



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 12, 2024 – 01:23 AM EDT

PDB ID : 1D1Z
Title : CRYSTAL STRUCTURE OF THE XLP PROTEIN SAP
Authors : Poy, F.; Yaffe, M.B.; Sayos, J.; Saxena, K.; Eck, M.J.
Deposited on : 1999-09-22
Resolution : 1.40 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.20.1
EDS : 2.36.2
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) : 2.36.2

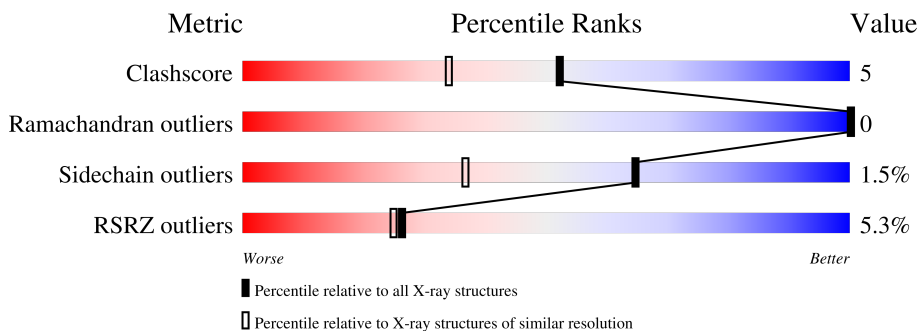
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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))
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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	104	 3% (Poor fit), 86% (0-1 outliers), 11% (2-3 outliers), 0% (4+ outliers), 0% (Not modelled)
1	B	104	 7% (Poor fit), 76% (0-1 outliers), 12% (2-3 outliers), 10% (4+ outliers), 0% (Not modelled)
1	C	104	 7% (Poor fit), 80% (0-1 outliers), 17% (2-3 outliers), 0% (4+ outliers), 0% (Not modelled)
1	D	104	 4% (Poor fit), 84% (0-1 outliers), 13% (2-3 outliers), 0% (4+ outliers), 0% (Not modelled)

2 Entry composition [i](#)

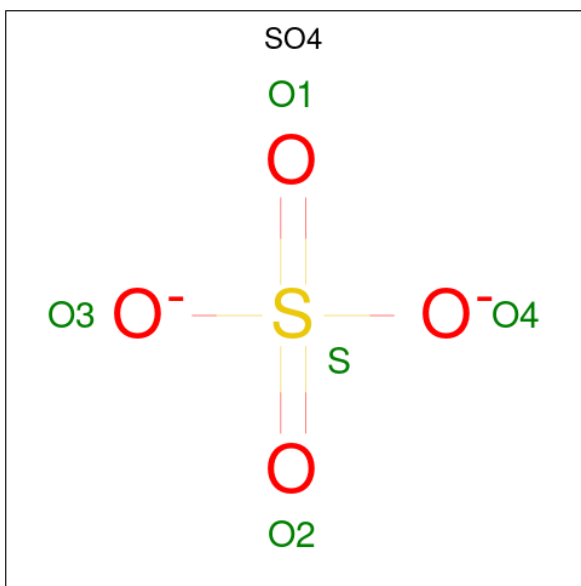
There are 3 unique types of molecules in this entry. The entry contains 3823 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 SAP SH2 DOMAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	101	Total 804	C 519	N 135	O 148	S 2	0	0	0
1	B	94	Total 748	C 485	N 126	O 135	S 2	0	0	0
1	C	101	Total 804	C 519	N 135	O 148	S 2	0	0	0
1	D	101	Total 804	C 519	N 135	O 148	S 2	0	0	0

- Molecule 2 is SULFATE ION (three-letter code: SO₄) (formula: O₄S).



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

Continued on next page...

Continued from previous page...

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

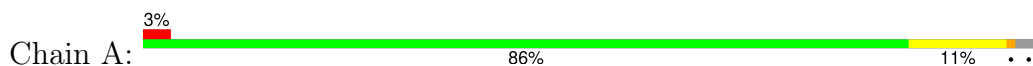
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	170	Total	O	0	0
			170	170		
3	B	152	Total	O	0	0
			152	152		
3	C	153	Total	O	0	0
			153	153		
3	D	153	Total	O	0	0
			153	153		

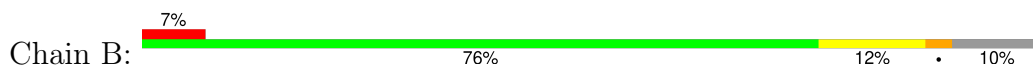
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

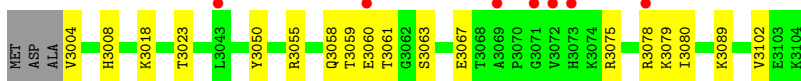
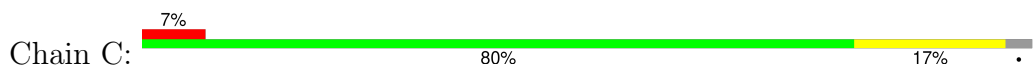
- Molecule 1: SAP SH2 DOMAIN



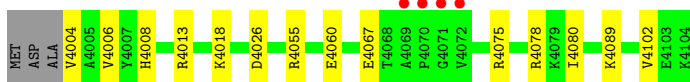
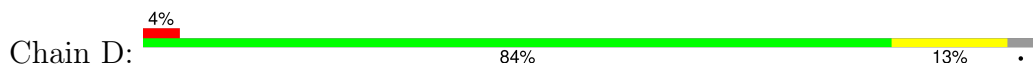
- Molecule 1: SAP SH2 DOMAIN



- Molecule 1: SAP SH2 DOMAIN



- Molecule 1: SAP SH2 DOMAIN



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	32.48Å 62.27Å 65.09Å 65.71° 78.67° 89.66°	Depositor
Resolution (Å)	10.00 – 1.40 21.95 – 1.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-1.40) 92.8 (21.95-1.40)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.74 (at 1.40Å)	Xtrriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.139 , 0.187 0.146 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	14.1	Xtrriage
Anisotropy	0.343	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.45 , 75.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3823	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.63% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: SO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.57	0/824	1.14	5/1116 (0.4%)
1	B	0.61	0/765	1.32	6/1033 (0.6%)
1	C	0.54	0/824	1.27	5/1116 (0.4%)
1	D	0.59	0/824	1.25	6/1116 (0.5%)
All	All	0.58	0/3237	1.25	22/4381 (0.5%)

There are no bond length outliers.

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	2055	ARG	NE-CZ-NH2	12.67	126.63	120.30
1	C	3055	ARG	NE-CZ-NH2	9.56	125.08	120.30
1	D	4013	ARG	NE-CZ-NH1	9.17	124.89	120.30
1	C	3075	ARG	NE-CZ-NH2	-8.71	115.95	120.30
1	B	2013	ARG	NE-CZ-NH1	8.07	124.34	120.30
1	B	2075	ARG	NE-CZ-NH2	-7.47	116.56	120.30
1	A	1032	ARG	NE-CZ-NH2	-6.66	116.97	120.30
1	C	3050	TYR	CB-CG-CD2	-6.64	117.02	121.00
1	D	4055	ARG	CD-NE-CZ	6.44	132.61	123.60
1	D	4075	ARG	CD-NE-CZ	-6.37	114.69	123.60
1	D	4013	ARG	NH1-CZ-NH2	-6.21	112.58	119.40
1	C	3059	THR	O-C-N	6.19	132.61	122.70
1	D	4026	ASP	CB-CG-OD1	5.86	123.57	118.30
1	A	1026	ASP	CB-CG-OD2	-5.70	113.17	118.30
1	A	1013	ARG	CD-NE-CZ	5.65	131.51	123.60
1	B	2075	ARG	NE-CZ-NH1	5.57	123.08	120.30
1	B	2055	ARG	NH1-CZ-NH2	-5.54	113.30	119.40
1	A	1035	GLU	OE1-CD-OE2	-5.16	117.11	123.30
1	D	4055	ARG	NE-CZ-NH2	5.15	122.88	120.30
1	A	1026	ASP	CB-CG-OD1	5.13	122.91	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	3078	ARG	NE-CZ-NH1	-5.10	117.75	120.30
1	B	2067	GLU	CA-CB-CG	5.03	124.46	113.40

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	804	0	807	6	0
1	B	748	0	760	12	0
1	C	804	0	807	10	0
1	D	804	0	807	9	0
2	A	5	0	0	0	0
2	B	10	0	0	0	0
2	C	5	0	0	0	0
2	D	15	0	0	1	0
3	A	170	0	0	4	0
3	B	152	0	0	6	0
3	C	153	0	0	3	0
3	D	153	0	0	4	0
All	All	3823	0	3181	35	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (35) 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:D:4089:LYS:HD2	3:D:6274:HOH:O	1.86	0.74
1:C:3060:GLU:HG2	1:C:3061:THR:HG23	1.78	0.63
1:C:3008:HIS:HE1	1:C:3102:VAL:O	1.83	0.62
1:D:4008:HIS:HE1	1:D:4102:VAL:O	1.84	0.61
1:B:2008:HIS:HE1	1:B:2102:VAL:O	1.83	0.61
1:C:3067:GLU:HG2	3:C:6320:HOH:O	2.01	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:3060:GLU:HG2	1:C:3061:THR:N	2.15	0.60
1:B:2080:ILE:HD13	3:B:6539:HOH:O	2.04	0.57
1:C:3018:LYS:HG3	3:C:6459:HOH:O	2.04	0.56
1:B:2018:LYS:HD3	3:B:6170:HOH:O	2.05	0.56
1:C:3058:GLN:HA	1:C:3063:SER:O	2.08	0.54
1:B:2004:VAL:N	3:B:6579:HOH:O	2.45	0.50
1:C:3023:THR:O	1:D:4008:HIS:HD2	1.96	0.48
1:B:2079:LYS:HB2	1:B:2082:ASN:ND2	2.29	0.48
1:A:1059:THR:OG1	1:A:1061:THR:HG22	2.15	0.47
1:D:4004:VAL:HG21	1:D:4080:ILE:HG21	1.97	0.47
1:A:1018:LYS:HG3	3:A:6197:HOH:O	2.15	0.46
1:D:4006:VAL:HG13	3:D:6268:HOH:O	2.15	0.46
1:A:1055:ARG:NE	3:A:6468:HOH:O	2.49	0.46
1:C:3060:GLU:CG	1:C:3061:THR:HG23	2.46	0.46
1:C:3004:VAL:HG21	1:C:3080:ILE:HG23	1.98	0.45
1:B:2065:SER:N	3:B:6549:HOH:O	2.50	0.45
1:A:1004:VAL:N	3:A:6528:HOH:O	2.50	0.45
1:B:2089:LYS:NZ	3:B:6485:HOH:O	2.50	0.44
1:B:2040:VAL:HG22	1:B:2057:SER:HB2	2.00	0.43
1:D:4060:GLU:HG3	2:D:5006:SO4:S	2.58	0.43
1:B:2067:GLU:OE2	1:D:4067:GLU:OE1	2.37	0.43
1:A:1104:LYS:NZ	3:A:6531:HOH:O	2.52	0.42
1:C:3079:LYS:NZ	3:C:6538:HOH:O	2.50	0.42
1:D:4089:LYS:HD2	3:D:6419:HOH:O	2.20	0.41
1:B:2004:VAL:HA	3:B:6626:HOH:O	2.20	0.41
1:B:2004:VAL:HG21	1:B:2080:ILE:HG22	2.02	0.41
1:D:4018:LYS:HG3	3:D:6529:HOH:O	2.20	0.41
1:A:1058:GLN:HA	1:A:1063:SER:O	2.20	0.41
1:B:2080:ILE:O	1:B:2080:ILE:HG23	2.21	0.40

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	99/104 (95%)	97 (98%)	2 (2%)	0	100	100
1	B	90/104 (86%)	87 (97%)	3 (3%)	0	100	100
1	C	99/104 (95%)	97 (98%)	2 (2%)	0	100	100
1	D	99/104 (95%)	97 (98%)	2 (2%)	0	100	100
All	All	387/416 (93%)	378 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

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	87/89 (98%)	86 (99%)	1 (1%)	73	50
1	B	81/89 (91%)	79 (98%)	2 (2%)	47	14
1	C	87/89 (98%)	86 (99%)	1 (1%)	73	50
1	D	87/89 (98%)	86 (99%)	1 (1%)	73	50
All	All	342/356 (96%)	337 (98%)	5 (2%)	65	37

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1055	ARG
1	B	2079	LYS
1	B	2080	ILE
1	C	3089	LYS
1	D	4078	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1099	GLN
1	B	2008	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	3008	HIS
1	D	4008	HIS
1	D	4048	HIS
1	D	4058	GLN
1	D	4082	ASN
1	D	4088	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

7 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	SO4	C	5003	-	4,4,4	0.33	0	6,6,6	0.23	0
2	SO4	D	5006	-	4,4,4	0.37	0	6,6,6	0.12	0
2	SO4	D	5004	-	4,4,4	0.30	0	6,6,6	0.36	0
2	SO4	A	5001	-	4,4,4	0.50	0	6,6,6	0.33	0
2	SO4	D	5005	-	4,4,4	0.32	0	6,6,6	0.13	0
2	SO4	B	5007	-	4,4,4	0.31	0	6,6,6	0.17	0
2	SO4	B	5002	-	4,4,4	0.42	0	6,6,6	0.46	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	5006	SO4	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, 95th 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 ²)	Q<0.9
1	A	101/104 (97%)	-0.11	3 (2%) 50 49	11, 16, 30, 39	0
1	B	94/104 (90%)	0.14	7 (7%) 14 13	11, 16, 39, 46	0
1	C	101/104 (97%)	0.33	7 (6%) 16 15	12, 19, 39, 52	0
1	D	101/104 (97%)	0.06	4 (3%) 38 39	10, 18, 31, 41	0
All	All	397/416 (95%)	0.11	21 (5%) 26 25	10, 18, 36, 52	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	3078	ARG	5.1
1	C	3073	HIS	5.0
1	D	4071	GLY	4.3
1	B	2069	ALA	4.2
1	B	2073	HIS	4.1
1	C	3060	GLU	3.7
1	B	2071	GLY	3.6
1	B	2078	ARG	3.6
1	B	2072	VAL	3.5
1	D	4070	PRO	3.1
1	A	1072	VAL	2.8
1	C	3072	VAL	2.7
1	B	2070	PRO	2.6
1	C	3071	GLY	2.5
1	A	1061	THR	2.3
1	A	1071	GLY	2.3
1	D	4072	VAL	2.3
1	C	3043	LEU	2.2
1	D	4069	ALA	2.2
1	B	2080	ILE	2.1
1	C	3069	ALA	2.1

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 monosaccharides 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, 95th 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(\AA^2)	Q<0.9
2	SO4	B	5007	5/5	0.98	0.13	33,37,43,44	0
2	SO4	D	5006	5/5	0.98	0.15	30,31,36,37	0
2	SO4	A	5001	5/5	0.99	0.06	16,18,24,27	0
2	SO4	D	5004	5/5	0.99	0.05	15,16,22,22	0
2	SO4	D	5005	5/5	0.99	0.04	17,17,18,20	0
2	SO4	B	5002	5/5	0.99	0.04	17,19,24,26	0
2	SO4	C	5003	5/5	1.00	0.04	17,17,22,23	0

6.5 Other polymers [i](#)

There are no such residues in this entry.