



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 15, 2018 – 12:56 AM EST

PDB ID : 4KC7  
Title : Crystal Structure of Endo-1,5-alpha-L-arabinanase from *Thermotoga petrophila* RKU-1  
Authors : Nascimento, A.F.Z.; Polo, C.C.; Santos, C.R.; Costa, M.C.M.F.; Mesa, A.N.; Prade, R.A.; Ruller, R.; Squina, F.M.; Murakami, M.T.  
Deposited on : 2013-04-24  
Resolution : 1.75 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030736  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

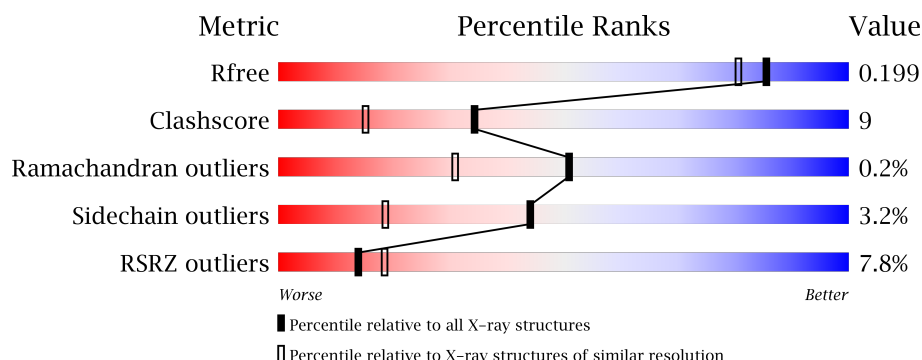
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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))
$R_{free}$	100719	1762 (1.76-1.76)
Clashscore	112137	1889 (1.76-1.76)
Ramachandran outliers	110173	1868 (1.76-1.76)
Sidechain outliers	110143	1868 (1.76-1.76)
RSRZ outliers	101464	1770 (1.76-1.76)

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. 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	474	
1	B	474	
1	C	474	

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
2	PEG	B	502	-	-	-	X
3	CA	A	502	-	-	-	X
3	CA	B	503	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11493 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 Glycoside hydrolase, family 43.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	449	Total	C	N	O	S	0	6	0
			3623	2337	602	674	10			
1	B	449	Total	C	N	O	S	0	7	0
			3632	2345	605	672	10			
1	C	433	Total	C	N	O	S	0	3	0
			3476	2244	582	640	10			

There are 69 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	INITIATING METHIONINE	UNP A5IKD4
A	-1	GLY	-	EXPRESSION TAG	UNP A5IKD4
A	0	SER	-	EXPRESSION TAG	UNP A5IKD4
A	1	SER	-	EXPRESSION TAG	UNP A5IKD4
A	2	HIS	-	EXPRESSION TAG	UNP A5IKD4
A	3	HIS	-	EXPRESSION TAG	UNP A5IKD4
A	4	HIS	-	EXPRESSION TAG	UNP A5IKD4
A	5	HIS	-	EXPRESSION TAG	UNP A5IKD4
A	6	HIS	-	EXPRESSION TAG	UNP A5IKD4
A	7	HIS	-	EXPRESSION TAG	UNP A5IKD4
A	8	SER	-	EXPRESSION TAG	UNP A5IKD4
A	9	SER	-	EXPRESSION TAG	UNP A5IKD4
A	10	GLY	-	EXPRESSION TAG	UNP A5IKD4
A	11	LEU	-	EXPRESSION TAG	UNP A5IKD4
A	12	VAL	-	EXPRESSION TAG	UNP A5IKD4
A	13	PRO	-	EXPRESSION TAG	UNP A5IKD4
A	14	ARG	-	EXPRESSION TAG	UNP A5IKD4
A	15	GLY	-	EXPRESSION TAG	UNP A5IKD4
A	16	SER	-	EXPRESSION TAG	UNP A5IKD4
A	17	HIS	-	EXPRESSION TAG	UNP A5IKD4
A	18	MET	-	EXPRESSION TAG	UNP A5IKD4
A	19	ALA	-	EXPRESSION TAG	UNP A5IKD4
A	20	SER	-	EXPRESSION TAG	UNP A5IKD4

*Continued on next page...*

*Continued from previous page...*

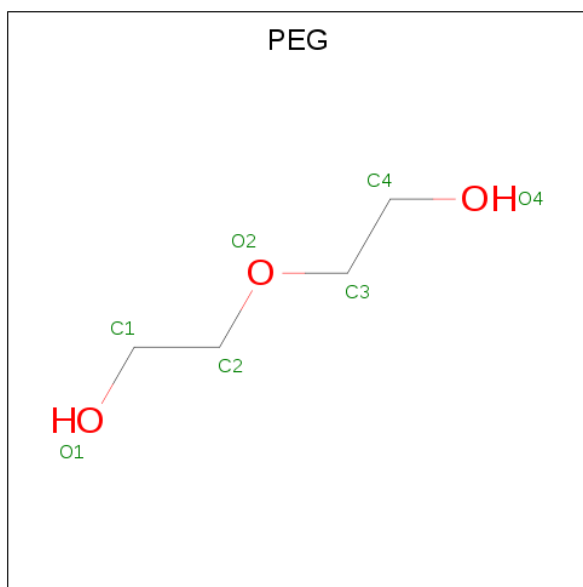
Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	MET	-	INITIATING METHIONINE	UNP A5IKD4
B	-1	GLY	-	EXPRESSION TAG	UNP A5IKD4
B	0	SER	-	EXPRESSION TAG	UNP A5IKD4
B	1	SER	-	EXPRESSION TAG	UNP A5IKD4
B	2	HIS	-	EXPRESSION TAG	UNP A5IKD4
B	3	HIS	-	EXPRESSION TAG	UNP A5IKD4
B	4	HIS	-	EXPRESSION TAG	UNP A5IKD4
B	5	HIS	-	EXPRESSION TAG	UNP A5IKD4
B	6	HIS	-	EXPRESSION TAG	UNP A5IKD4
B	7	HIS	-	EXPRESSION TAG	UNP A5IKD4
B	8	SER	-	EXPRESSION TAG	UNP A5IKD4
B	9	SER	-	EXPRESSION TAG	UNP A5IKD4
B	10	GLY	-	EXPRESSION TAG	UNP A5IKD4
B	11	LEU	-	EXPRESSION TAG	UNP A5IKD4
B	12	VAL	-	EXPRESSION TAG	UNP A5IKD4
B	13	PRO	-	EXPRESSION TAG	UNP A5IKD4
B	14	ARG	-	EXPRESSION TAG	UNP A5IKD4
B	15	GLY	-	EXPRESSION TAG	UNP A5IKD4
B	16	SER	-	EXPRESSION TAG	UNP A5IKD4
B	17	HIS	-	EXPRESSION TAG	UNP A5IKD4
B	18	MET	-	EXPRESSION TAG	UNP A5IKD4
B	19	ALA	-	EXPRESSION TAG	UNP A5IKD4
B	20	SER	-	EXPRESSION TAG	UNP A5IKD4
C	-2	MET	-	INITIATING METHIONINE	UNP A5IKD4
C	-1	GLY	-	EXPRESSION TAG	UNP A5IKD4
C	0	SER	-	EXPRESSION TAG	UNP A5IKD4
C	1	SER	-	EXPRESSION TAG	UNP A5IKD4
C	2	HIS	-	EXPRESSION TAG	UNP A5IKD4
C	3	HIS	-	EXPRESSION TAG	UNP A5IKD4
C	4	HIS	-	EXPRESSION TAG	UNP A5IKD4
C	5	HIS	-	EXPRESSION TAG	UNP A5IKD4
C	6	HIS	-	EXPRESSION TAG	UNP A5IKD4
C	7	HIS	-	EXPRESSION TAG	UNP A5IKD4
C	8	SER	-	EXPRESSION TAG	UNP A5IKD4
C	9	SER	-	EXPRESSION TAG	UNP A5IKD4
C	10	GLY	-	EXPRESSION TAG	UNP A5IKD4
C	11	LEU	-	EXPRESSION TAG	UNP A5IKD4
C	12	VAL	-	EXPRESSION TAG	UNP A5IKD4
C	13	PRO	-	EXPRESSION TAG	UNP A5IKD4
C	14	ARG	-	EXPRESSION TAG	UNP A5IKD4
C	15	GLY	-	EXPRESSION TAG	UNP A5IKD4
C	16	SER	-	EXPRESSION TAG	UNP A5IKD4

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	17	HIS	-	EXPRESSION TAG	UNP A5IKD4
C	18	MET	-	EXPRESSION TAG	UNP A5IKD4
C	19	ALA	-	EXPRESSION TAG	UNP A5IKD4
C	20	SER	-	EXPRESSION TAG	UNP A5IKD4

- Molecule 2 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 7 4 3	0	0
2	B	1	Total C O 7 4 3	0	0
2	B	1	Total C O 7 4 3	0	0
2	C	1	Total C O 7 4 3	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Ca 1 1	0	0
3	A	1	Total Ca 1 1	0	0
3	C	1	Total Ca 1 1	0	0

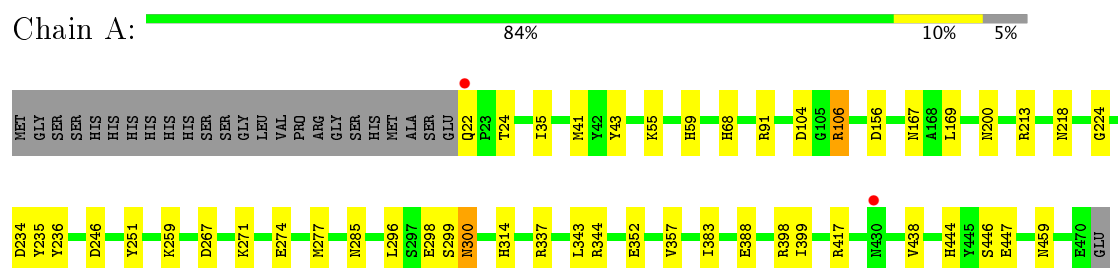
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	355	Total 355	O 355	0	0
4	B	176	Total 176	O 176	0	0
4	C	200	Total 200	O 200	0	0

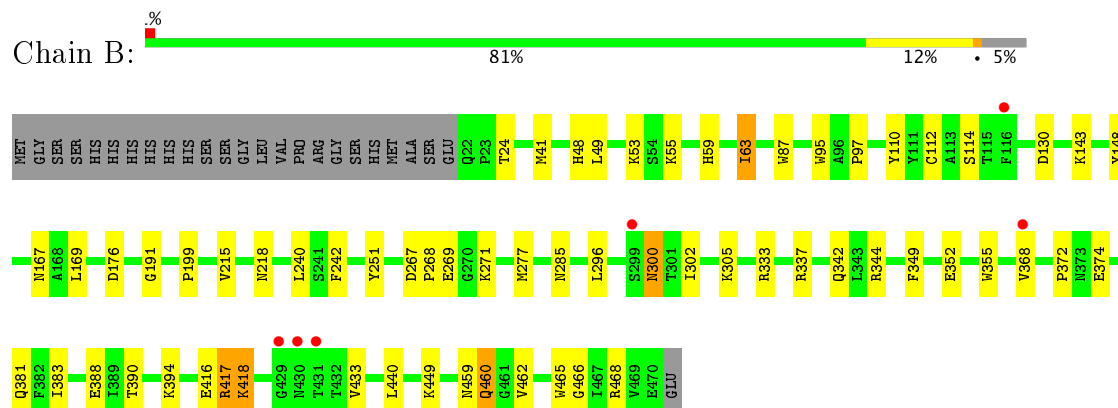
### 3 Residue-property plots [i](#)

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.

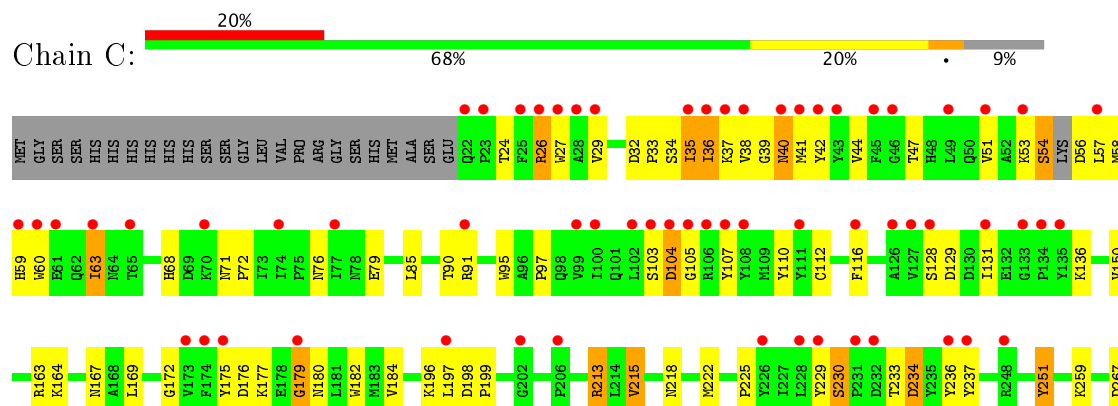
- Molecule 1: Glycoside hydrolase, family 43



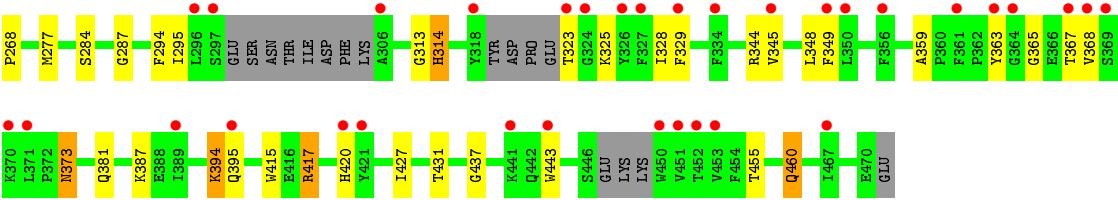
- Molecule 1: Glycoside hydrolase, family 43



- Molecule 1: Glycoside hydrolase, family 43







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	41.67Å 86.91Å 194.85Å 90.00° 90.15° 90.00°	Depositor
Resolution (Å)	40.77 – 1.75 40.77 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.1 (40.77-1.75) 99.1 (40.77-1.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.49 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.160 , 0.197 0.168 , 0.199	Depositor DCC
$R_{free}$ test set	7010 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	20.1	Xtriage
Anisotropy	0.534	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 33.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.398 for h,-k,-l	Xtriage
Reported twinning fraction	0.612 for H, K, L 0.388 for -h,-k,l	Depositor
Outliers	0 of 139453 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11493	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.27% 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: CA, PEG

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	3/3766 (0.1%)	1.14	11/5117 (0.2%)
1	B	0.94	1/3780 (0.0%)	1.03	5/5134 (0.1%)
1	C	0.86	1/3598 (0.0%)	0.98	8/4885 (0.2%)
All	All	0.99	5/11144 (0.0%)	1.05	24/15136 (0.2%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	182	TRP	CB-CG	-7.43	1.36	1.50
1	A	344	ARG	CZ-NH2	5.92	1.40	1.33
1	A	224	GLY	N-CA	-5.74	1.37	1.46
1	B	465	TRP	CG-CD1	-5.37	1.29	1.36
1	A	344	ARG	NE-CZ	-5.15	1.26	1.33

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	213	ARG	NE-CZ-NH1	10.61	125.61	120.30
1	A	106	ARG	NE-CZ-NH1	9.51	125.06	120.30
1	A	106	ARG	NE-CZ-NH2	-9.16	115.72	120.30
1	C	213	ARG	NE-CZ-NH2	-8.22	116.19	120.30
1	C	35	ILE	CB-CA-C	-7.19	97.22	111.60
1	A	343	LEU	CB-CG-CD2	-7.01	99.08	111.00
1	B	468	ARG	NE-CZ-NH2	-6.97	116.81	120.30
1	A	169	LEU	CA-CB-CG	6.44	130.12	115.30
1	C	373[A]	ASN	CB-CA-C	6.34	123.09	110.40
1	C	373[B]	ASN	CB-CA-C	6.34	123.09	110.40
1	A	156	ASP	CB-CG-OD1	5.91	123.61	118.30
1	B	267	ASP	CB-CG-OD2	-5.47	113.38	118.30
1	A	236	TYR	CD1-CE1-CZ	5.43	124.68	119.80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	267	ASP	CB-CG-OD1	5.35	123.12	118.30
1	C	36	ILE	CB-CA-C	-5.34	100.91	111.60
1	A	213	ARG	NE-CZ-NH1	5.33	122.96	120.30
1	B	333	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	A	234	ASP	CB-CG-OD2	-5.30	113.53	118.30
1	A	344	ARG	NE-CZ-NH1	-5.27	117.66	120.30
1	C	169	LEU	CA-CB-CG	5.17	127.18	115.30
1	A	200	ASN	CB-CA-C	-5.14	100.11	110.40
1	B	176	ASP	CB-CG-OD1	5.12	122.91	118.30
1	B	242	PHE	CB-CG-CD1	5.03	124.32	120.80
1	C	234	ASP	CB-CG-OD1	5.01	122.81	118.30

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	3623	0	3479	30	0
1	B	3632	0	3501	46	0
1	C	3476	0	3335	123	2
2	A	7	0	10	0	0
2	B	14	0	20	3	0
2	C	7	0	10	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	355	0	0	12	2
4	B	176	0	0	10	0
4	C	200	0	0	46	0
All	All	11493	0	10355	200	2

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 9.

All (200) 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:C:58:MET:HB3	4:C:779:HOH:O	1.29	1.25
1:B:459:ASN:HB3	4:B:707:HOH:O	1.39	1.23
1:C:136:LYS:HB3	4:C:621:HOH:O	1.51	1.08
1:C:38:VAL:HG21	1:C:107:TYR:OH	1.55	1.05
1:C:175:TYR:CD1	4:C:789:HOH:O	2.12	1.02
1:C:39:GLY:O	1:C:41:MET:N	1.96	0.98
1:A:459:ASN:ND2	4:A:866:HOH:O	1.97	0.96
1:C:373[A]:ASN:OD1	1:C:417:ARG:NH2	2.01	0.93
1:C:175:TYR:HD1	4:C:789:HOH:O	1.46	0.92
1:A:274:GLU:OE2	4:A:788:HOH:O	1.88	0.91
1:C:229:TYR:OH	1:C:234:ASP:OD1	1.87	0.91
1:C:345:VAL:HG21	4:C:677:HOH:O	1.72	0.90
1:C:38:VAL:HG13	1:C:39:GLY:CA	2.03	0.88
1:C:51:VAL:HB	1:C:63:ILE:HG22	1.55	0.88
1:B:433:VAL:HG11	2:B:502:PEG:H41	1.58	0.85
1:A:300:ASN:HB3	4:A:863:HOH:O	1.76	0.85
1:C:230:SER:OG	1:C:233:THR:HG22	1.79	0.83
1:B:24[A]:THR:HG23	1:B:59:HIS:NE2	1.96	0.81
1:B:416:GLU:OE2	4:B:732:HOH:O	1.99	0.79
1:B:352:GLU:OE2	1:B:418:LYS:NZ	2.15	0.78
1:C:38:VAL:HG13	1:C:39:GLY:HA3	1.65	0.78
1:B:296:LEU:CD2	1:B:383:ILE:HD11	2.16	0.75
1:C:395:GLN:HA	4:C:759:HOH:O	1.85	0.75
1:C:236:TYR:CE2	4:C:790:HOH:O	2.41	0.73
1:C:131:ILE:HG22	4:C:714:HOH:O	1.88	0.73
1:C:215:VAL:HG22	1:C:222:MET:SD	2.29	0.71
1:C:27:TRP:CD1	4:C:666:HOH:O	2.42	0.71
1:A:296:LEU:CD2	1:A:383:ILE:HD11	2.21	0.71
1:C:179:GLY:O	4:C:789:HOH:O	2.08	0.71
1:C:36:ILE:HG22	4:C:672:HOH:O	1.91	0.70
1:C:363:TYR:CD2	4:C:780:HOH:O	2.43	0.70
1:A:235:TYR:CE1	1:A:259:LYS:HE2	2.27	0.70
1:C:415:TRP:HZ3	1:C:417:ARG:HG3	1.56	0.70
1:C:51:VAL:HB	1:C:63:ILE:CG2	2.21	0.70
1:C:177:LYS:HG2	4:C:788:HOH:O	1.91	0.70
1:C:36:ILE:HD11	4:C:778:HOH:O	1.92	0.69
1:C:417:ARG:HD3	1:C:417:ARG:C	2.12	0.69
1:C:38:VAL:HG13	1:C:39:GLY:HA2	1.74	0.69
1:B:459:ASN:CB	4:B:707:HOH:O	2.15	0.68
1:C:27:TRP:NE1	4:C:666:HOH:O	2.27	0.68
1:B:271:LYS:HE3	1:B:285:ASN:O	1.93	0.67
1:C:54:SER:HG	1:C:56:ASP:N	1.92	0.67

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:105:GLY:O	4:C:799:HOH:O	2.12	0.67
1:C:417:ARG:HD3	1:C:417:ARG:O	1.93	0.67
1:C:38:VAL:CG1	1:C:39:GLY:HA3	2.25	0.66
1:B:342:GLN:OE1	1:B:344:ARG:NH1	2.22	0.64
1:C:236:TYR:CD2	4:C:790:HOH:O	2.50	0.63
1:C:60:TRP:HZ2	4:C:677:HOH:O	1.79	0.63
1:A:24[B]:THR:HG21	4:A:888:HOH:O	1.98	0.63
1:C:233:THR:HG21	1:C:237:TYR:OH	1.99	0.63
1:C:39:GLY:C	1:C:41:MET:N	2.52	0.62
1:C:323:THR:HB	1:C:325:LYS:HE3	1.82	0.61
1:C:40:ASN:HB2	4:C:800:HOH:O	1.99	0.61
1:C:90:THR:O	1:C:91:ARG:HG2	2.01	0.61
1:B:296:LEU:HD23	1:B:383:ILE:HD11	1.82	0.60
1:A:235:TYR:CZ	1:A:259:LYS:HE2	2.36	0.60
1:A:246:ASP:OD1	4:A:859:HOH:O	2.17	0.59
1:C:287:GLY:O	1:C:387:LYS:HE3	2.03	0.58
1:C:344:ARG:NH2	4:C:666:HOH:O	2.36	0.58
1:C:58:MET:CB	4:C:779:HOH:O	2.10	0.57
1:A:41:MET:SD	1:A:55:LYS:HG2	2.44	0.57
1:A:68:HIS:HD2	4:A:813:HOH:O	1.88	0.56
1:B:191:GLY:HA2	4:B:691:HOH:O	2.06	0.55
1:C:437:GLY:HA3	1:C:455:THR:O	2.05	0.55
1:C:349:PHE:CE1	1:C:359:ALA:HB2	2.42	0.55
1:C:349:PHE:HE1	1:C:359:ALA:HB2	1.72	0.55
1:B:388:GLU:HG2	1:B:390:THR:HG22	1.88	0.55
1:C:387:LYS:NZ	1:C:460:GLN:O	2.31	0.55
1:B:53:LYS:HE2	1:B:63:ILE:HD11	1.88	0.55
1:C:365:GLY:HA2	4:C:782:HOH:O	2.07	0.55
1:C:395:GLN:CA	4:C:759:HOH:O	2.47	0.55
1:C:294:PHE:HA	1:C:394:LYS:O	2.07	0.55
1:C:90:THR:C	1:C:91:ARG:HG2	2.28	0.55
1:C:38:VAL:HG21	1:C:107:TYR:HH	1.67	0.54
1:C:97:PRO:HA	1:C:110:TYR:O	2.07	0.54
1:A:296:LEU:HD23	1:A:383:ILE:HD11	1.89	0.54
1:C:26:ARG:CZ	1:C:26:ARG:HB3	2.36	0.54
1:C:229:TYR:CZ	1:C:234:ASP:HA	2.43	0.53
1:C:35:ILE:HG12	1:C:329:PHE:CD1	2.43	0.53
1:C:417:ARG:O	1:C:417:ARG:CD	2.56	0.53
1:C:363:TYR:CG	4:C:780:HOH:O	2.61	0.52
1:C:68:HIS:O	1:C:71:ASN:ND2	2.42	0.52
1:C:175:TYR:CE1	4:C:789:HOH:O	2.53	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:236:TYR:HE2	4:C:790:HOH:O	1.85	0.52
1:C:95:TRP:HB2	1:C:112:CYS:SG	2.49	0.52
1:B:112:CYS:SG	1:B:169:LEU:HD22	2.50	0.52
1:C:345:VAL:HG11	4:C:677:HOH:O	2.10	0.52
1:C:35:ILE:HD11	1:C:329:PHE:CD1	2.44	0.52
1:C:24:THR:OG1	1:C:59:HIS:NE2	2.42	0.52
1:C:116:PHE:HB3	4:C:675:HOH:O	2.09	0.52
1:A:444:HIS:CD2	1:A:447:GLU:H	2.28	0.52
1:C:36:ILE:CD1	4:C:778:HOH:O	2.55	0.51
1:C:42:TYR:O	1:C:53:LYS:HA	2.10	0.51
1:B:41:MET:SD	1:B:55:LYS:HG3	2.50	0.51
1:C:325:LYS:HE2	4:C:644:HOH:O	2.08	0.51
1:C:39:GLY:C	1:C:41:MET:H	2.11	0.51
1:C:196:LYS:NZ	4:C:731:HOH:O	2.42	0.51
1:A:24[B]:THR:HG22	1:A:59:HIS:NE2	2.26	0.51
1:A:337:ARG:HD2	4:A:707:HOH:O	2.10	0.51
1:A:337:ARG:NH1	4:A:877:HOH:O	2.44	0.51
1:B:352:GLU:CD	1:B:418:LYS:HZ3	2.13	0.51
1:C:34:SER:O	1:C:44:VAL:HA	2.10	0.50
1:A:271:LYS:HE3	1:A:285:ASN:O	2.11	0.50
1:C:443:TRP:HB3	4:C:783:HOH:O	2.11	0.50
1:C:85:LEU:HD22	1:C:91:ARG:HA	1.94	0.50
1:A:444:HIS:HD2	1:A:446:SER:H	1.59	0.50
1:C:328:ILE:HB	1:C:348:LEU:HD11	1.94	0.49
1:B:296:LEU:HD21	1:B:383:ILE:HD11	1.92	0.49
1:C:42:TYR:CD2	1:C:57:LEU:HD21	2.47	0.49
1:C:417:ARG:C	1:C:417:ARG:CD	2.80	0.49
1:C:136:LYS:CB	4:C:621:HOH:O	2.32	0.49
1:C:177:LYS:HD3	1:C:229:TYR:CE2	2.48	0.49
1:B:95:TRP:CZ3	1:B:114:SER:HB3	2.48	0.49
1:C:79:GLU:HG3	4:C:687:HOH:O	2.13	0.48
1:C:163:ARG:HE	1:C:218:ASN:ND2	2.10	0.48
1:C:427:ILE:HD12	1:C:427:ILE:N	2.28	0.47
1:C:284:SER:O	1:C:387:LYS:HG3	2.15	0.47
1:A:104:ASP:OD2	1:A:106:ARG:HD3	2.15	0.47
1:B:374[A]:GLU:H	1:B:374[A]:GLU:CD	2.17	0.47
1:C:35:ILE:HG12	1:C:329:PHE:CE1	2.50	0.47
1:B:352:GLU:HG3	4:B:641:HOH:O	2.12	0.47
1:B:352:GLU:HG3	4:B:605:HOH:O	2.13	0.47
1:A:104:ASP:OD2	1:A:106:ARG:CD	2.63	0.47
1:B:48:HIS:NE2	2:B:501:PEG:H22	2.30	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:271:LYS:NZ	4:A:873:HOH:O	2.48	0.47
1:B:349:PHE:CE1	1:B:368:VAL:CG2	2.97	0.46
1:C:26:ARG:NH1	1:C:60:TRP:O	2.47	0.46
1:C:329:PHE:CD1	1:C:329:PHE:N	2.83	0.46
1:B:372:PRO:HB2	1:B:374[A]:GLU:HG2	1.98	0.46
1:B:269:GLU:OE2	1:B:460:GLN:NE2	2.49	0.46
1:C:236:TYR:HE1	1:C:259:LYS:O	1.99	0.46
1:C:42:TYR:CE2	1:C:57:LEU:HD21	2.50	0.46
1:B:337[A]:ARG:HD2	4:B:710:HOH:O	2.15	0.46
1:C:176:ASP:CG	4:C:788:HOH:O	2.54	0.46
1:C:128:SER:OG	1:C:129:ASP:N	2.50	0.45
1:A:298[B]:GLU:OE2	1:A:398:ARG:NH2	2.49	0.45
1:C:367:THR:N	4:C:616:HOH:O	2.48	0.45
1:B:215:VAL:HG21	1:B:240:LEU:HD11	1.97	0.45
1:C:40:ASN:C	4:C:800:HOH:O	2.55	0.45
1:C:32:ASP:N	1:C:33:PRO:CD	2.80	0.45
1:C:40:ASN:CB	4:C:800:HOH:O	2.58	0.45
1:C:314:HIS:HD2	4:C:797:HOH:O	1.98	0.45
1:C:29:VAL:CG1	4:C:668:HOH:O	2.65	0.45
1:C:35:ILE:HG22	1:C:36:ILE:N	2.32	0.44
1:C:180:ASN:C	4:C:789:HOH:O	2.56	0.44
1:C:39:GLY:O	1:C:41:MET:CA	2.64	0.44
1:A:24[B]:THR:CG2	4:A:889:HOH:O	2.65	0.44
1:C:104:ASP:N	1:C:104:ASP:OD1	2.49	0.44
1:C:38:VAL:CG2	1:C:107:TYR:OH	2.45	0.44
1:C:229:TYR:HB2	4:C:790:HOH:O	2.17	0.44
1:B:449:LYS:NZ	4:B:725:HOH:O	2.51	0.43
1:C:267:ASP:HB2	1:C:268:PRO:HD2	2.00	0.43
1:C:295:ILE:HG23	1:C:395:GLN:HG2	2.00	0.43
1:C:32:ASP:OD2	1:C:47:THR:OG1	2.27	0.43
1:C:58:MET:CA	4:C:779:HOH:O	2.55	0.43
1:B:218:ASN:OD1	1:B:277:MET:CE	2.67	0.43
1:C:184:VAL:CG2	1:C:225:PRO:HB2	2.49	0.43
1:B:462:VAL:HB	2:B:502:PEG:H42	1.99	0.43
1:B:87:TRP:CD2	1:B:143:LYS:HD2	2.53	0.43
1:B:277:MET:HB2	1:B:277:MET:HE3	1.86	0.43
1:B:388:GLU:HG2	1:B:390:THR:CG2	2.48	0.43
1:C:277:MET:HB2	1:C:277:MET:HE2	1.91	0.43
1:B:388:GLU:CD	1:B:394:LYS:HZ1	2.22	0.43
1:A:314:HIS:CE1	4:A:840:HOH:O	2.72	0.42
1:B:148:TYR:C	1:B:148:TYR:CD1	2.93	0.42

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:97:PRO:HA	1:B:110:TYR:O	2.19	0.42
1:C:381:GLN:HG2	4:C:795:HOH:O	2.20	0.42
1:A:444:HIS:HE1	4:A:897:HOH:O	2.03	0.42
1:B:114:SER:HB2	1:B:169:LEU:HD21	2.02	0.42
1:C:71:ASN:HA	1:C:72:PRO:HD3	1.84	0.42
1:A:41:MET:SD	1:A:55:LYS:CG	3.08	0.41
1:B:268:PRO:HG3	1:B:355:TRP:CE2	2.55	0.41
1:C:35:ILE:CG1	1:C:329:PHE:CD1	3.04	0.41
1:B:300:ASN:OD1	1:B:302:ILE:O	2.37	0.41
1:C:365:GLY:CA	4:C:782:HOH:O	2.66	0.41
1:C:197:LEU:O	1:C:199:PRO:HD3	2.21	0.41
1:C:234:ASP:CG	1:C:234:ASP:O	2.58	0.41
1:C:35:ILE:CD1	1:C:329:PHE:CD1	3.03	0.41
1:C:198:ASP:O	1:C:199:PRO:C	2.58	0.41
1:C:39:GLY:O	1:C:41:MET:CB	2.69	0.41
1:C:54:SER:CB	4:C:678:HOH:O	2.69	0.41
1:A:277:MET:HB2	1:A:277:MET:HE3	1.89	0.41
1:C:172:GLY:N	1:C:184:VAL:O	2.45	0.41
1:B:271:LYS:CE	1:B:285:ASN:O	2.65	0.41
1:C:251:TYR:CD2	1:C:313:GLY:HA3	2.56	0.41
1:C:38:VAL:H	1:C:39:GLY:HA3	1.85	0.41
1:A:35:ILE:HD12	1:A:43:TYR:O	2.21	0.41
1:C:267:ASP:HB2	1:C:268:PRO:CD	2.51	0.41
1:C:26:ARG:HH11	1:C:26:ARG:CG	2.34	0.41
1:B:130:ASP:CG	4:B:760:HOH:O	2.59	0.40
1:B:49:LEU:HA	1:B:49:LEU:HD23	1.93	0.40
1:A:218:ASN:OD1	1:A:277:MET:HE1	2.21	0.40
1:A:357:VAL:HG21	1:A:438[B]:VAL:HG12	2.03	0.40
1:C:368:VAL:HG11	1:C:420:HIS:CG	2.56	0.40
1:A:399:ILE:HG21	1:A:399:ILE:HD13	1.91	0.40
1:B:381:GLN:O	1:B:466:GLY:HA2	2.21	0.40

All (2) 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:C:213:ARG:NH2	4:A:725:HOH:O[2_555]	1.23	0.97
1:C:213:ARG:CZ	4:A:725:HOH:O[2_555]	1.92	0.28

## 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	453/474 (96%)	436 (96%)	17 (4%)	0	100	100
1	B	454/474 (96%)	441 (97%)	13 (3%)	0	100	100
1	C	426/474 (90%)	407 (96%)	16 (4%)	3 (1%)	25	9
All	All	1333/1422 (94%)	1284 (96%)	46 (4%)	3 (0%)	51	31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	40	ASN
1	C	314	HIS
1	C	179	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	391/406 (96%)	382 (98%)	9 (2%)	56	31
1	B	392/406 (97%)	381 (97%)	11 (3%)	49	24
1	C	372/406 (92%)	355 (95%)	17 (5%)	31	10
All	All	1155/1218 (95%)	1118 (97%)	37 (3%)	44	19

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	22	GLN
1	A	91	ARG
1	A	167	ASN
1	A	251	TYR
1	A	299	SER
1	A	300	ASN
1	A	352	GLU
1	A	388	GLU
1	A	417	ARG
1	B	63	ILE
1	B	167	ASN
1	B	199	PRO
1	B	251	TYR
1	B	300	ASN
1	B	305	LYS
1	B	417[A]	ARG
1	B	417[B]	ARG
1	B	418	LYS
1	B	440	LEU
1	B	460	GLN
1	C	26	ARG
1	C	37	LYS
1	C	54	SER
1	C	63	ILE
1	C	76	ASN
1	C	103	SER
1	C	104	ASP
1	C	150	VAL
1	C	164	LYS
1	C	167	ASN
1	C	215	VAL
1	C	230	SER
1	C	251	TYR
1	C	394	LYS
1	C	417	ARG
1	C	431	THR
1	C	460	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	68	HIS
1	A	167	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	275	ASN
1	A	373	ASN
1	A	430	ASN
1	A	444	HIS
1	B	76	ASN
1	B	167	ASN
1	B	314	HIS
1	B	444	HIS
1	C	76	ASN
1	C	167	ASN
1	C	200	ASN
1	C	347	GLN
1	C	430	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 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	$\# Z  > 2$
2	PEG	A	501	-	6,6,6	0.59	0	5,5,5	0.79	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	PEG	B	501	-	6,6,6	0.43	0	5,5,5	1.43	1 (20%)
2	PEG	B	502	-	6,6,6	0.62	0	5,5,5	0.70	0
2	PEG	C	501	-	6,6,6	0.48	0	5,5,5	1.01	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	PEG	A	501	-	-	0/4/4/4	0/0/0/0
2	PEG	B	501	-	-	0/4/4/4	0/0/0/0
2	PEG	B	502	-	-	0/4/4/4	0/0/0/0
2	PEG	C	501	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	PEG	O2-C2-C1	2.29	120.71	110.15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	501	PEG	1	0
2	B	502	PEG	2	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	449/474 (94%)	-0.02	2 (0%) 92 95	10, 16, 29, 54	2 (0%)
1	B	449/474 (94%)	0.11	6 (1%) 77 84	17, 25, 38, 54	1 (0%)
1	C	433/474 (91%)	1.20	96 (22%) 1 1	22, 34, 51, 63	5 (1%)
All	All	1331/1422 (93%)	0.42	104 (7%) 14 18	10, 25, 46, 63	8 (0%)

All (104) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	133	GLY	7.0
1	C	443	TRP	6.2
1	C	175	TYR	5.9
1	C	127	VAL	5.6
1	C	179	GLY	5.6
1	C	63	ILE	5.5
1	C	371	LEU	5.5
1	C	323	THR	5.2
1	C	296	LEU	4.8
1	C	363	TYR	4.4
1	C	108	TYR	4.3
1	C	327	PHE	4.2
1	A	22	GLN	4.2
1	C	45	PHE	4.2
1	C	116	PHE	4.1
1	C	60	TRP	4.1
1	C	42	TYR	4.1
1	C	49	LEU	4.1
1	C	103	SER	4.0
1	B	116	PHE	3.9
1	C	29	VAL	3.9
1	C	318	TYR	3.9
1	C	43	TYR	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	51	VAL	3.8
1	C	306	ALA	3.8
1	C	453	VAL	3.7
1	C	22	GLN	3.7
1	C	26	ARG	3.7
1	C	370	LYS	3.7
1	C	326	TYR	3.6
1	C	28	ALA	3.6
1	C	450	TRP	3.6
1	C	368	VAL	3.5
1	C	104	ASP	3.4
1	C	131	ILE	3.4
1	C	40	ASN	3.4
1	C	111	TYR	3.3
1	C	46	GLY	3.3
1	C	25	PHE	3.3
1	C	134	PRO	3.3
1	C	91	ARG	3.2
1	C	53	LYS	3.1
1	C	99	VAL	3.1
1	C	231	PRO	3.1
1	C	23	PRO	3.1
1	C	349	PHE	3.1
1	C	107	TYR	3.1
1	C	367	THR	3.1
1	C	226	TYR	3.0
1	C	38	VAL	3.0
1	C	77	ILE	3.0
1	C	297	SER	2.9
1	C	173	VAL	2.9
1	C	126	ALA	2.9
1	C	57	LEU	2.8
1	C	65	THR	2.8
1	C	369	SER	2.8
1	B	299	SER	2.7
1	A	430	ASN	2.7
1	C	356	PHE	2.7
1	C	202	GLY	2.6
1	C	229	TYR	2.6
1	C	74	ILE	2.6
1	C	364	GLY	2.6
1	C	345	VAL	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	237	TYR	2.6
1	C	395	GLN	2.6
1	B	430	ASN	2.5
1	C	35	ILE	2.5
1	C	70	LYS	2.5
1	C	361	PHE	2.5
1	C	100	ILE	2.4
1	C	41	MET	2.4
1	C	350	LEU	2.4
1	C	389	ILE	2.4
1	B	429	GLY	2.3
1	C	27	TRP	2.3
1	C	452	THR	2.3
1	C	451	VAL	2.3
1	C	135	TYR	2.3
1	C	61	GLU	2.3
1	C	59	HIS	2.3
1	C	36	ILE	2.3
1	B	368	VAL	2.2
1	C	236	TYR	2.2
1	C	324	GLY	2.2
1	C	102	LEU	2.2
1	C	206	PRO	2.2
1	C	105	GLY	2.2
1	C	37	LYS	2.2
1	C	174	PHE	2.1
1	C	334	PHE	2.1
1	C	441	LYS	2.1
1	C	421	TYR	2.1
1	C	197	LEU	2.1
1	C	420	HIS	2.1
1	C	248	ARG	2.1
1	C	232	ASP	2.1
1	C	128	SER	2.1
1	C	467	ILE	2.1
1	B	431	THR	2.0
1	C	106	ARG	2.0
1	C	228	LEU	2.0
1	C	329	PHE	2.0



## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.3 Carbohydrates [i](#)

There are no 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	CA	B	503	1/1	0.94	0.21	8.68	39,39,39,39	0
3	CA	A	502	1/1	0.95	0.13	5.31	39,39,39,39	0
2	PEG	B	502	7/7	0.87	0.25	3.02	31,36,40,46	0
2	PEG	A	501	7/7	0.93	0.13	1.87	28,29,34,39	0
2	PEG	B	501	7/7	0.85	0.14	0.66	33,35,43,48	0
2	PEG	C	501	7/7	0.84	0.11	0.56	40,41,42,42	0
3	CA	C	502	1/1	0.96	0.07	-1.98	41,41,41,41	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.