



# Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 04:07 AM GMT

PDB ID : 4I34  
Title : Crystal Structure of W-W-W ClpX Hexamer  
Authors : Glynn, S.E.; Nager, A.R.; Stinson, B.S.; Schmitz, K.R.; Baker, T.A.; Sauer, R.T.  
Deposited on : 2012-11-23  
Resolution : 4.12 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

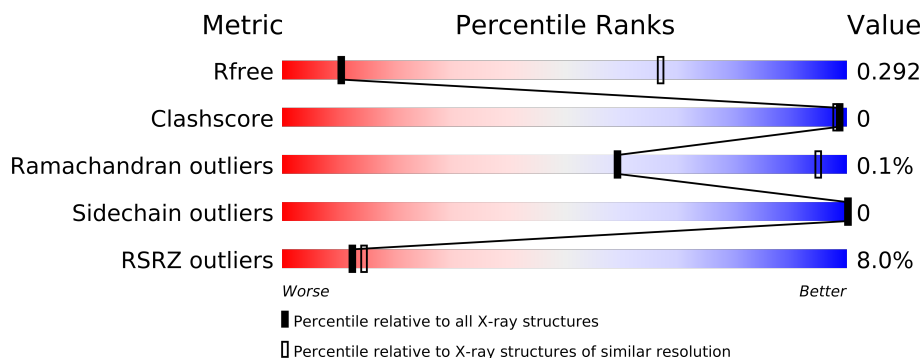
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 4.12 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))
$R_{free}$	66092	1003 (4.76-3.48)
Clashscore	79885	1248 (4.70-3.50)
Ramachandran outliers	78287	1183 (4.70-3.50)
Sidechain outliers	78261	1168 (4.70-3.50)
RSRZ outliers	66119	1003 (4.76-3.48)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	363	
1	B	363	
1	C	363	
1	D	363	
1	E	363	
1	F	363	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	D	501	-	X

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13815 atoms, of which 0 are hydrogen and 0 are deuterium.

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 ATP-dependent Clp protease ATP-binding subunit ClpX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total	C	N	O	S	0	0	0
			2371	1508	392	465	6			
1	B	312	Total	C	N	O	S	0	0	0
			2338	1487	385	460	6			
1	C	303	Total	C	N	O	S	0	0	0
			2275	1450	376	443	6			
1	D	295	Total	C	N	O	S	0	0	0
			2229	1420	367	436	6			
1	E	301	Total	C	N	O	S	0	0	0
			2271	1446	373	446	6			
1	F	306	Total	C	N	O	S	0	0	0
			2301	1463	381	451	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	408	GLU	LYS	CONFLICT	UNP P0A6H1
B	408	GLU	LYS	CONFLICT	UNP P0A6H1
C	408	GLU	LYS	CONFLICT	UNP P0A6H1
D	408	GLU	LYS	CONFLICT	UNP P0A6H1
E	408	GLU	LYS	CONFLICT	UNP P0A6H1
F	408	GLU	LYS	CONFLICT	UNP P0A6H1

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



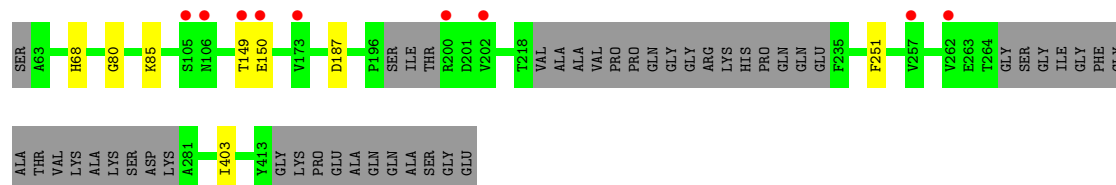
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		

### 3 Residue-property plots

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.

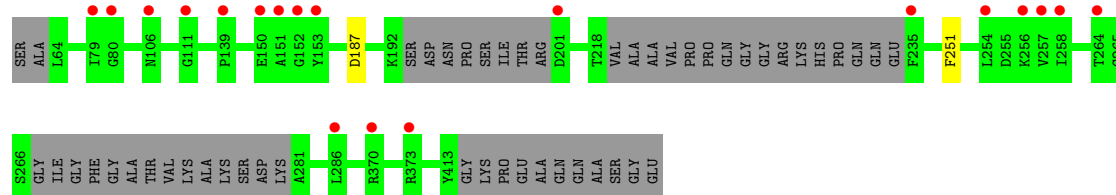
- Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

Chain A: 



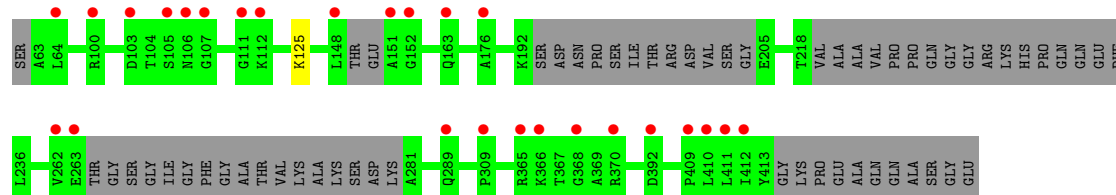
- Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

Chain B: 



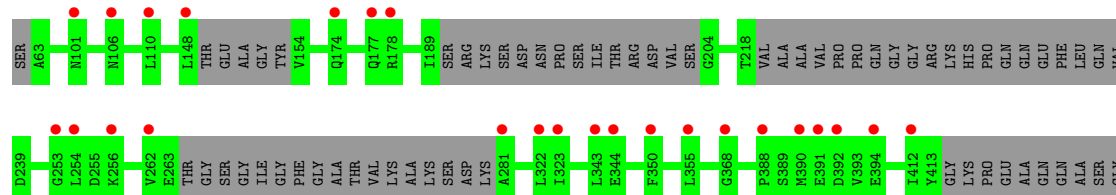
- Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

Chain C: 



- Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

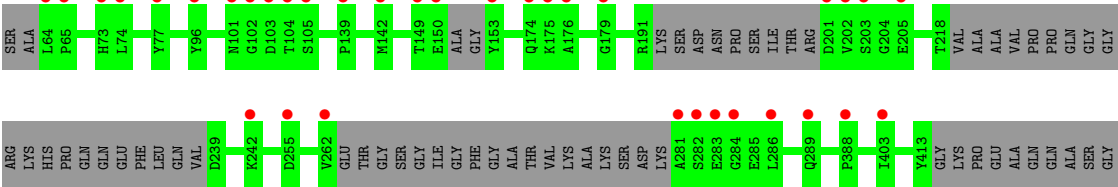
Chain D: 



GLU

• Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

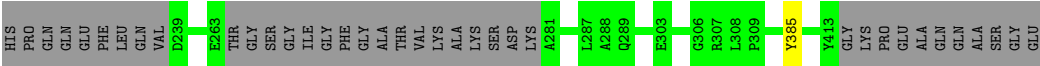
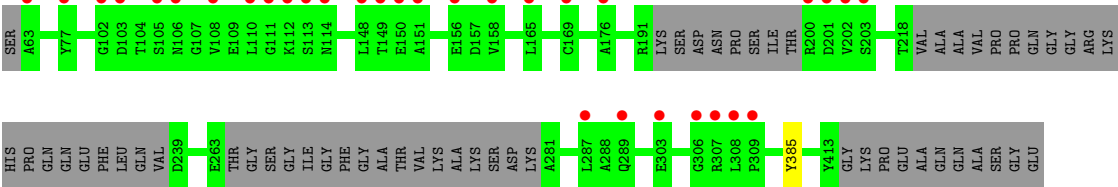
Chain E: 



GLU

• Molecule 1: ATP-dependent Clp protease ATP-binding subunit ClpX

Chain F: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.34Å 199.61Å 203.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.34 – 4.12 38.34 – 4.12	Depositor EDS
% Data completeness (in resolution range)	96.5 (38.34-4.12) 96.6 (38.34-4.12)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.96 (at 4.13Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, $R_{free}$	0.274 , 0.308 0.267 , 0.292	Depositor DCC
$R_{free}$ test set	945 reflections (5.12%)	DCC
Wilson B-factor (Å <sup>2</sup> )	180.6	Xtriage
Anisotropy	0.420	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 172.2	EDS
Estimated twinning fraction	0.006 for -h,l,k	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	2 of 18487 reflections (0.011%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	13815	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	241.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

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	0.23	0/2398	0.40	0/3251
1	B	0.22	0/2364	0.38	0/3203
1	C	0.21	0/2299	0.37	0/3114
1	D	0.21	0/2252	0.37	0/3048
1	E	0.22	0/2295	0.37	0/3108
1	F	0.22	0/2326	0.39	0/3151
All	All	0.22	0/13934	0.38	0/18875

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2371	0	0	4	0
1	B	2338	0	0	1	0
1	C	2275	0	0	1	0
1	D	2229	0	0	0	0
1	E	2271	0	0	0	0
1	F	2301	0	0	1	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	1	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
All	All	13815	0	0	6	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 0.

All (6) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:187:ASP:OD2	1:A:251:PHE:N	2.48	0.46
1:C:125:LYS:NZ	2:C:501:SO4:O4	2.50	0.45
1:B:187:ASP:OD2	1:B:251:PHE:N	2.49	0.45
1:A:68:HIS:CE1	1:F:385:TYR:CE2	3.08	0.42
1:A:149:THR:O	1:A:150:GLU:CB	2.70	0.40
1:A:80:GLY:O	1:A:85:LYS:NZ	2.54	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	308/363 (85%)	288 (94%)	19 (6%)	1 (0%)	50	91
1	B	304/363 (84%)	287 (94%)	17 (6%)	0	100	100
1	C	293/363 (81%)	277 (94%)	16 (6%)	0	100	100
1	D	285/363 (78%)	275 (96%)	10 (4%)	0	100	100
1	E	291/363 (80%)	278 (96%)	13 (4%)	0	100	100
1	F	298/363 (82%)	283 (95%)	15 (5%)	0	100	100
All	All	1779/2178 (82%)	1688 (95%)	90 (5%)	1 (0%)	59	95

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	403	ILE

### 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	247/299 (83%)	247 (100%)	0	100	100
1	B	244/299 (82%)	244 (100%)	0	100	100
1	C	237/299 (79%)	237 (100%)	0	100	100
1	D	235/299 (79%)	235 (100%)	0	100	100
1	E	240/299 (80%)	240 (100%)	0	100	100
1	F	241/299 (81%)	241 (100%)	0	100	100
All	All	1444/1794 (80%)	1444 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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 chains 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 ⓘ

6 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	A	501	-	4,4,4	0.22	0	6,6,6	0.10	0
2	SO4	B	501	-	4,4,4	0.24	0	6,6,6	0.08	0
2	SO4	C	501	-	4,4,4	0.31	0	6,6,6	0.07	0
2	SO4	D	501	-	4,4,4	0.16	0	6,6,6	0.09	0
2	SO4	E	501	-	4,4,4	0.21	0	6,6,6	0.06	0
2	SO4	F	501	-	4,4,4	0.22	0	6,6,6	0.09	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	SO4	A	501	-	-	0/0/0/0	0/0/0/0
2	SO4	B	501	-	-	0/0/0/0	0/0/0/0
2	SO4	C	501	-	-	0/0/0/0	0/0/0/0
2	SO4	D	501	-	-	0/0/0/0	0/0/0/0
2	SO4	E	501	-	-	0/0/0/0	0/0/0/0
2	SO4	F	501	-	-	0/0/0/0	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.

## 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	316/363 (87%)	0.42	9 (2%)	50	40	146, 193, 288, 354	0
1	B	312/363 (85%)	0.46	19 (6%)	21	20	151, 212, 264, 355	0
1	C	303/363 (83%)	0.59	26 (8%)	11	13	169, 244, 303, 355	0
1	D	295/363 (81%)	0.68	25 (8%)	11	13	213, 264, 317, 361	0
1	E	301/363 (82%)	0.85	35 (11%)	5	8	216, 283, 352, 382	0
1	F	306/363 (84%)	0.65	32 (10%)	7	10	147, 235, 316, 361	0
All	All	1833/2178 (84%)	0.61	146 (7%)	12	14	146, 243, 319, 382	0

All (146) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	200	ARG	6.3
1	F	112	LYS	6.2
1	E	281	ALA	5.7
1	E	77	TYR	5.3
1	C	151	ALA	5.1
1	F	201	ASP	5.1
1	E	150	GLU	4.9
1	A	149	THR	4.7
1	D	178	ARG	4.6
1	E	176	ALA	4.6
1	F	103	ASP	4.3
1	E	149	THR	4.2
1	D	392	ASP	4.1
1	E	283	GLU	4.1
1	E	262	VAL	4.1
1	E	139	PRO	4.1
1	D	390	MET	4.0
1	E	96	TYR	4.0
1	F	309	PRO	3.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	111	GLY	3.9
1	D	391	GLU	3.9
1	C	107	GLY	3.8
1	E	202	VAL	3.7
1	F	307	ARG	3.5
1	D	350	PHE	3.5
1	C	410	LEU	3.5
1	E	205	GLU	3.4
1	C	152	GLY	3.4
1	E	201	ASP	3.3
1	F	114	ASN	3.2
1	C	411	LEU	3.2
1	F	202	VAL	3.2
1	E	64	LEU	3.2
1	F	287	LEU	3.2
1	B	258	ILE	3.2
1	D	177	GLN	3.1
1	C	111	GLY	3.1
1	E	103	ASP	3.1
1	F	303	GLU	3.1
1	F	156	GLU	3.1
1	C	289	GLN	3.1
1	E	403	ILE	3.1
1	D	106	ASN	3.0
1	F	113	SER	3.0
1	B	111	GLY	3.0
1	A	150	GLU	3.0
1	E	282	SER	3.0
1	B	150	GLU	2.9
1	B	235	PHE	2.9
1	C	100	ARG	2.9
1	C	176	ALA	2.9
1	E	102	GLY	2.9
1	A	106	ASN	2.9
1	E	73	HIS	2.9
1	C	262	VAL	2.9
1	C	148	LEU	2.8
1	F	200	ARG	2.8
1	B	79	ILE	2.8
1	B	254	LEU	2.8
1	F	176	ALA	2.8
1	A	173	VAL	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	108	VAL	2.7
1	B	151	ALA	2.7
1	E	105	SER	2.7
1	C	412	ILE	2.7
1	C	106	ASN	2.6
1	D	343	LEU	2.6
1	F	306	GLY	2.6
1	A	105	SER	2.6
1	F	289	GLN	2.6
1	B	257	VAL	2.6
1	E	175	LYS	2.6
1	F	102	GLY	2.6
1	E	174	GLN	2.6
1	F	110	LEU	2.6
1	E	142	MET	2.6
1	D	368	GLY	2.6
1	C	365	ARG	2.6
1	D	412	ILE	2.5
1	F	105	SER	2.5
1	C	103	ASP	2.5
1	F	148	LEU	2.5
1	D	256	LYS	2.5
1	F	308	LEU	2.5
1	D	322	LEU	2.5
1	B	153	TYR	2.5
1	D	148	LEU	2.5
1	F	150	GLU	2.5
1	B	373	ARG	2.5
1	B	264	THR	2.4
1	F	63	ALA	2.4
1	A	262	VAL	2.4
1	C	370	ARG	2.4
1	D	388	PRO	2.4
1	F	151	ALA	2.4
1	F	158	VAL	2.4
1	A	257	VAL	2.4
1	B	370	ARG	2.4
1	F	203	SER	2.3
1	C	392	ASP	2.3
1	F	106	ASN	2.3
1	E	284	GLY	2.3
1	D	262	VAL	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	242	LYS	2.3
1	F	165	LEU	2.3
1	E	255	ASP	2.3
1	B	152	GLY	2.3
1	D	101	ASN	2.3
1	D	253	GLY	2.3
1	C	105	SER	2.2
1	D	323	ILE	2.2
1	F	77	TYR	2.2
1	E	74	LEU	2.2
1	F	149	THR	2.2
1	D	110	LEU	2.2
1	E	179	GLY	2.2
1	F	169	CYS	2.2
1	D	344	GLU	2.2
1	B	201	ASP	2.2
1	B	106	ASN	2.2
1	C	263	GLU	2.2
1	E	388	PRO	2.2
1	C	366	LYS	2.2
1	E	104	THR	2.1
1	C	309	PRO	2.1
1	E	203	SER	2.1
1	E	286	LEU	2.1
1	E	289	GLN	2.1
1	D	254	LEU	2.1
1	E	101	ASN	2.1
1	B	139	PRO	2.1
1	C	163	GLN	2.1
1	D	281	ALA	2.1
1	D	394	GLU	2.1
1	D	355	LEU	2.1
1	E	65	PRO	2.1
1	B	286	LEU	2.1
1	C	112	LYS	2.1
1	C	409	PRO	2.1
1	A	202	VAL	2.0
1	B	80	GLY	2.0
1	D	174	GLN	2.0
1	E	153	TYR	2.0
1	B	256	LYS	2.0
1	C	368	GLY	2.0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	64	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

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	D	501	5/5	0.59	2.52	253,253,256,264	0
2	SO4	B	501	5/5	0.31	1.74	176,177,180,189	0
2	SO4	E	501	5/5	0.48	1.67	237,240,241,243	0
2	SO4	A	501	5/5	0.34	0.66	173,176,180,187	0
2	SO4	C	501	5/5	0.22	-0.27	177,180,182,185	0
2	SO4	F	501	5/5	0.14	-1.74	187,190,194,195	0

## 6.5 Other polymers

There are no such residues in this entry.