



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 01:17 AM GMT

PDB ID : 2CKO  
Title : CRYSTAL STRUCTURE OF HUMAN CHOLINE KINASE ALPHA 2  
Authors : Malito, E.; Lavie, A.  
Deposited on : 2006-04-20  
Resolution : 2.15 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 (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

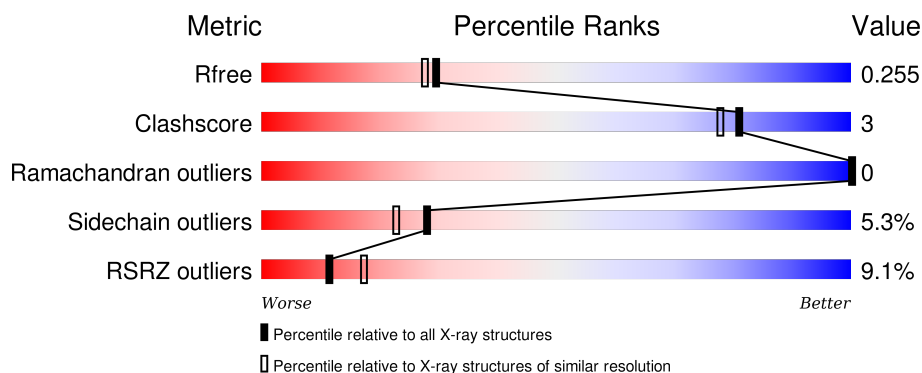
# 1 Overall quality at a glance

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	390	<div> <div>9%</div> <div>79%</div> <div>8%</div> <div>12%</div> </div>
1	B	390	<div> <div>7%</div> <div>79%</div> <div>9%</div> <div>11%</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	DTT	A	1458	-	-	-	X

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6119 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.

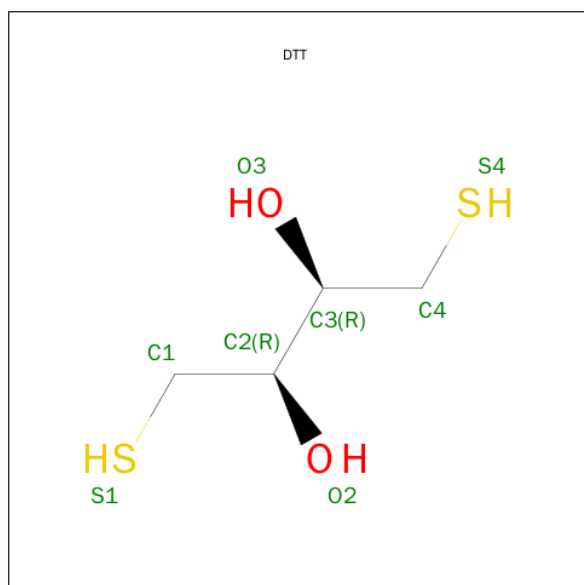
- Molecule 1 is a protein called CHOLINE KINASE ALPHA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	344	Total	C	N	O	S	0	0	0
			2830	1836	476	502	16			
1	B	346	Total	C	N	O	S	0	0	0
			2856	1852	480	508	16			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	220	SER	GLY	CONFLICT	UNP P35790
A	422	LEU	GLN	CONFLICT	UNP P35790
B	220	SER	GLY	CONFLICT	UNP P35790
B	422	LEU	GLN	CONFLICT	UNP P35790

- Molecule 2 is 2,3-DIHYDROXY-1,4-DITHIOBUTANE (three-letter code: DTT) (formula:  $C_4H_{10}O_2S_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 8	C 4	O 2	S 2	0	0
2	B	1	Total 8	C 4	O 2	S 2	0	0

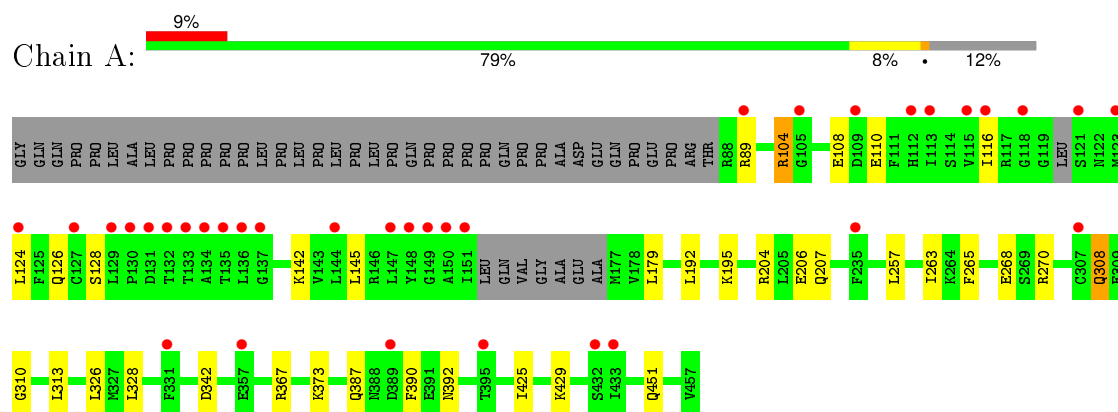
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	214	Total 214	O 214	0	0
3	B	203	Total 203	O 203	0	0

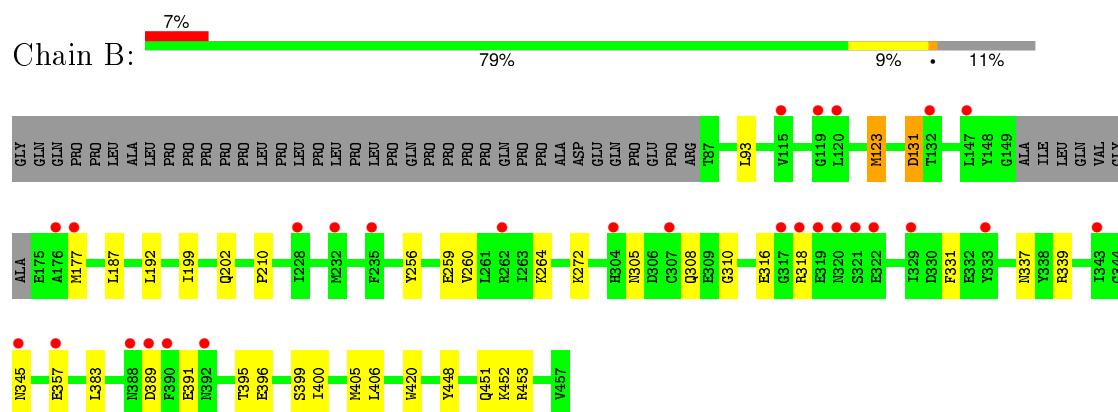
### 3 Residue-property plots [i](#)

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: CHOLINE KINASE ALPHA



#### • Molecule 1: CHOLINE KINASE ALPHA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	56.18Å 122.82Å 132.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.15 29.59 – 2.15	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-2.15) 100.0 (29.59-2.15)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.72 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.199 , 0.256 0.199 , 0.255	Depositor DCC
$R_{free}$ test set	5080 reflections (11.17%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.4	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 50.6	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 50560 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6119	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.37% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: DTT

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	0.68	0/2903	0.67	0/3907
1	B	0.67	0/2930	0.67	0/3945
All	All	0.67	0/5833	0.67	0/7852

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	2830	0	2789	14	0
1	B	2856	0	2820	22	0
2	A	8	0	10	0	0
2	B	8	0	10	1	0
3	A	214	0	0	2	0
3	B	203	0	0	4	0
All	All	6119	0	5629	37	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 3.

All (37) 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:104:ARG:HG2	1:A:104:ARG:HH11	1.42	0.82
1:B:339:ARG:CZ	1:B:405:MET:HE2	2.10	0.82
1:B:123:MET:HE2	1:B:123:MET:H	1.48	0.78
1:B:331:PHE:HE2	3:B:2023:HOH:O	1.70	0.74
1:A:392:ASN:OD1	3:A:2183:HOH:O	2.12	0.68
1:B:305:ASN:HD21	1:B:337:ASN:HD22	1.47	0.62
1:B:331:PHE:CE2	3:B:2023:HOH:O	2.47	0.61
1:A:195:LYS:H	1:A:207:GLN:HE21	1.49	0.60
1:B:389:ASP:OD1	3:B:2169:HOH:O	2.17	0.59
1:B:308:GLN:HE21	1:B:310:GLY:H	1.50	0.57
1:B:305:ASN:ND2	1:B:337:ASN:HD22	2.05	0.54
1:A:308:GLN:HE21	1:A:310:GLY:H	1.56	0.54
1:A:265:PHE:CD1	1:A:270:ARG:HD3	2.43	0.54
1:B:210:PRO:HB2	1:B:316:GLU:HG2	1.91	0.53
1:B:453:ARG:NH1	3:B:2200:HOH:O	2.28	0.52
1:B:383:LEU:HD21	1:B:405:MET:CE	2.39	0.51
1:B:339:ARG:NH2	1:B:405:MET:HE2	2.26	0.50
1:A:387:GLN:NE2	3:A:2179:HOH:O	2.32	0.50
1:A:116:ILE:HD11	1:A:126:GLN:HB2	1.93	0.50
1:B:396:GLU:O	1:B:400:ILE:HG12	2.13	0.49
1:A:126:GLN:HE21	1:A:142:LYS:NZ	2.11	0.49
1:B:131:ASP:N	1:B:131:ASP:OD1	2.44	0.48
1:A:104:ARG:HG2	1:A:104:ARG:NH1	2.18	0.48
1:A:204:ARG:HD2	1:A:206:GLU:OE1	2.14	0.48
1:A:313:LEU:O	1:A:326:LEU:HA	2.13	0.47
1:B:187:LEU:HD13	1:B:331:PHE:CD2	2.50	0.47
1:A:263:ILE:HD11	1:A:425:ILE:HG22	1.97	0.46
1:B:345:ASN:ND2	1:B:420:TRP:HE1	2.13	0.46
1:B:187:LEU:HD13	1:B:331:PHE:CE2	2.52	0.45
1:B:357:GLU:H	1:B:357:GLU:CD	2.19	0.45
2:B:1458:DTT:H12	2:B:1458:DTT:H41	1.57	0.44
1:A:89:ARG:HG3	1:A:108:GLU:OE2	2.20	0.42
1:B:448:TYR:CE2	1:B:452:LYS:HD2	2.55	0.41
1:A:387:GLN:HG3	1:A:390:PHE:HB2	2.01	0.41
1:B:383:LEU:HD21	1:B:405:MET:HE1	2.01	0.41
1:B:256:TYR:O	1:B:260:VAL:HG23	2.20	0.41
1:B:123:MET:HE2	1:B:123:MET:N	2.27	0.40

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	338/390 (87%)	323 (96%)	15 (4%)	0	100	100
1	B	342/390 (88%)	329 (96%)	13 (4%)	0	100	100
All	All	680/780 (87%)	652 (96%)	28 (4%)	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	301/349 (86%)	285 (95%)	16 (5%)	28	23
1	B	305/349 (87%)	289 (95%)	16 (5%)	29	23
All	All	606/698 (87%)	574 (95%)	32 (5%)	28	23

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	104	ARG
1	A	110	GLU
1	A	124	LEU
1	A	128	SER
1	A	145	LEU
1	A	179	LEU
1	A	192	LEU
1	A	257	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	268	GLU
1	A	308	GLN
1	A	328	LEU
1	A	342	ASP
1	A	367	ARG
1	A	373	LYS
1	A	429	LYS
1	A	451	GLN
1	B	93	LEU
1	B	123	MET
1	B	131	ASP
1	B	177	MET
1	B	192	LEU
1	B	199	ILE
1	B	202	GLN
1	B	259	GLU
1	B	264	LYS
1	B	272	LYS
1	B	318	ARG
1	B	391	GLU
1	B	395	THR
1	B	399	SER
1	B	406	LEU
1	B	451	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	112	HIS
1	A	126	GLN
1	A	207	GLN
1	A	308	GLN
1	A	377	HIS
1	A	392	ASN
1	A	451	GLN
1	B	305	ASN
1	B	308	GLN
1	B	345	ASN
1	B	392	ASN
1	B	410	ASN
1	B	451	GLN

### 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 ⓘ

2 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	DTT	A	1458	-	7,7,7	0.61	0	4,8,8	3.02	1 (25%)
2	DTT	B	1458	-	7,7,7	0.57	0	4,8,8	0.50	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	DTT	A	1458	-	-	0/8/8/8	0/0/0/0
2	DTT	B	1458	-	-	0/8/8/8	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( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	A	1458	DTT	C3-C4-S4	-5.62	104.59	113.91

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	B	1458	DTT	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, 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	344/390 (88%)	0.37	35 (10%) 9 15	17, 31, 68, 75	0
1	B	346/390 (88%)	0.28	28 (8%) 15 21	19, 33, 56, 65	0
All	All	690/780 (88%)	0.32	63 (9%) 11 17	17, 32, 63, 75	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	132	THR	6.2
1	B	320	ASN	5.7
1	A	150	ALA	5.3
1	A	149	GLY	5.2
1	A	148	TYR	5.0
1	B	120	LEU	5.0
1	A	151	ILE	4.5
1	B	388	ASN	4.3
1	B	317	GLY	4.2
1	B	119	GLY	4.2
1	A	118	GLY	4.1
1	A	121	SER	4.1
1	B	321	SER	4.0
1	A	115	VAL	3.9
1	A	109	ASP	3.9
1	A	131	ASP	3.7
1	A	116	ILE	3.4
1	B	333	TYR	3.3
1	B	115	VAL	3.3
1	A	433	ILE	3.1
1	B	389	ASP	3.0
1	A	136	LEU	3.0
1	B	176	ALA	2.9
1	A	135	THR	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	112	HIS	2.9
1	B	228	ILE	2.9
1	A	432	SER	2.8
1	B	322	GLU	2.8
1	A	147	LEU	2.8
1	A	235	PHE	2.7
1	B	319	GLU	2.7
1	B	392	ASN	2.7
1	A	123	MET	2.7
1	A	130	PRO	2.7
1	A	129	LEU	2.6
1	A	113	ILE	2.6
1	B	329	ILE	2.6
1	B	307	CYS	2.5
1	A	137	GLY	2.5
1	B	343	ILE	2.5
1	A	357	GLU	2.5
1	A	89	ARG	2.5
1	B	318	ARG	2.4
1	B	235	PHE	2.4
1	A	127	CYS	2.4
1	A	395	THR	2.4
1	A	105	GLY	2.4
1	A	134	ALA	2.3
1	A	331	PHE	2.3
1	B	357	GLU	2.3
1	A	144	LEU	2.3
1	A	389	ASP	2.2
1	A	124	LEU	2.2
1	B	177	MET	2.2
1	B	262	ARG	2.1
1	A	133	THR	2.1
1	A	307	CYS	2.1
1	B	232	MET	2.1
1	B	390	PHE	2.1
1	B	132	THR	2.0
1	B	345	ASN	2.0
1	B	147	LEU	2.0
1	B	304	HIS	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( $\text{\AA}^2$ )	Q<0.9
2	DTT	A	1458	8/8	0.85	0.31	7.07	68,69,70,73	0
2	DTT	B	1458	8/8	0.87	0.14	0.44	66,66,67,68	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.