



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:13 AM GMT

PDB ID : 2PRR  
Title : Crystal structure of alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein (YP\_296737.1) from *Ralstonia eutropha* JMP134 at 2.15 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2007-05-04  
Resolution : 2.15 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  
<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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

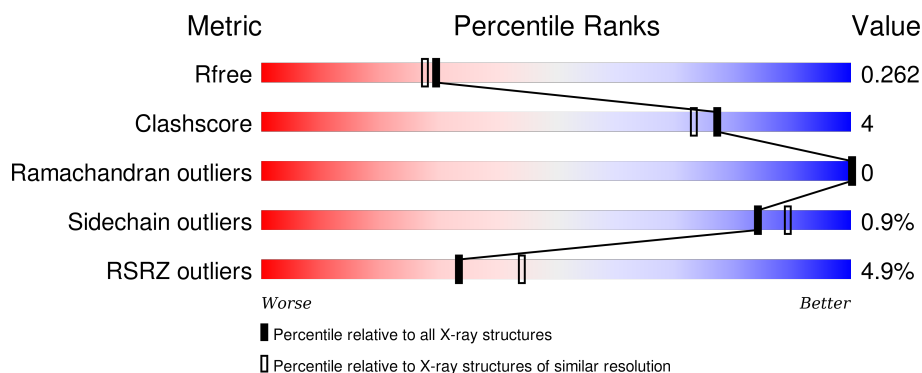
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
$R_{free}$	91344	1045 (2.16-2.16)
Clashscore	102246	1152 (2.16-2.16)
Ramachandran outliers	100387	1131 (2.16-2.16)
Sidechain outliers	100360	1131 (2.16-2.16)
RSRZ outliers	91569	1050 (2.16-2.16)

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	197	<div> <div>6%</div> <div>84%</div> <div>12%</div> <div>• •</div> </div>
1	B	197	<div> <div>4%</div> <div>85%</div> <div>11%</div> <div>• •</div> </div>
1	C	197	<div> <div>3%</div> <div>90%</div> <div>5%</div> <div>• •</div> </div>
1	D	197	<div> <div>6%</div> <div>91%</div> <div>5%</div> <div>• •</div> </div>
1	E	197	<div> <div>2%</div> <div>90%</div> <div>6%</div> <div>• •</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	197	
1	G	197	
1	H	197	
1	I	197	
1	J	197	
1	K	197	
1	L	197	

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	PGE	B	197	-	-	-	X
2	PGE	C	197	-	-	-	X
2	PGE	E	197	-	-	-	X
2	PGE	G	197	-	-	-	X
2	PGE	H	197	-	-	-	X
3	PEG	A	198	-	-	-	X
3	PEG	B	198	-	-	-	X
3	PEG	C	199	-	-	-	X
3	PEG	E	198	-	-	-	X
3	PEG	F	199	-	-	-	X
3	PEG	G	199	-	-	-	X
3	PEG	H	199	-	-	-	X
3	PEG	J	199	-	-	-	X
3	PEG	K	198	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 19292 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 Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	190	Total	C	N	O	S	Se	0	1	0
			1497	960	260	268	3	6			
1	B	190	Total	C	N	O	S	Se	0	1	0
			1493	955	259	270	3	6			
1	C	189	Total	C	N	O	S	Se	0	0	0
			1483	951	257	266	3	6			
1	D	191	Total	C	N	O	S	Se	0	0	0
			1481	954	257	261	3	6			
1	E	190	Total	C	N	O	S	Se	0	0	0
			1486	952	260	265	3	6			
1	F	190	Total	C	N	O	S	Se	0	0	0
			1478	949	257	263	3	6			
1	G	189	Total	C	N	O	S	Se	0	0	0
			1474	945	256	264	3	6			
1	H	190	Total	C	N	O	S	Se	0	0	0
			1475	948	258	260	3	6			
1	I	189	Total	C	N	O	S	Se	0	0	0
			1462	935	255	263	3	6			
1	J	188	Total	C	N	O	S	Se	0	0	0
			1455	932	253	261	3	6			
1	K	191	Total	C	N	O	S	Se	0	0	0
			1463	937	257	260	3	6			
1	L	189	Total	C	N	O	S	Se	0	0	0
			1462	940	253	260	3	6			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
A	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
A	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
A	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
A	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
A	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
B	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
B	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
B	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
B	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
B	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
B	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
B	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
C	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
C	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
C	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
C	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
C	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
C	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
C	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
D	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
D	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
D	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
D	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
D	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
D	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
D	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
E	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
E	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
E	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
E	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
E	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
E	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
E	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
E	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
F	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
F	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
F	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
F	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
F	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
F	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90

*Continued on next page...*

*Continued from previous page...*

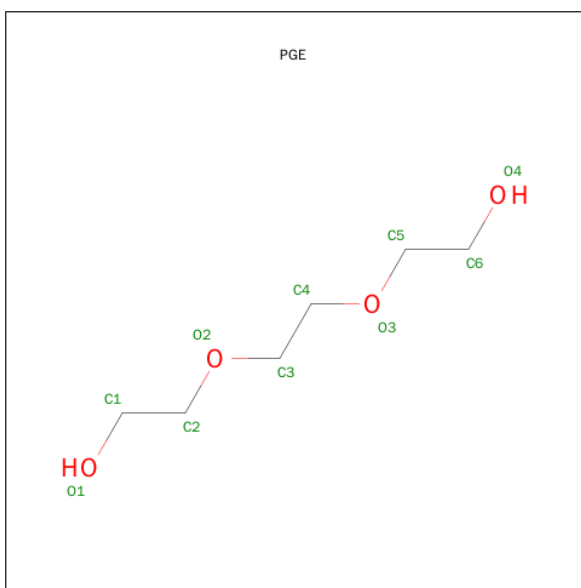
Chain	Residue	Modelled	Actual	Comment	Reference
F	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
F	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
G	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
G	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
G	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
G	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
G	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
G	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
G	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
G	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
H	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
H	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
H	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
H	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
H	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
H	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
H	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
H	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
I	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
I	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
I	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
I	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
I	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
I	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
I	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
I	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
J	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
J	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
J	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
J	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
J	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
J	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
J	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
J	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
K	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
K	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
K	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
K	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
K	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
K	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
K	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
K	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90

*Continued on next page...*

Continued from previous page...

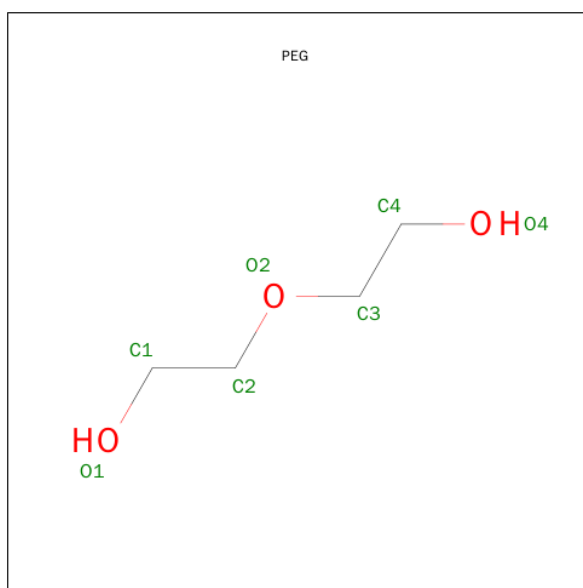
Chain	Residue	Modelled	Actual	Comment	Reference
L	0	GLY	-	LEADER SEQUENCE	UNP Q46Y90
L	1	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
L	62	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
L	75	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
L	126	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
L	174	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
L	180	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90
L	189	MSE	MET	MODIFIED RESIDUE	UNP Q46Y90

- Molecule 2 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	G	1	Total	C	O	0	0
			10	6	4		
2	E	1	Total	C	O	0	0
			10	6	4		
2	H	1	Total	C	O	0	0
			10	6	4		
2	C	1	Total	C	O	0	0
			10	6	4		
2	B	1	Total	C	O	0	0
			10	6	4		
2	D	1	Total	C	O	0	0
			10	6	4		
2	F	1	Total	C	O	0	0
			10	6	4		

- Molecule 3 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
3	F	1	Total	C	O	0	0
			7	4	3		
3	A	1	Total	C	O	0	0
			7	4	3		
3	A	1	Total	C	O	0	0
			7	4	3		
3	E	1	Total	C	O	0	0
			7	4	3		
3	L	1	Total	C	O	0	0
			7	4	3		
3	I	1	Total	C	O	0	0
			7	4	3		
3	G	1	Total	C	O	0	0
			7	4	3		
3	J	1	Total	C	O	0	0
			7	4	3		
3	J	1	Total	C	O	0	0
			7	4	3		
3	H	1	Total	C	O	0	0
			7	4	3		
3	K	1	Total	C	O	0	0
			7	4	3		
3	C	1	Total	C	O	0	0
			7	4	3		
3	H	1	Total	C	O	0	0
			7	4	3		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	E	1	Total C O 7 4 3	0	0
3	B	1	Total C O 7 4 3	0	0
3	C	1	Total C O 7 4 3	0	0
3	F	1	Total C O 7 4 3	0	0
3	G	1	Total C O 7 4 3	0	0
3	J	1	Total C O 7 4 3	0	0
3	K	1	Total C O 7 4 3	0	0
3	C	1	Total C O 7 4 3	0	0
3	E	1	Total C O 7 4 3	0	0
3	B	1	Total C O 7 4 3	0	0
3	D	1	Total C O 7 4 3	0	0
3	A	1	Total C O 7 4 3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	133	Total O 133 133	0	0
4	B	133	Total O 133 133	0	0
4	C	122	Total O 122 122	0	0
4	D	117	Total O 117 117	0	0
4	E	130	Total O 130 130	0	0
4	F	116	Total O 116 116	0	0
4	G	102	Total O 102 102	0	0

*Continued on next page...*

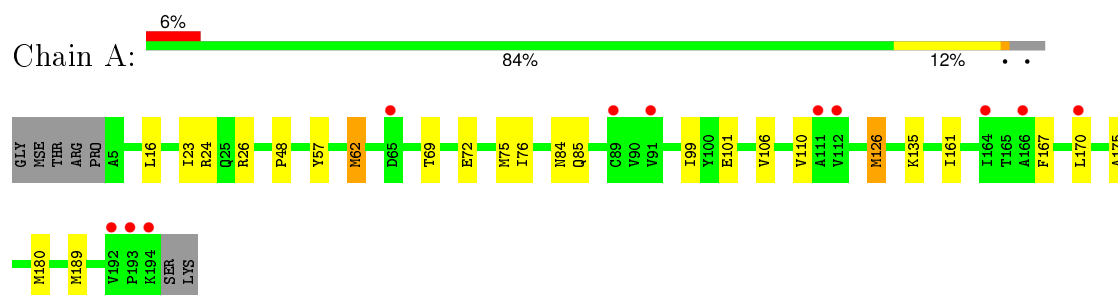
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	130	Total 130	O 130	0	0
4	I	107	Total 107	O 107	0	0
4	J	81	Total 81	O 81	0	0
4	K	88	Total 88	O 88	0	0
4	L	79	Total 79	O 79	0	0

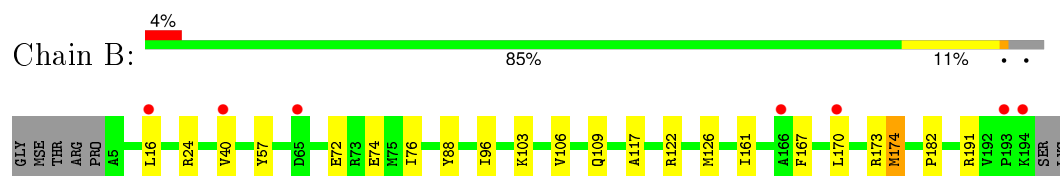
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

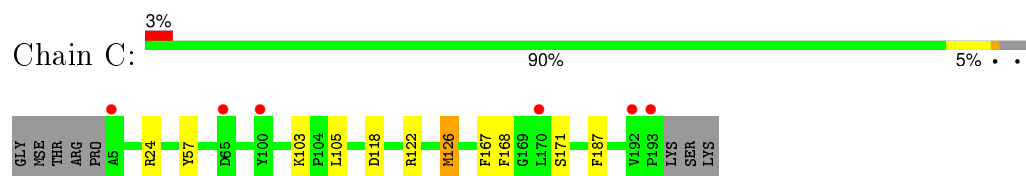
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



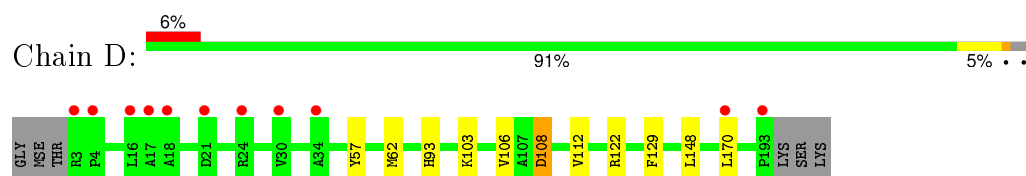
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



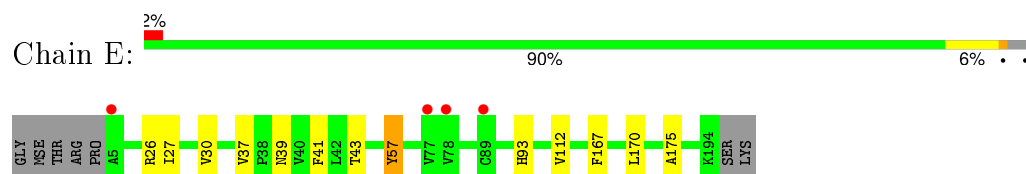
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



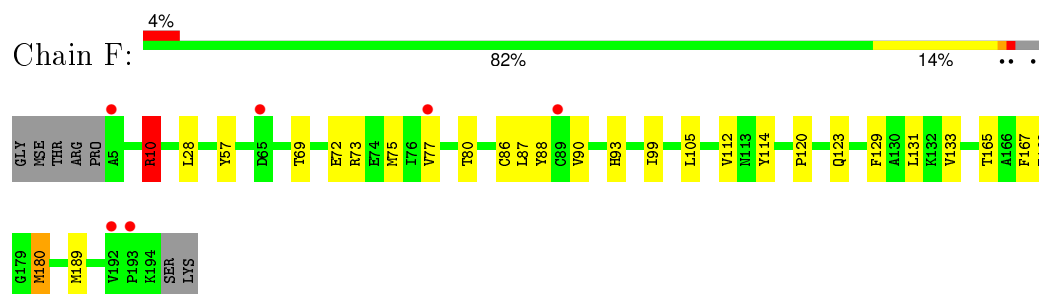
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



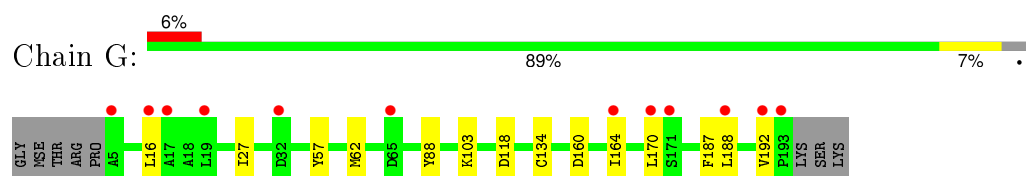
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



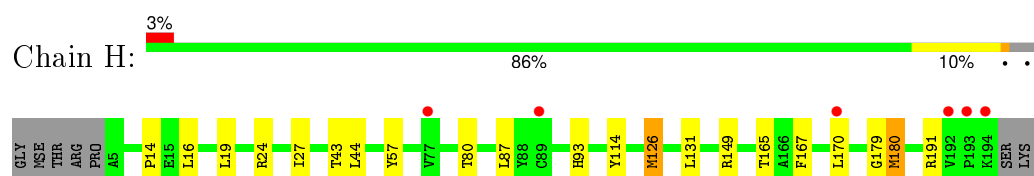
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



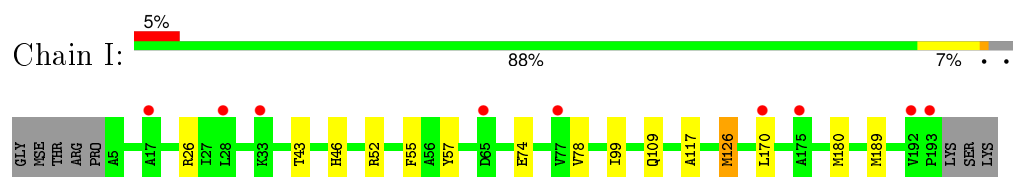
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



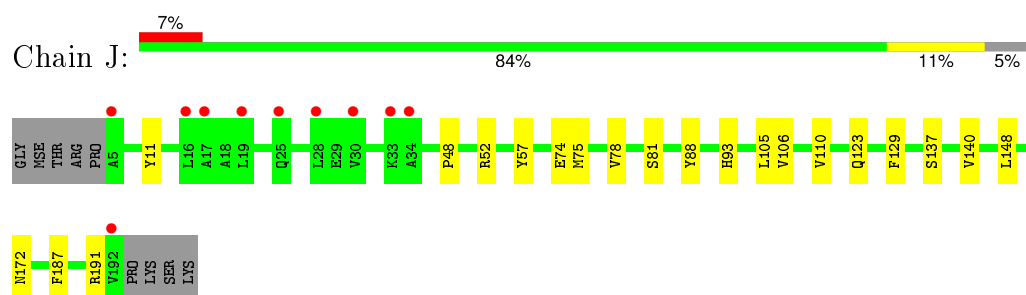
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



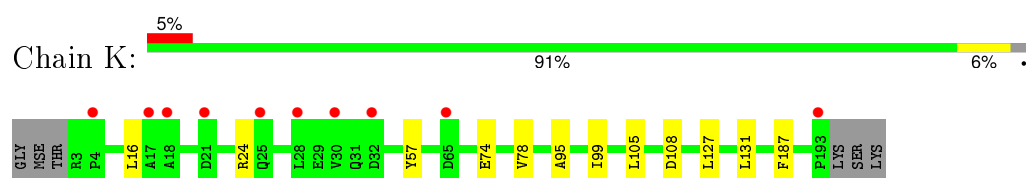
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



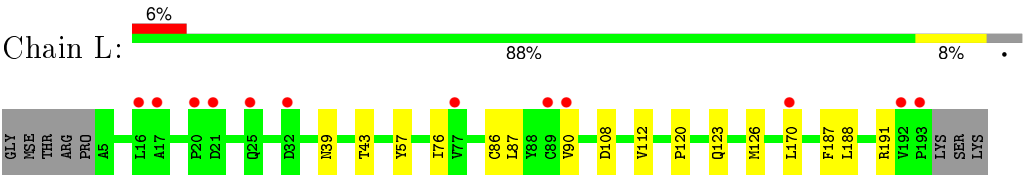
- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



- Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



● Molecule 1: Alkylhydroperoxidase AhpD core: uncharacterized peroxidase-related protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.11Å 93.95Å 120.90Å 84.26° 81.64° 77.73°	Depositor
Resolution (Å)	48.91 – 2.15 29.71 – 2.09	Depositor EDS
% Data completeness (in resolution range)	94.7 (48.91-2.15) 91.7 (29.71-2.09)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.03 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.211 , 0.262 0.213 , 0.262	Depositor DCC
$R_{free}$ test set	7215 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.4	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 55.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 152364 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	19292	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.67% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: PEG, PGE

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	0.76	2/1526 (0.1%)	0.81	2/2060 (0.1%)
1	B	0.79	1/1522 (0.1%)	0.86	4/2057 (0.2%)
1	C	0.69	1/1512 (0.1%)	0.81	2/2041 (0.1%)
1	D	0.74	1/1511 (0.1%)	0.79	0/2042
1	E	0.69	0/1515	0.78	0/2046
1	F	0.72	1/1507 (0.1%)	0.81	2/2036 (0.1%)
1	G	0.74	3/1503 (0.2%)	0.76	0/2030
1	H	0.70	2/1504 (0.1%)	0.80	3/2032 (0.1%)
1	I	0.72	1/1491 (0.1%)	0.81	1/2015 (0.0%)
1	J	0.71	1/1483 (0.1%)	0.75	0/2002
1	K	0.68	0/1493	0.76	0/2020
1	L	0.68	0/1491	0.77	0/2015
All	All	0.72	13/18058 (0.1%)	0.79	14/24396 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	126	MSE	SE-CE	8.90	2.48	1.95
1	I	126	MSE	CG-SE	7.02	2.19	1.95
1	G	88	TYR	CD2-CE2	-6.77	1.29	1.39
1	G	88	TYR	CD1-CE1	-6.74	1.29	1.39
1	C	126	MSE	SE-CE	6.63	2.34	1.95

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	10	ARG	NE-CZ-NH1	7.82	124.21	120.30
1	B	122[A]	ARG	NE-CZ-NH2	6.91	123.75	120.30
1	B	122[B]	ARG	NE-CZ-NH2	6.91	123.75	120.30
1	B	191	ARG	NE-CZ-NH1	5.87	123.23	120.30
1	A	126	MSE	CG-SE-CE	5.81	111.68	98.90

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	1497	0	1475	17	0
1	B	1493	0	1450	17	0
1	C	1483	0	1458	6	0
1	D	1481	0	1450	6	0
1	E	1486	0	1459	11	0
1	F	1478	0	1445	23	0
1	G	1474	0	1438	10	0
1	H	1475	0	1444	14	0
1	I	1462	0	1410	10	0
1	J	1455	0	1406	15	0
1	K	1463	0	1405	6	0
1	L	1462	0	1423	10	0
2	B	10	0	14	0	0
2	C	10	0	14	0	0
2	D	10	0	14	0	0
2	E	10	0	14	0	0
2	F	10	0	14	2	0
2	G	10	0	14	0	0
2	H	10	0	14	0	0
3	A	21	0	30	0	0
3	B	14	0	20	1	0
3	C	21	0	30	0	0
3	D	7	0	10	1	0
3	E	21	0	30	4	0
3	F	14	0	20	1	0
3	G	14	0	20	0	0
3	H	14	0	20	1	0
3	I	7	0	10	0	0
3	J	21	0	30	1	0
3	K	14	0	20	0	0
3	L	7	0	10	0	0
4	A	133	0	0	1	0
4	B	133	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	122	0	0	0	0
4	D	117	0	0	1	0
4	E	130	0	0	1	0
4	F	116	0	0	0	0
4	G	102	0	0	0	0
4	H	130	0	0	0	0
4	I	107	0	0	1	0
4	J	81	0	0	0	0
4	K	88	0	0	0	0
4	L	79	0	0	0	0
All	All	19292	0	17611	132	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 4.

The worst 5 of 132 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:180:MSE:CE	1:A:180:MSE:SE	2.16	1.43
1:I:126:MSE:CG	1:I:126:MSE:SE	2.19	1.41
1:H:180:MSE:CE	1:H:180:MSE:SE	2.25	1.34
1:H:126:MSE:SE	1:H:126:MSE:CE	2.30	1.29
1:A:62:MSE:SE	1:A:62:MSE:CE	2.30	1.28

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	189/197 (96%)	185 (98%)	4 (2%)	0	100	100
1	B	189/197 (96%)	185 (98%)	4 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	187/197 (95%)	185 (99%)	2 (1%)	0	100	100
1	D	189/197 (96%)	184 (97%)	5 (3%)	0	100	100
1	E	188/197 (95%)	184 (98%)	4 (2%)	0	100	100
1	F	188/197 (95%)	186 (99%)	2 (1%)	0	100	100
1	G	187/197 (95%)	185 (99%)	2 (1%)	0	100	100
1	H	188/197 (95%)	184 (98%)	4 (2%)	0	100	100
1	I	187/197 (95%)	181 (97%)	6 (3%)	0	100	100
1	J	186/197 (94%)	183 (98%)	3 (2%)	0	100	100
1	K	189/197 (96%)	186 (98%)	3 (2%)	0	100	100
1	L	187/197 (95%)	182 (97%)	5 (3%)	0	100	100
All	All	2254/2364 (95%)	2210 (98%)	44 (2%)	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	152/154 (99%)	151 (99%)	1 (1%)	88	93
1	B	150/154 (97%)	149 (99%)	1 (1%)	88	93
1	C	150/154 (97%)	149 (99%)	1 (1%)	88	93
1	D	146/154 (95%)	144 (99%)	2 (1%)	74	80
1	E	150/154 (97%)	149 (99%)	1 (1%)	88	93
1	F	147/154 (96%)	144 (98%)	3 (2%)	63	67
1	G	147/154 (96%)	146 (99%)	1 (1%)	88	93
1	H	146/154 (95%)	145 (99%)	1 (1%)	88	93
1	I	144/154 (94%)	143 (99%)	1 (1%)	88	93
1	J	142/154 (92%)	141 (99%)	1 (1%)	88	93
1	K	142/154 (92%)	140 (99%)	2 (1%)	74	80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	L	144/154 (94%)	143 (99%)	1 (1%)	88	93
All	All	1760/1848 (95%)	1744 (99%)	16 (1%)	84	89

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	28	LEU
1	F	57	TYR
1	J	57	TYR
1	F	10	ARG
1	K	57	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

32 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	PEG	A	197	-	6,6,6	0.38	0	5,5,5	0.23	0
3	PEG	A	198	-	6,6,6	0.60	0	5,5,5	0.29	0
3	PEG	A	199	-	6,6,6	0.57	0	5,5,5	0.32	0
2	PGE	B	197	-	9,9,9	0.47	0	8,8,8	0.36	0
3	PEG	B	198	-	6,6,6	0.57	0	5,5,5	0.43	0
3	PEG	B	199	-	6,6,6	0.46	0	5,5,5	0.47	0
2	PGE	C	197	-	9,9,9	0.46	0	8,8,8	0.45	0
3	PEG	C	198	-	6,6,6	0.67	0	5,5,5	0.54	0
3	PEG	C	199	-	6,6,6	0.57	0	5,5,5	0.26	0
3	PEG	C	200	-	6,6,6	0.52	0	5,5,5	0.41	0
2	PGE	D	197	-	9,9,9	0.50	0	8,8,8	0.32	0
3	PEG	D	198	-	6,6,6	0.58	0	5,5,5	0.61	0
2	PGE	E	197	-	9,9,9	0.31	0	8,8,8	0.63	0
3	PEG	E	198	-	6,6,6	0.69	0	5,5,5	0.33	0
3	PEG	E	199	-	6,6,6	0.38	0	5,5,5	0.74	0
3	PEG	E	200	-	6,6,6	0.46	0	5,5,5	0.33	0
2	PGE	F	197	-	9,9,9	0.42	0	8,8,8	0.38	0
3	PEG	F	198	-	6,6,6	0.57	0	5,5,5	0.18	0
3	PEG	F	199	-	6,6,6	0.48	0	5,5,5	0.70	0
2	PGE	G	197	-	9,9,9	0.40	0	8,8,8	0.35	0
3	PEG	G	198	-	6,6,6	0.51	0	5,5,5	0.19	0
3	PEG	G	199	-	6,6,6	0.57	0	5,5,5	0.63	0
2	PGE	H	197	-	9,9,9	0.48	0	8,8,8	0.60	0
3	PEG	H	198	-	6,6,6	0.46	0	5,5,5	0.37	0
3	PEG	H	199	-	6,6,6	0.53	0	5,5,5	0.75	0
3	PEG	I	197	-	6,6,6	0.60	0	5,5,5	0.36	0
3	PEG	J	197	-	6,6,6	0.45	0	5,5,5	0.58	0
3	PEG	J	198	-	6,6,6	0.36	0	5,5,5	0.54	0
3	PEG	J	199	-	6,6,6	0.45	0	5,5,5	0.71	0
3	PEG	K	197	-	6,6,6	0.43	0	5,5,5	0.36	0
3	PEG	K	198	-	6,6,6	0.48	0	5,5,5	0.58	0
3	PEG	L	197	-	6,6,6	0.49	0	5,5,5	0.49	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
3	PEG	A	197	-	-	0/4/4/4	0/0/0/0
3	PEG	A	198	-	-	0/4/4/4	0/0/0/0
3	PEG	A	199	-	-	0/4/4/4	0/0/0/0
2	PGE	B	197	-	-	0/7/7/7	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PEG	B	198	-	-	0/4/4/4	0/0/0/0
3	PEG	B	199	-	-	0/4/4/4	0/0/0/0
2	PGE	C	197	-	-	0/7/7/7	0/0/0/0
3	PEG	C	198	-	-	0/4/4/4	0/0/0/0
3	PEG	C	199	-	-	0/4/4/4	0/0/0/0
3	PEG	C	200	-	-	0/4/4/4	0/0/0/0
2	PGE	D	197	-	-	0/7/7/7	0/0/0/0
3	PEG	D	198	-	-	0/4/4/4	0/0/0/0
2	PGE	E	197	-	-	0/7/7/7	0/0/0/0
3	PEG	E	198	-	-	0/4/4/4	0/0/0/0
3	PEG	E	199	-	-	0/4/4/4	0/0/0/0
3	PEG	E	200	-	-	0/4/4/4	0/0/0/0
2	PGE	F	197	-	-	0/7/7/7	0/0/0/0
3	PEG	F	198	-	-	0/4/4/4	0/0/0/0
3	PEG	F	199	-	-	0/4/4/4	0/0/0/0
2	PGE	G	197	-	-	0/7/7/7	0/0/0/0
3	PEG	G	198	-	-	0/4/4/4	0/0/0/0
3	PEG	G	199	-	-	0/4/4/4	0/0/0/0
2	PGE	H	197	-	-	0/7/7/7	0/0/0/0
3	PEG	H	198	-	-	0/4/4/4	0/0/0/0
3	PEG	H	199	-	-	0/4/4/4	0/0/0/0
3	PEG	I	197	-	-	0/4/4/4	0/0/0/0
3	PEG	J	197	-	-	0/4/4/4	0/0/0/0
3	PEG	J	198	-	-	0/4/4/4	0/0/0/0
3	PEG	J	199	-	-	0/4/4/4	0/0/0/0
3	PEG	K	197	-	-	0/4/4/4	0/0/0/0
3	PEG	K	198	-	-	0/4/4/4	0/0/0/0
3	PEG	L	197	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	198	PEG	1	0
3	D	198	PEG	1	0
3	E	199	PEG	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	200	PEG	1	0
2	F	197	PGE	2	0
3	F	199	PEG	1	0
3	H	199	PEG	1	0
3	J	199	PEG	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	184/197 (93%)	0.20	11 (5%) 25 34	15, 32, 48, 66	0
1	B	184/197 (93%)	0.21	7 (3%) 44 54	12, 33, 52, 74	0
1	C	183/197 (92%)	0.12	6 (3%) 50 60	19, 34, 50, 61	0
1	D	185/197 (93%)	0.30	11 (5%) 26 36	22, 34, 54, 68	0
1	E	184/197 (93%)	0.15	4 (2%) 65 73	22, 33, 53, 69	0
1	F	184/197 (93%)	0.13	7 (3%) 44 54	24, 34, 51, 71	0
1	G	183/197 (92%)	0.35	12 (6%) 22 30	22, 35, 52, 61	0
1	H	184/197 (93%)	0.24	6 (3%) 50 60	23, 33, 53, 71	0
1	I	183/197 (92%)	0.36	9 (4%) 33 44	23, 35, 53, 67	0
1	J	182/197 (92%)	0.46	14 (7%) 16 22	26, 36, 53, 59	0
1	K	185/197 (93%)	0.30	10 (5%) 29 40	21, 36, 54, 69	0
1	L	183/197 (92%)	0.37	12 (6%) 22 30	26, 36, 53, 67	0
All	All	2204/2364 (93%)	0.27	109 (4%) 33 44	12, 34, 53, 74	0

The worst 5 of 109 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	17	ALA	5.7
1	L	192	VAL	4.7
1	B	194	LYS	4.6
1	J	34	ALA	4.6
1	H	193	PRO	4.5

### 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 ⓘ

There are no carbohydrates 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. 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	PEG	C	199	7/7	0.74	0.41	9.51	56,64,70,70	0
3	PEG	J	199	7/7	0.88	0.43	6.67	51,53,58,62	0
3	PEG	G	199	7/7	0.81	0.37	5.34	64,67,68,70	0
2	PGE	H	197	10/10	0.91	0.14	5.25	39,42,46,55	0
3	PEG	K	198	7/7	0.85	0.30	4.43	49,54,63,64	0
3	PEG	F	199	7/7	0.84	0.28	3.40	48,51,64,68	0
3	PEG	H	199	7/7	0.81	0.27	3.19	44,50,65,65	0
2	PGE	B	197	10/10	0.91	0.23	3.04	34,46,52,56	0
2	PGE	E	197	10/10	0.90	0.24	2.67	36,40,45,46	0
2	PGE	C	197	10/10	0.92	0.23	2.55	38,45,50,50	0
3	PEG	A	198	7/7	0.87	0.20	2.52	36,47,56,56	0
3	PEG	E	198	7/7	0.86	0.16	2.48	46,46,51,52	0
2	PGE	G	197	10/10	0.89	0.22	2.23	47,59,65,66	0
3	PEG	B	198	7/7	0.88	0.23	2.22	42,44,53,58	0
3	PEG	G	198	7/7	0.88	0.17	1.96	44,48,61,63	0
3	PEG	A	199	7/7	0.90	0.22	1.76	42,48,60,63	0
3	PEG	I	197	7/7	0.84	0.22	1.76	45,50,55,56	0
3	PEG	F	198	7/7	0.77	0.18	1.68	42,59,64,66	0
3	PEG	B	199	7/7	0.84	0.18	1.64	41,47,63,63	0
3	PEG	C	200	7/7	0.91	0.16	1.63	49,56,68,76	0
3	PEG	J	198	7/7	0.92	0.14	1.57	38,41,47,53	0
3	PEG	E	199	7/7	0.92	0.20	1.31	43,46,52,54	0
2	PGE	F	197	10/10	0.90	0.20	1.31	30,39,53,55	0
2	PGE	D	197	10/10	0.91	0.20	0.85	31,38,48,49	0
3	PEG	C	198	7/7	0.92	0.14	0.71	27,34,44,47	0
3	PEG	L	197	7/7	0.91	0.14	0.45	50,63,68,75	0
3	PEG	J	197	7/7	0.93	0.12	0.38	37,42,45,46	0
3	PEG	E	200	7/7	0.90	0.14	0.37	46,49,55,57	0
3	PEG	A	197	7/7	0.90	0.13	0.34	35,40,43,51	0
3	PEG	D	198	7/7	0.91	0.14	0.01	34,36,44,54	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PEG	K	197	7/7	0.87	0.14	-	46,51,54,56	0
3	PEG	H	198	7/7	0.91	0.14	-	38,42,47,49	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.