



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2019 – 03:10 PM EDT

PDB ID : 4XRA  
Title : Salmonella typhimurium AhpC T43S mutant  
Authors : Perkins, A.; Brereton, A.E.; Nelson, K.; Parsonage, D.; Poole, L.; Karplus, P.A.  
Deposited on : 2015-01-20  
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

<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 : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.4  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

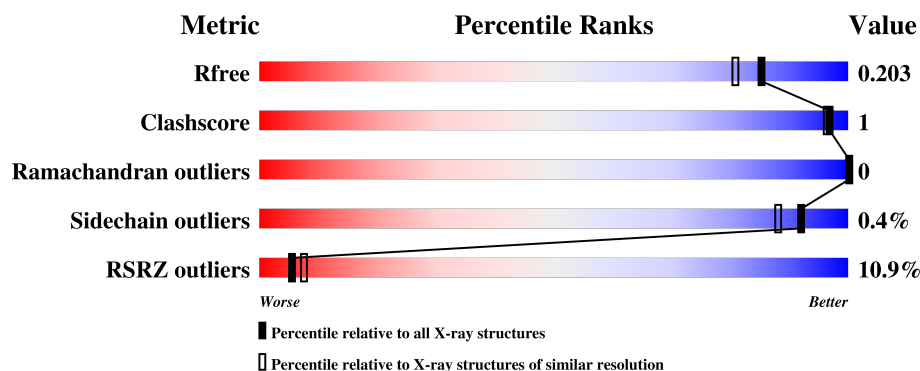
# 1 Overall quality at a glance

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))
$R_{free}$	111664	1952 (1.76-1.76)
Clashscore	122126	2072 (1.76-1.76)
Ramachandran outliers	120053	2050 (1.76-1.76)
Sidechain outliers	120020	2050 (1.76-1.76)
RSRZ outliers	108989	1913 (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	186	<div> <div>5%</div> <div>89%</div> <div>10%</div> </div>
1	B	186	<div> <div>11%</div> <div>97%</div> <div>.</div> </div>
1	C	186	<div> <div>10%</div> <div>98%</div> <div>..</div> </div>
1	D	186	<div> <div>11%</div> <div>97%</div> <div>.</div> </div>
1	E	186	<div> <div>16%</div> <div>99%</div> <div>.</div> </div>

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	CL	A	202	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15227 atoms, of which 7067 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 Alkyl hydroperoxide reductase subunit C.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	168	Total	C	H	N	O	S	0	9	0
			2651	869	1283	226	268	5			
1	B	186	Total	C	H	N	O	S	0	9	0
			2963	956	1461	250	292	4			
1	C	186	Total	C	H	N	O	S	0	9	0
			2947	949	1452	247	295	4			
1	D	186	Total	C	H	N	O	S	0	7	0
			2931	945	1442	245	295	4			
1	E	186	Total	C	H	N	O	S	0	2	0
			2900	939	1429	242	286	4			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	43	SER	THR	engineered mutation	UNP P0A251
B	43	SER	THR	engineered mutation	UNP P0A251
C	43	SER	THR	engineered mutation	UNP P0A251
D	43	SER	THR	engineered mutation	UNP P0A251
E	43	SER	THR	engineered mutation	UNP P0A251

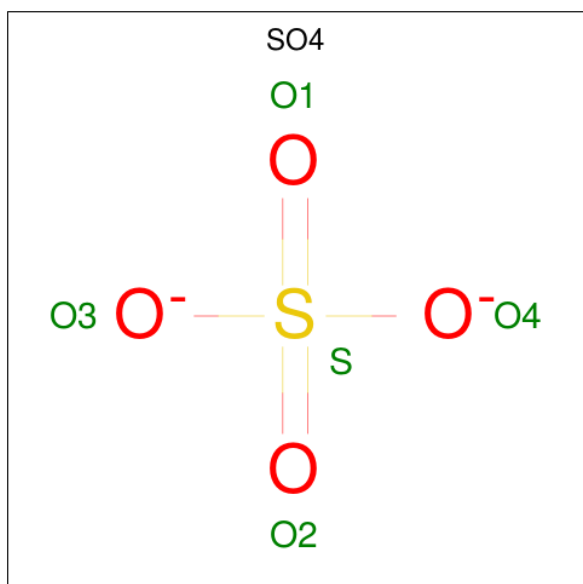
- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Cl	0	0
			1	1		
2	A	3	Total	Cl	0	0
			3	3		
2	D	2	Total	Cl	0	1
			3	3		
2	C	2	Total	Cl	0	1
			3	3		
2	E	1	Total	Cl	0	0
			1	1		

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total K 1 1	0	0
3	A	1	Total K 1 1	0	0
3	D	1	Total K 1 1	0	0

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



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total O S 5 4 1	0	0

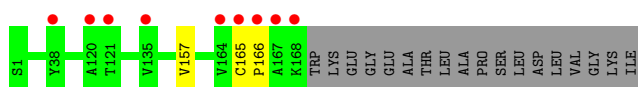
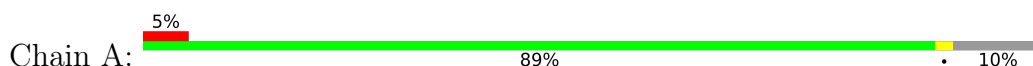
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	174	Total O 179 179	0	6
5	B	174	Total O 180 180	0	7
5	C	141	Total O 148 148	0	7
5	D	189	Total O 195 195	0	6
5	E	112	Total O 114 114	0	3

### 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.

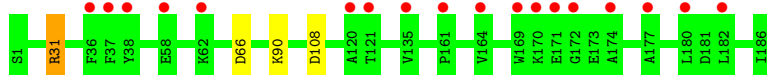
- Molecule 1: Alkyl hydroperoxide reductase subunit C



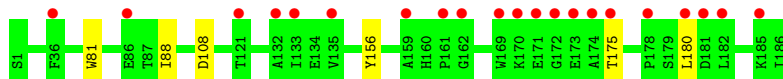
- Molecule 1: Alkyl hydroperoxide reductase subunit C



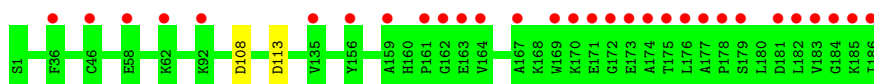
- Molecule 1: Alkyl hydroperoxide reductase subunit C



- Molecule 1: Alkyl hydroperoxide reductase subunit C



- Molecule 1: Alkyl hydroperoxide reductase subunit C



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.07Å 171.92Å 135.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.28 – 1.75 36.28 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.9 (36.28-1.75) 99.9 (36.28-1.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.12 (at 1.75Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.171 , 0.200 0.175 , 0.203	Depositor DCC
$R_{free}$ test set	7380 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.1	Xtriage
Anisotropy	0.156	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 47.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	15227	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.20% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: K, SO4, CL

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	0.57	0/1423	0.63	0/1929
1	B	0.56	0/1558	0.62	0/2111
1	C	0.53	0/1562	0.61	0/2114
1	D	0.53	0/1547	0.59	0/2095
1	E	0.47	0/1509	0.55	0/2044
All	All	0.53	0/7599	0.60	0/10293

There are no bond length outliers.

There are no bond angle outliers.

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	1368	1283	1299	3	0
1	B	1502	1461	1434	3	0
1	C	1495	1452	1423	2	0
1	D	1489	1442	1423	5	0
1	E	1471	1429	1428	2	0
2	A	3	0	0	3	0
2	B	1	0	0	0	0
2	C	3	0	0	0	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	3	0	0	0	0
2	E	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
4	B	5	0	0	0	0
5	A	179	0	0	2	0
5	B	180	0	0	0	0
5	C	148	0	0	1	0
5	D	195	0	0	2	0
5	E	114	0	0	1	0
All	All	8160	7067	7007	15	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 1.

All (15) 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)
2:A:202:CL:CL	5:A:321:HOH:O	2.09	1.07
2:A:202:CL:CL	5:A:322[B]:HOH:O	2.17	1.00
1:D:108[A]:ASP:OD2	5:D:420[A]:HOH:O	2.04	0.74
1:D:108[B]:ASP:OD2	5:D:420[B]:HOH:O	2.10	0.69
1:A:165[B]:CYS:SG	2:A:203:CL:CL	2.90	0.55
1:C:31:ARG:HD3	1:C:66:ASP:OD2	2.11	0.51
1:E:108:ASP:OD2	5:E:301:HOH:O	2.20	0.48
1:B:19:GLU:HA	1:D:180:LEU:HD13	1.97	0.46
1:B:109:ASN:OD1	1:B:118:ASP:HB2	2.17	0.45
1:A:165[B]:CYS:HA	1:A:166:PRO:HD3	1.81	0.43
1:D:81:TRP:CD2	1:D:88:ILE:HG13	2.54	0.42
1:A:157[A]:VAL:HG11	1:B:139:GLY:HA3	2.02	0.41
1:C:108[A]:ASP:OD2	5:C:338:HOH:O	2.22	0.41
1:D:156:TYR:CE1	1:D:175:THR:HG21	2.54	0.41
1:E:113[A]:ASP:OD1	1:E:113[A]:ASP:N	2.52	0.41

There are no symmetry-related clashes.

## 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	175/186 (94%)	168 (96%)	7 (4%)	0	100	100
1	B	193/186 (104%)	191 (99%)	2 (1%)	0	100	100
1	C	193/186 (104%)	187 (97%)	6 (3%)	0	100	100
1	D	191/186 (103%)	186 (97%)	5 (3%)	0	100	100
1	E	186/186 (100%)	183 (98%)	3 (2%)	0	100	100
All	All	938/930 (101%)	915 (98%)	23 (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	149/154 (97%)	149 (100%)	0	100	100
1	B	160/154 (104%)	159 (99%)	1 (1%)	87	80
1	C	162/154 (105%)	160 (99%)	2 (1%)	74	59
1	D	161/154 (104%)	161 (100%)	0	100	100
1	E	156/154 (101%)	156 (100%)	0	100	100
All	All	788/770 (102%)	785 (100%)	3 (0%)	92	88

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	84	SER
1	C	31	ARG
1	C	90	LYS

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

Mol	Chain	Res	Type
1	D	127	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 15 ligands modelled in this entry, 14 are monoatomic - leaving 1 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	$\# Z  > 2$
4	SO4	B	201	-	4,4,4	0.15	0	6,6,6	0.24	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.

No monomer is involved in short contacts.

## 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	168/186 (90%)	0.21	9 (5%) 26 32	24, 34, 56, 114	0
1	B	186/186 (100%)	0.22	21 (11%) 5 6	23, 36, 65, 92	0
1	C	186/186 (100%)	0.41	18 (9%) 8 10	24, 40, 67, 90	0
1	D	186/186 (100%)	0.30	21 (11%) 5 6	26, 39, 59, 82	0
1	E	186/186 (100%)	0.65	30 (16%) 1 3	31, 51, 79, 112	0
All	All	912/930 (98%)	0.36	99 (10%) 5 7	23, 39, 70, 114	0

All (99) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	165[A]	CYS	8.6
1	E	174	ALA	8.1
1	A	164	VAL	6.0
1	A	167	ALA	5.8
1	E	172	GLY	5.8
1	B	174	ALA	5.6
1	B	46	CYS	5.4
1	D	172	GLY	5.1
1	D	174	ALA	4.9
1	E	185	LYS	4.8
1	B	182[A]	LEU	4.8
1	C	174	ALA	4.8
1	E	169	TRP	4.8
1	B	172	GLY	4.6
1	C	171[A]	GLU	4.5
1	E	178	PRO	4.3
1	B	180	LEU	4.3
1	D	180	LEU	4.2
1	E	156	TYR	4.0
1	B	181[A]	ASP	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	62	LYS	3.9
1	E	173	GLU	3.9
1	E	181	ASP	3.8
1	C	170	LYS	3.7
1	E	170	LYS	3.7
1	E	175	THR	3.7
1	B	177	ALA	3.6
1	E	171	GLU	3.6
1	C	172[A]	GLY	3.6
1	A	168	LYS	3.6
1	B	169	TRP	3.5
1	E	167	ALA	3.5
1	E	176	LEU	3.4
1	E	159	ALA	3.3
1	B	175	THR	3.2
1	D	181	ASP	3.2
1	E	164	VAL	3.2
1	E	186	ILE	3.1
1	E	182	LEU	3.1
1	E	58	GLU	3.1
1	B	170	LYS	3.0
1	B	185	LYS	3.0
1	C	62	LYS	3.0
1	C	121	THR	3.0
1	C	164	VAL	3.0
1	B	86	GLU	2.9
1	A	166	PRO	2.9
1	B	179	SER	2.9
1	D	185	LYS	2.9
1	C	37	PHE	2.8
1	C	135	VAL	2.8
1	D	178	PRO	2.8
1	C	120	ALA	2.8
1	E	179	SER	2.8
1	D	175	THR	2.8
1	C	36	PHE	2.8
1	C	38	TYR	2.7
1	C	182	LEU	2.7
1	C	169	TRP	2.6
1	B	178	PRO	2.6
1	E	184	GLY	2.6
1	B	161	PRO	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	169	TRP	2.5
1	A	120	ALA	2.5
1	D	133	ILE	2.5
1	B	44	PHE	2.5
1	C	177	ALA	2.4
1	D	36	PHE	2.4
1	D	161	PRO	2.4
1	E	161	PRO	2.4
1	E	183	VAL	2.4
1	D	170	LYS	2.4
1	D	182	LEU	2.3
1	B	183	VAL	2.3
1	D	135	VAL	2.3
1	D	173	GLU	2.3
1	E	162	GLY	2.3
1	D	132	ALA	2.3
1	E	177	ALA	2.3
1	E	92	LYS	2.3
1	E	135	VAL	2.2
1	E	36	PHE	2.2
1	B	171	GLU	2.2
1	D	171	GLU	2.2
1	C	161	PRO	2.2
1	A	38	TYR	2.2
1	B	173	GLU	2.2
1	D	121	THR	2.2
1	B	186	ILE	2.1
1	B	162	GLY	2.1
1	D	162	GLY	2.1
1	D	86[A]	GLU	2.1
1	E	163	GLU	2.1
1	A	135	VAL	2.1
1	E	46	CYS	2.1
1	A	121	THR	2.0
1	C	58	GLU	2.0
1	D	159	ALA	2.0
1	C	180	LEU	2.0

## 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(Å <sup>2</sup> )	Q<0.9
2	CL	A	203	1/1	0.78	0.19	74,74,74,74	1
2	CL	A	202	1/1	0.80	0.30	80,80,80,80	1
4	SO4	B	201	5/5	0.89	0.19	101,103,105,108	0
2	CL	C	202[B]	1/1	0.92	0.22	57,57,57,57	1
2	CL	C	202[A]	1/1	0.92	0.22	53,53,53,53	1
2	CL	D	202[A]	1/1	0.96	0.18	60,60,60,60	1
2	CL	D	202[B]	1/1	0.96	0.18	56,56,56,56	1
2	CL	E	201	1/1	0.99	0.03	47,47,47,47	0
3	K	B	203	1/1	0.99	0.14	52,52,52,52	0
2	CL	B	202	1/1	0.99	0.04	38,38,38,38	0
2	CL	D	201	1/1	0.99	0.06	40,40,40,40	0
3	K	A	204	1/1	0.99	0.05	52,52,52,52	1
2	CL	A	201	1/1	0.99	0.07	34,34,34,34	0
3	K	D	203	1/1	0.99	0.22	60,60,60,60	0
2	CL	C	201	1/1	0.99	0.08	42,42,42,42	0

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

There are no such residues in this entry.