55	111.50	118.30
1	C	24	ASP	CB-CG-OD2	-7.20	111.82	118.30
1	C	65	ASP	CB-CG-OD1	7.09	124.68	118.30
1	A	14	ASP	CB-CG-OD2	-7.04	111.97	118.30
1	C	53	TYR	CB-CG-CD1	7.00	125.20	121.00
1	C	53	TYR	CB-CG-CD2	-6.91	116.85	121.00
1	B	65	ASP	CB-CG-OD2	-6.88	112.11	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	64	ARG	NE-CZ-NH1	6.83	123.72	120.30
1	A	117	LEU	CB-CG-CD1	6.46	121.98	111.00
1	A	104	ARG	NE-CZ-NH2	-6.38	117.11	120.30
1	B	64	ARG	NE-CZ-NH2	-5.93	117.34	120.30
1	A	24	ASP	CB-CG-OD2	-5.79	113.09	118.30
1	C	99	ARG	NE-CZ-NH1	5.63	123.11	120.30
1	A	120	PHE	CB-CG-CD1	-5.40	117.02	120.80
1	C	139	PHE	CB-CG-CD2	-5.37	117.04	120.80
1	B	62	TYR	CB-CG-CD2	5.37	124.22	121.00
1	C	112	TYR	CB-CG-CD2	-5.29	117.83	121.00
1	A	133	TYR	CZ-CE2-CD2	5.27	124.54	119.80
1	B	50	ASP	CB-CG-OD1	5.22	123.00	118.30
1	A	132	ASP	CB-CG-OD2	-5.21	113.61	118.30
1	C	46	TYR	CZ-CE2-CD2	-5.14	115.17	119.80
1	B	24	ASP	CB-CG-OD2	-5.06	113.75	118.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1140	0	1136	3	1
1	B	1218	0	1222	2	1
1	C	1260	0	1255	6	0
2	A	12	0	16	0	0
2	B	6	0	7	0	0
2	C	12	0	16	1	0
3	A	9	0	13	0	0
3	B	9	0	13	0	0
3	C	9	0	13	1	0
4	A	103	0	0	2	0
4	B	114	0	0	1	0
4	C	145	0	0	2	0
All	All	4037	0	3691	11	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 1.

All (11) 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:124:ASP:OD2	4:A:393:HOH:O	2.03	0.77
1:B:155:GLN:OE1	4:B:370:HOH:O	2.10	0.69
1:C:16:ALA:HA	1:C:19:LYS:HE2	1.75	0.67
1:B:77:VAL:HG11	1:B:150:LYS:HE2	1.88	0.55
1:A:5:ILE:HD11	1:A:102:VAL:HG23	1.96	0.48
1:C:144:ASP:O	1:C:145:LEU:HD23	2.15	0.46
1:C:103:LEU:HD21	3:C:203:HYH:H2	1.82	0.45
1:C:33:ASP:HB3	4:C:443:HOH:O	2.19	0.43
1:C:46:TYR:CD2	2:C:202:GOL:H11	2.54	0.43
1:A:124:ASP:HB2	4:A:393:HOH:O	2.20	0.41
1:C:79:SER:HB2	4:C:408:HOH:O	2.19	0.41

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)
1:A:7:GLN:NE2	1:B:9:ARG:NH2[2_664]	2.14	0.06

## 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	144/174 (83%)	142 (99%)	2 (1%)	0	100	100
1	B	154/174 (88%)	152 (99%)	2 (1%)	0	100	100
1	C	159/174 (91%)	159 (100%)	0	0	100	100
All	All	457/522 (88%)	453 (99%)	4 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	123/147 (84%)	123 (100%)	0	100	100
1	B	131/147 (89%)	130 (99%)	1 (1%)	81	66
1	C	136/147 (92%)	135 (99%)	1 (1%)	84	69
All	All	390/441 (88%)	388 (100%)	2 (0%)	88	78

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	7	GLN
1	C	7	GLN

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

Mol	Chain	Res	Type
1	C	7	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

8 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	HYH	B	202	-	8,8,8	1.30	2 (25%)	7,9,9	3.48	4 (57%)
2	GOL	B	201	-	5,5,5	0.91	0	5,5,5	0.85	0
2	GOL	C	201	-	5,5,5	0.84	0	5,5,5	0.77	0
3	HYH	C	203	-	8,8,8	0.95	1 (12%)	7,9,9	4.64	3 (42%)
2	GOL	A	201	-	5,5,5	0.47	0	5,5,5	0.56	0
2	GOL	A	202	-	5,5,5	0.90	0	5,5,5	0.40	0
2	GOL	C	202	-	5,5,5	0.50	0	5,5,5	1.85	2 (40%)
3	HYH	A	203	-	8,8,8	1.74	2 (25%)	7,9,9	4.22	3 (42%)

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	HYH	B	202	-	-	3/8/8/8	-
2	GOL	B	201	-	-	3/4/4/4	-
2	GOL	C	201	-	-	0/4/4/4	-
3	HYH	C	203	-	-	3/8/8/8	-
2	GOL	A	201	-	-	0/4/4/4	-
2	GOL	A	202	-	-	0/4/4/4	-
2	GOL	C	202	-	-	2/4/4/4	-
3	HYH	A	203	-	-	4/8/8/8	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	203	HYH	O-C4	3.84	1.51	1.42
3	B	202	HYH	O-C4	2.62	1.48	1.42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	203	HYH	C4-C5	2.37	1.55	1.52
3	C	203	HYH	O-C4	2.27	1.47	1.42
3	B	202	HYH	C4-C5	-2.22	1.50	1.52

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	203	HYH	C4-C5-N	10.21	125.59	117.25
3	A	203	HYH	C4-C5-N	8.94	124.55	117.25
3	B	202	HYH	C4-C5-N	7.18	123.11	117.25
3	C	203	HYH	O1-C5-N	-5.13	114.08	123.00
3	A	203	HYH	O-C4-C3	4.85	122.97	109.42
3	B	202	HYH	O1-C5-N	-4.28	115.56	123.00
3	C	203	HYH	O-C4-C3	4.14	120.99	109.42
3	A	203	HYH	O1-C5-N	-4.11	115.85	123.00
2	C	202	GOL	O1-C1-C2	2.95	124.37	110.20
2	C	202	GOL	O3-C3-C2	2.70	123.14	110.20
3	B	202	HYH	O1-C5-C4	2.59	121.32	118.94
3	B	202	HYH	O-C4-C3	2.48	116.35	109.42

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	202	HYH	O-C4-C5-O1
3	B	202	HYH	O-C4-C5-N
2	B	201	GOL	C1-C2-C3-O3
2	B	201	GOL	O2-C2-C3-O3
3	C	203	HYH	O-C4-C5-N
2	C	202	GOL	C1-C2-C3-O3
3	A	203	HYH	O-C4-C5-N
2	C	202	GOL	O2-C2-C3-O3
2	B	201	GOL	O1-C1-C2-O2
3	C	203	HYH	O-C4-C5-O1
3	A	203	HYH	O-C4-C5-O1
3	B	202	HYH	C3-C4-C5-O1
3	A	203	HYH	C3-C4-C5-O1
3	C	203	HYH	C3-C4-C5-O1
3	A	203	HYH	C3-C4-C5-N

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	203	HYH	1	0
2	C	202	GOL	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	144/174 (82%)	0.01	3 (2%) 63 68	12, 18, 32, 56	0
1	B	155/174 (89%)	0.08	7 (4%) 33 36	11, 17, 37, 58	0
1	C	160/174 (91%)	0.10	4 (2%) 57 62	9, 15, 26, 52	0
All	All	459/522 (87%)	0.06	14 (3%) 49 54	9, 16, 32, 58	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	7	GLN	5.8
1	B	158	ILE	5.0
1	B	159	SER	4.1
1	C	164	ASN	3.6
1	B	153	PRO	3.3
1	B	6	GLU	3.1
1	C	6	GLU	2.6
1	B	5	ILE	2.6
1	A	6	GLU	2.5
1	B	145	LEU	2.4
1	C	5	ILE	2.4
1	A	5	ILE	2.2
1	B	157	LEU	2.2
1	C	7	GLN	2.1

### 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 [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
2	GOL	B	201	6/6	0.86	0.22	34,41,46,51	0
3	HYH	B	202	9/9	0.88	0.14	19,26,32,41	0
2	GOL	C	202	6/6	0.91	0.16	29,38,43,48	0
3	HYH	A	203	9/9	0.92	0.15	20,31,36,47	0
2	GOL	A	201	6/6	0.95	0.09	17,20,22,23	0
2	GOL	A	202	6/6	0.95	0.09	17,23,23,24	0
2	GOL	C	201	6/6	0.96	0.09	14,20,22,26	0
3	HYH	C	203	9/9	0.96	0.10	16,23,26,34	0

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

There are no such residues in this entry.