



Full wwPDB X-ray Structure Validation Report

Apr 25, 2014 – 01:18 AM EDT

PDB ID : 4C0B
Title : Structure of wild-type Clp1p-Pcf11p (454 -563) complex
Authors : Fribourg, S.; Dupin, A.F.
Deposited on : 2013-08-01
Resolution : 2.77 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

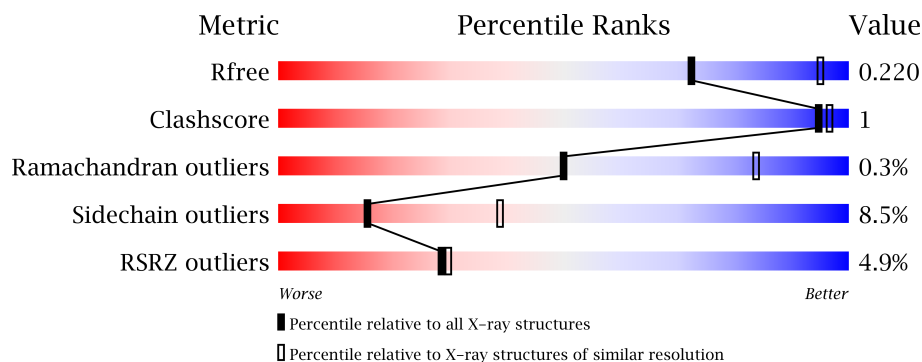
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.16 November 2013
Xtriage (Phenix) : dev-1439
EDS : stable22978
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) : stable22978

1 Overall quality at a glance

The reported resolution of this entry is 2.77 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	2193 (2.80-2.76)
Clashscore	79885	2751 (2.80-2.76)
Ramachandran outliers	78287	2699 (2.80-2.76)
Sidechain outliers	78261	2701 (2.80-2.76)
RSRZ outliers	66119	2196 (2.80-2.76)

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 ≥ 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	446	
1	B	446	
2	C	110	
2	D	110	

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

Mol	Type	Chain	Res	Geometry	Electron density
4	MG	A	1447	-	X
4	MG	B	1448	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7464 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 MRNA CLEAVAGE AND POLYADENYLATION FACTOR CLP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	428	Total	C	N	O	S	0	0	0
			3410	2191	571	636	12			
1	B	430	Total	C	N	O	S	0	0	0
			3426	2201	574	639	12			

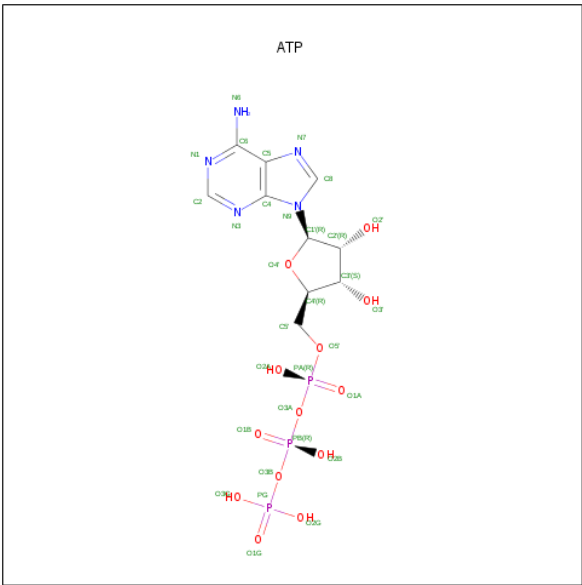
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	446	GLN	-	EXPRESSION TAG	UNP Q08685
B	446	GLN	-	EXPRESSION TAG	UNP Q08685

- Molecule 2 is a protein called PCF11P.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	27	Total	C	N	O	0	0	0
			230	143	41	46			
2	D	36	Total	C	N	O	0	0	0
			290	177	51	62			

- Molecule 3 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: C₁₀H₁₆N₅O₁₃P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		
4	A	1	Total	Mg	0	0
			1	1		

- Molecule 5 is water.

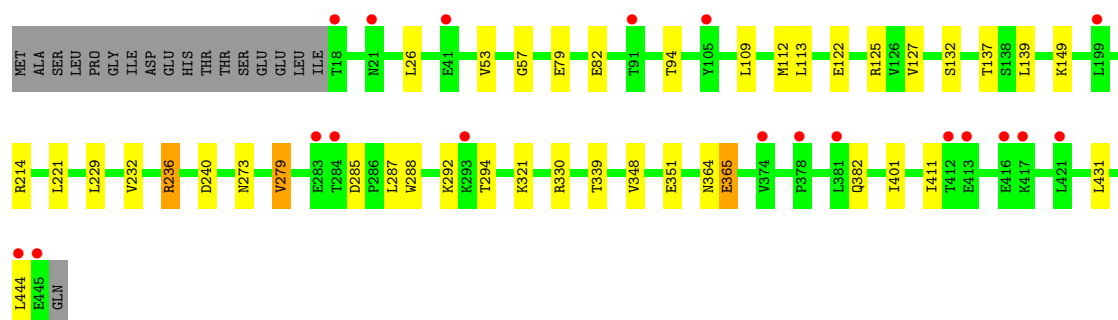
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	19	Total	O	0	0
			19	19		
5	B	18	Total	O	0	0
			18	18		
5	C	1	Total	O	0	0
			1	1		
5	D	6	Total	O	0	0
			6	6		

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.

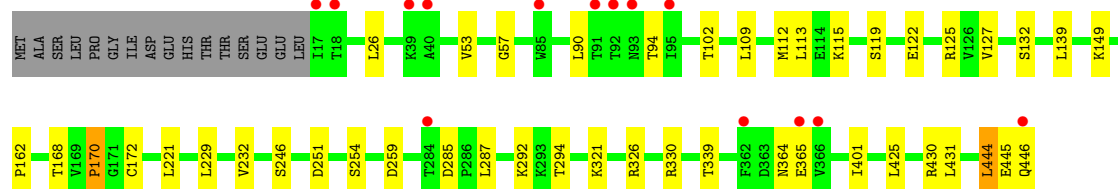
• Molecule 1: MRNA CLEAVAGE AND POLYADENYLATION FACTOR CLP1

Chain A: 



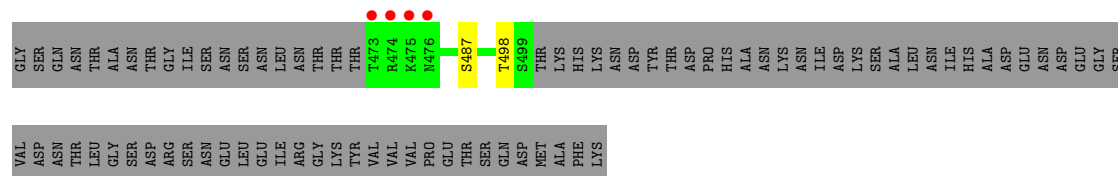
• Molecule 1: MRNA CLEAVAGE AND POLYADENYLATION FACTOR CLP1

Chain B: 



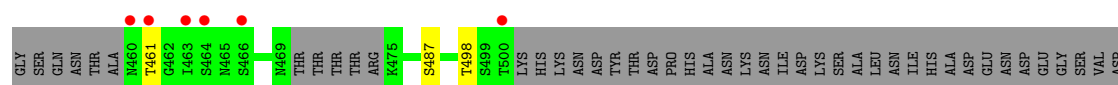
• Molecule 2: PCF11P

Chain C: 



• Molecule 2: PCF11P

Chain D: 



ASN	THR	LEU	GLY	SER	ASP	ARG	SER	ASN	GLU	LEU	GLU	ILE	ARG	GLY	LYS	TYR	VAL	VAL	PRO	GLU	THR	SER	GLN	ASP	MET	ALA	PHE	LYS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	89.35Å 95.41Å 182.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.55 – 2.77 40.46 – 2.77	Depositor EDS
% Data completeness (in resolution range)	99.6 (40.55-2.77) 99.6 (40.46-2.77)	Depositor EDS
R_{merge}	0.01	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.54 (at 2.77Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.179 , 0.214 0.191 , 0.220	Depositor DCC
R_{free} test set	2018 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	58.4	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 40.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 40190 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7464	wwPDB-VP
Average B, all atoms (Å ²)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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: MG, ATP

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	0.48	0/3487	0.78	0/4744
1	B	0.51	0/3503	0.81	3/4766 (0.1%)
2	C	0.53	0/235	0.69	0/317
2	D	0.71	0/294	0.90	0/397
All	All	0.51	0/7519	0.79	3/10224 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	246	SER	C-N-CA	-6.03	109.64	122.30
1	B	172	CYS	N-CA-C	5.36	125.48	111.00
1	B	170	PRO	N-CA-C	5.27	125.81	112.10

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	3410	0	0	6	0
1	B	3426	0	0	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	230	0	0	0	0
2	D	290	0	0	0	0
3	A	31	0	0	0	0
3	B	31	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	19	0	0	0	0
5	B	18	0	0	0	0
5	C	1	0	0	0	0
5	D	6	0	0	0	0
All	All	7464	0	0	10	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 1.

All (10) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:364:ASN:O	1:A:365:GLU:CB	2.36	0.74
1:A:57:GLY:O	1:A:149:LYS:NZ	2.34	0.61
1:B:57:GLY:O	1:B:149:LYS:NZ	2.34	0.60
1:B:326:ARG:NH2	1:B:444:LEU:O	2.36	0.58
1:B:162:PRO:O	1:B:170:PRO:O	2.25	0.54
1:A:125:ARG:NH2	1:B:122:GLU:OE2	2.42	0.53
1:B:364:ASN:O	1:B:365:GLU:CB	2.57	0.52
1:A:236:ARG:NH1	1:A:240:ASP:OD2	2.45	0.49
1:A:122:GLU:OE2	1:B:125:ARG:NH2	2.46	0.48
1:A:279:VAL:CG1	1:A:288:TRP:CZ3	3.05	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	426/446 (96%)	408 (96%)	18 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	428/446 (96%)	407 (95%)	21 (5%)	0	100	100
2	C	25/110 (23%)	23 (92%)	1 (4%)	1 (4%)	5	13
2	D	32/110 (29%)	26 (81%)	4 (12%)	2 (6%)	2	5
All	All	911/1112 (82%)	864 (95%)	44 (5%)	3 (0%)	50	84

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	498	THR
2	D	498	THR
2	D	461	THR

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	384/401 (96%)	350 (91%)	34 (9%)	14	36
1	B	386/401 (96%)	352 (91%)	34 (9%)	14	37
2	C	25/98 (26%)	24 (96%)	1 (4%)	42	78
2	D	33/98 (34%)	32 (97%)	1 (3%)	53	86
All	All	828/998 (83%)	758 (92%)	70 (8%)	15	39

All (70) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	26	LEU
1	A	53	VAL
1	A	79	GLU
1	A	82	GLU
1	A	94	THR
1	A	109	LEU
1	A	112	MET
1	A	113	LEU
1	A	127	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	132	SER
1	A	137	THR
1	A	139	LEU
1	A	214	ARG
1	A	221	LEU
1	A	229	LEU
1	A	232	VAL
1	A	236	ARG
1	A	273	ASN
1	A	279	VAL
1	A	285	ASP
1	A	287	LEU
1	A	292	LYS
1	A	294	THR
1	A	321	LYS
1	A	330	ARG
1	A	339	THR
1	A	348	VAL
1	A	351	GLU
1	A	365	GLU
1	A	382	GLN
1	A	401	ILE
1	A	411	ILE
1	A	431	LEU
1	A	444	LEU
1	B	26	LEU
1	B	53	VAL
1	B	90	LEU
1	B	94	THR
1	B	102	THR
1	B	109	LEU
1	B	112	MET
1	B	113	LEU
1	B	115	LYS
1	B	119	SER
1	B	127	VAL
1	B	132	SER
1	B	139	LEU
1	B	168	THR
1	B	221	LEU
1	B	229	LEU
1	B	232	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	251	ASP
1	B	254	SER
1	B	259	ASP
1	B	285	ASP
1	B	287	LEU
1	B	292	LYS
1	B	294	THR
1	B	321	LYS
1	B	330	ARG
1	B	339	THR
1	B	401	ILE
1	B	425	LEU
1	B	430	ARG
1	B	431	LEU
1	B	444	LEU
1	B	445	GLU
1	B	446	GLN
2	C	487	SER
2	D	487	SER

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 ⓘ

Of 4 ligands modelled in this entry, 2 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
3	ATP	A	1446	4	33,33,33	1.46	2 (6%)	52,52,52	0.93	2 (3%)
3	ATP	B	1447	4	33,33,33	1.46	2 (6%)	52,52,52	0.95	2 (3%)

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
3	ATP	A	1446	4	-	0/22/38/38	0/3/3/3
3	ATP	B	1447	4	-	0/22/38/38	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1446	ATP	PA-O3A	-5.38	1.50	1.59
3	B	1447	ATP	PB-O3B	-5.38	1.50	1.59
3	A	1446	ATP	PB-O3B	-5.31	1.50	1.59
3	B	1447	ATP	PA-O3A	-5.27	1.50	1.59

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1447	ATP	O3A-PB-O3B	3.94	109.67	101.66
3	A	1446	ATP	O3A-PB-O3B	3.75	109.30	101.66
3	B	1447	ATP	O3A-PA-O5'	2.56	109.69	102.91
3	A	1446	ATP	O3A-PA-O5'	2.43	109.35	102.91

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, 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	428/446 (95%)	0.14	19 (4%) 33 35	35, 70, 117, 144	0
1	B	430/446 (96%)	0.07	14 (3%) 44 47	33, 57, 93, 125	0
2	C	27/110 (24%)	0.52	4 (14%) 3 3	37, 54, 119, 128	0
2	D	36/110 (32%)	0.81	6 (16%) 2 2	35, 60, 122, 129	0
All	All	921/1112 (82%)	0.15	43 (4%) 28 32	33, 63, 111, 144	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	460	ASN	6.6
2	D	463	ILE	6.2
1	A	417	LYS	4.5
1	B	366	VAL	4.2
1	B	17	ILE	4.2
1	A	378	PRO	3.9
2	D	464	SER	3.8
2	D	461	THR	3.8
2	C	475	LYS	3.7
2	C	473	THR	3.6
1	A	18	THR	3.6
2	C	474	ARG	3.6
1	A	199	LEU	3.5
1	B	18	THR	3.4
1	B	284	THR	3.3
1	A	283	GLU	3.3
1	B	446	GLN	3.1
1	A	381	LEU	2.9
1	A	374	VAL	2.9
1	B	91	THR	2.9
1	A	91	THR	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	476	ASN	2.8
1	A	413	GLU	2.7
2	D	466	SER	2.6
1	B	365	GLU	2.6
1	B	39	LYS	2.5
1	A	284	THR	2.5
1	A	445	GLU	2.4
2	D	500	THR	2.4
1	A	293	LYS	2.4
1	A	21	ASN	2.4
1	A	444	LEU	2.3
1	B	92	THR	2.3
1	A	421	LEU	2.2
1	B	40	ALA	2.2
1	B	93	ASN	2.2
1	B	95	ILE	2.2
1	B	362	PHE	2.0
1	A	412	THR	2.0
1	A	41	GLU	2.0
1	A	416	GLU	2.0
1	B	85	TRP	2.0
1	A	105	TYR	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, 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	MG	A	1447	1/1	0.42	11.38	67,67,67,67	0
4	MG	B	1448	1/1	0.42	7.62	66,66,66,66	0
3	ATP	B	1447	31/31	0.19	0.51	103,110,117,118	0
3	ATP	A	1446	31/31	0.18	-0.32	120,124,134,136	0

6.5 Other polymers

There are no such residues in this entry.