



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 27, 2023 – 09:14 PM EDT

PDB ID : 3HLT  
Title : The crystal structure of human haloacid dehalogenase-like hydrolase domain containing 2 (HDHD2)  
Authors : Ugochukwu, E.; Shafqat, N.; Yue, W.W.; Cocking, R.; Bray, J.E.; Muniz, J.R.C.; Krojer, T.; von Delft, F.; Bountra, C.; Arrowsmith, C.H.; Weigelt, J.; Edwards, A.; Oppermann, U.; Structural Genomics Consortium (SGC)  
Deposited on : 2009-05-28  
Resolution : 2.30 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 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

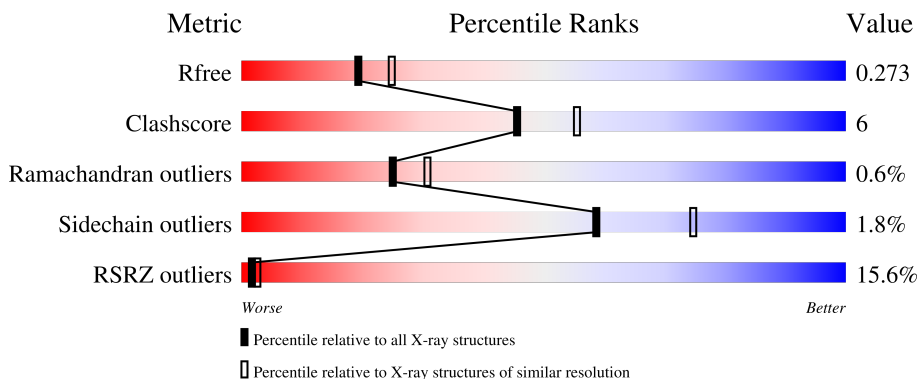
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	266	 6% 80% 16%
1	C	266	 24% 78% 14% 9%

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 3742 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 HDHD2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	256	1947	1242	336	362	7	0	3	0
1	C	243	1717	1095	298	317	7	0	2	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	260	ALA	-	expression tag	UNP Q9H0R4
A	261	GLU	-	expression tag	UNP Q9H0R4
A	262	ASN	-	expression tag	UNP Q9H0R4
A	263	LEU	-	expression tag	UNP Q9H0R4
A	264	TYR	-	expression tag	UNP Q9H0R4
A	265	PHE	-	expression tag	UNP Q9H0R4
A	266	GLN	-	expression tag	UNP Q9H0R4
C	260	ALA	-	expression tag	UNP Q9H0R4
C	261	GLU	-	expression tag	UNP Q9H0R4
C	262	ASN	-	expression tag	UNP Q9H0R4
C	263	LEU	-	expression tag	UNP Q9H0R4
C	264	TYR	-	expression tag	UNP Q9H0R4
C	265	PHE	-	expression tag	UNP Q9H0R4
C	266	GLN	-	expression tag	UNP Q9H0R4

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Na 1 1	0	0
4	C	1	Total Na 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	53	Total O 53 53	0	0

*Continued on next page...*

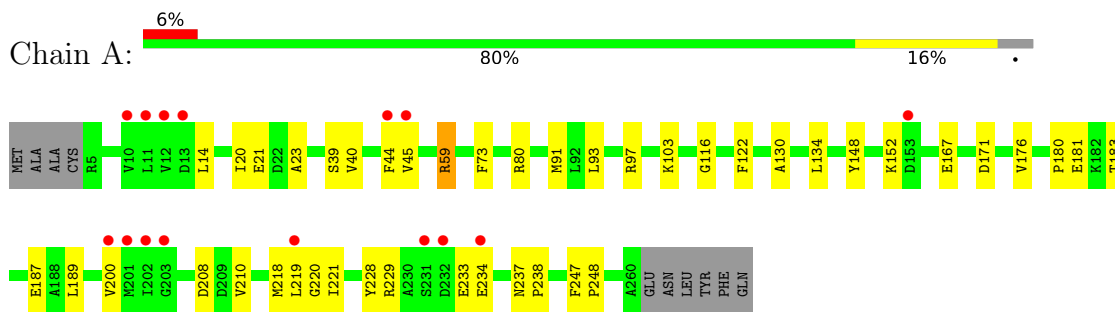
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	C	11	Total	O	0	0
			11	11		

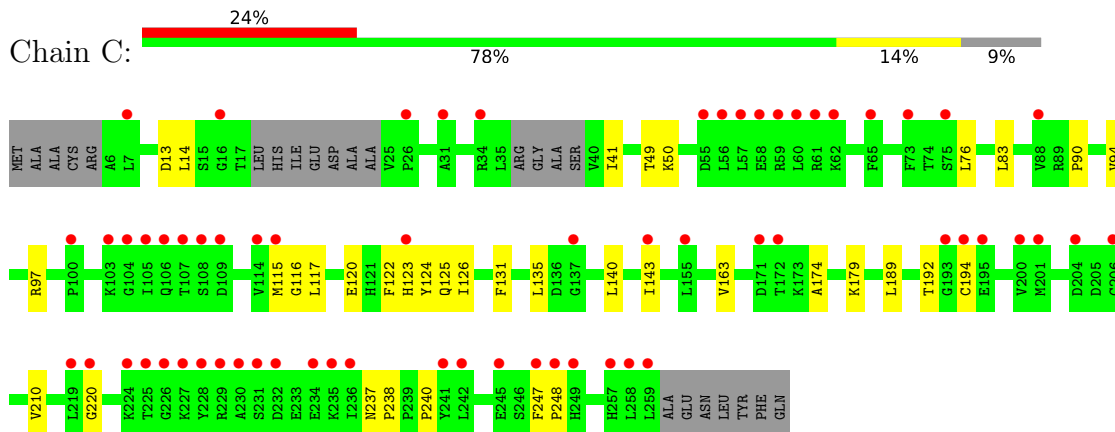
### 3 Residue-property plots [i](#)

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: HDHD2



- Molecule 1: HDHD2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.94Å 77.68Å 103.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.36 – 2.30 28.36 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.5 (28.36-2.30) 99.5 (28.36-2.30)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.65 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.5.0089	Depositor
R, $R_{free}$	0.213 , 0.266 0.220 , 0.273	Depositor DCC
$R_{free}$ test set	1264 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	55.1	Xtrriage
Anisotropy	0.099	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 53.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3742	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.64% 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: SO4, CL, NA

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	0.71	0/1992	0.74	0/2704
1	C	0.52	0/1755	0.62	0/2394
All	All	0.63	0/3747	0.69	0/5098

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

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	1947	0	1922	23	0
1	C	1717	0	1552	24	0
2	A	5	0	0	0	0
2	C	5	0	0	0	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
5	A	53	0	0	0	0
5	C	11	0	0	0	1
All	All	3742	0	3474	46	1



The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (46) 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:59:ARG:HH11	1:A:59:ARG:HG2	1.33	0.91
1:A:200:VAL:HG22	1:A:219:LEU:HB2	1.57	0.84
1:A:189:LEU:HD11	1:A:218:MET:HE2	1.76	0.68
1:A:59:ARG:HH11	1:A:59:ARG:CG	2.08	0.66
1:C:192:THR:HG22	1:C:194:CYS:SG	2.44	0.58
1:C:123[A]:HIS:CE1	1:C:125:GLN:H	2.23	0.56
1:A:210:VAL:HG13	1:A:220:GLY:HA3	1.87	0.55
1:A:228:TYR:O	1:A:229:ARG:NH1	2.39	0.55
1:C:41:ILE:HG21	1:C:192:THR:HG23	1.87	0.55
1:C:13:ASP:OD1	1:C:179:LYS:NZ	2.41	0.53
1:A:14:LEU:HD21	1:A:44:PHE:HB3	1.91	0.51
1:C:123[A]:HIS:CG	1:C:124:TYR:N	2.79	0.51
1:C:115:MET:HE1	1:C:163:VAL:HG23	1.95	0.49
1:A:91:MET:CE	1:A:93:LEU:HD21	2.44	0.48
1:A:200:VAL:HG11	1:A:221:ILE:HD12	1.96	0.48
1:A:148:TYR:C	1:A:148:TYR:CD1	2.87	0.48
1:C:76:LEU:HD23	1:C:143:ILE:HD12	1.95	0.48
1:C:210:VAL:HG11	1:C:240:PRO:HB3	1.95	0.47
1:C:50:LYS:HD3	1:C:97:ARG:NH1	2.29	0.47
1:C:189:LEU:O	1:C:192:THR:HB	2.14	0.47
1:A:97:ARG:HD3	1:C:97:ARG:HD3	1.97	0.47
1:C:123[B]:HIS:NE2	1:C:126:ILE:HD12	2.30	0.47
1:A:45:VAL:HA	1:A:73:PHE:O	2.15	0.46
1:A:59:ARG:HG2	1:A:59:ARG:NH1	2.14	0.46
1:C:116:GLY:O	1:C:122:PHE:HZ	1.99	0.45
1:C:237:ASN:HA	1:C:238:PRO:C	2.36	0.45
1:C:120:GLU:O	1:C:123[B]:HIS:CD2	2.69	0.45
1:C:210:VAL:HG13	1:C:220:GLY:HA3	1.97	0.45
1:A:176:VAL:O	1:A:181:GLU:HG3	2.17	0.44
1:C:123[B]:HIS:CE1	1:C:126:ILE:HD12	2.52	0.44
1:C:140:LEU:HD23	1:C:174:ALA:HB2	1.99	0.44
1:A:167:GLU:O	1:A:171:ASP:N	2.50	0.44
1:A:116:GLY:O	1:A:122:PHE:HZ	2.00	0.43
1:A:21:GLU:C	1:A:23:ALA:H	2.21	0.43
1:A:180:PRO:HB3	1:A:208:ASP:HA	2.01	0.43
1:C:49:THR:HG22	1:C:143:ILE:CG2	2.49	0.42
1:C:49:THR:HG22	1:C:143:ILE:HG21	2.02	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:130:ALA:O	1:A:134:LEU:HG	2.19	0.42
1:C:50:LYS:HE3	1:C:117:LEU:HD23	2.01	0.42
1:A:237:ASN:HA	1:A:238:PRO:C	2.40	0.42
1:C:247:PHE:N	1:C:248:PRO:HD2	2.34	0.42
1:A:80:ARG:NH1	1:A:103:LYS:O	2.41	0.41
1:C:83:LEU:HD13	1:C:90:PRO:HG3	2.03	0.41
1:C:131:PHE:CZ	1:C:135:LEU:HD11	2.56	0.41
1:A:247:PHE:HB3	1:A:248:PRO:HD3	2.03	0.41
1:A:183:THR:O	1:A:187:GLU:HG2	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:C:279:HOH:O	5:C:280:HOH:O[2_555]	1.66	0.54

## 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	257/266 (97%)	248 (96%)	7 (3%)	2 (1%)	19 23
1	C	239/266 (90%)	224 (94%)	14 (6%)	1 (0%)	34 42
All	All	496/532 (93%)	472 (95%)	21 (4%)	3 (1%)	25 31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	14	LEU
1	A	152	LYS
1	A	20	ILE

### 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	194/220 (88%)	189 (97%)	5 (3%)	46	63
1	C	147/220 (67%)	146 (99%)	1 (1%)	84	92
All	All	341/440 (78%)	335 (98%)	6 (2%)	59	75

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	39	SER
1	A	40	VAL
1	A	59	ARG
1	A	233	GLU
1	A	234	GLU
1	C	94	VAL

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

Mol	Chain	Res	Type
1	A	123	HIS

### 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](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	SO4	C	267	-	4,4,4	0.24	0	6,6,6	0.42	0
2	SO4	A	267	-	4,4,4	0.13	0	6,6,6	0.64	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 [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	256/266 (96%)	0.40	15 (5%) 22 28	18, 23, 29, 36	0
1	C	243/266 (91%)	1.11	63 (25%) 0 0	20, 25, 30, 34	0
All	All	499/532 (93%)	0.74	78 (15%) 2 2	18, 24, 30, 36	0

All (78) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	230	ALA	6.8
1	C	228	TYR	6.4
1	C	231	SER	5.2
1	C	106	GLN	4.8
1	C	155	LEU	4.7
1	A	202	ILE	4.5
1	C	26	PRO	4.3
1	C	109	ASP	4.2
1	C	16	GLY	3.8
1	C	123[A]	HIS	3.8
1	C	225	THR	3.7
1	C	226	GLY	3.7
1	C	249	HIS	3.7
1	C	229	ARG	3.6
1	A	12	VAL	3.6
1	C	171	ASP	3.6
1	C	227	LYS	3.5
1	C	234	GLU	3.5
1	A	201	MET	3.4
1	A	11	LEU	3.4
1	A	200	VAL	3.3
1	C	58	GLU	3.3
1	A	219	LEU	3.3
1	C	137	GLY	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	107	THR	3.3
1	C	115	MET	3.3
1	C	248	PRO	3.2
1	C	105	ILE	3.2
1	C	60	LEU	3.2
1	C	62	LYS	3.1
1	A	234	GLU	3.1
1	C	194	CYS	3.1
1	C	219	LEU	3.0
1	C	257	HIS	3.0
1	C	258	LEU	2.9
1	C	259	LEU	2.9
1	C	104	GLY	2.9
1	A	231	SER	2.8
1	C	232	ASP	2.8
1	C	103	LYS	2.8
1	A	232	ASP	2.7
1	C	7	LEU	2.7
1	A	203	GLY	2.7
1	C	206	CYS	2.6
1	C	108	SER	2.6
1	C	200	VAL	2.5
1	C	193	GLY	2.5
1	C	242	LEU	2.5
1	C	75	SER	2.5
1	C	235	LYS	2.4
1	C	59	ARG	2.4
1	C	65	PHE	2.4
1	C	61	ARG	2.3
1	C	220	GLY	2.3
1	C	241	TYR	2.3
1	A	45	VAL	2.3
1	C	56	LEU	2.3
1	A	153	ASP	2.2
1	A	10	VAL	2.2
1	C	73	PHE	2.2
1	C	195	GLU	2.2
1	C	31	ALA	2.2
1	C	57	LEU	2.2
1	C	201	MET	2.2
1	C	34	ARG	2.2
1	C	55	ASP	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	236	ILE	2.2
1	C	88	VAL	2.1
1	C	204	ASP	2.1
1	C	114	VAL	2.1
1	C	143	ILE	2.1
1	A	13	ASP	2.1
1	C	245	GLU	2.0
1	C	100	PRO	2.0
1	A	44	PHE	2.0
1	C	247	PHE	2.0
1	C	224	LYS	2.0
1	C	172	THR	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 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, 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
4	NA	C	269	1/1	0.85	0.11	59,59,59,59	0
2	SO4	C	267	5/5	0.86	0.27	43,44,45,48	5
4	NA	A	269	1/1	0.93	0.10	43,43,43,43	0
3	CL	A	268	1/1	0.94	0.07	63,63,63,63	0
3	CL	C	268	1/1	0.96	0.21	46,46,46,46	1
2	SO4	A	267	5/5	0.97	0.39	24,26,27,31	5

## 6.5 Other polymers [i](#)

There are no such residues in this entry.