Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2016 – 04:23 PM EDT

PDB ID	:	5JTJ
Title	:	USP7CD-CTP in complex with Ubiquitin
Authors	:	Murray, J.M.; Rouge, L.
Deposited on	:	2016-05-09
Resolution	:	3.32 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
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:

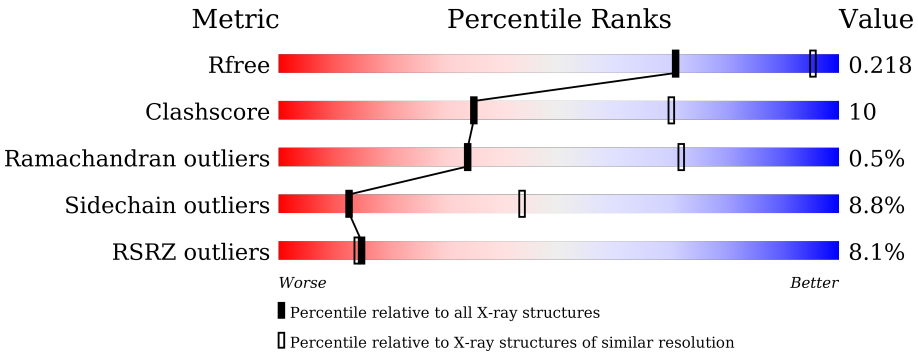
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.32 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	1198 (3.40-3.24)
Clashscore	102246	1280 (3.40-3.24)
Ramachandran outliers	100387	1260 (3.40-3.24)
Sidechain outliers	100360	1259 (3.40-3.24)
RSRZ outliers	91569	1203 (3.40-3.24)

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. 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 <=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	387	8% 63% 26% · 8%
2	B	76	5% 74% 24% ·

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3480 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 Ubiquitin carboxyl-terminal hydrolase 7,Ubiquitin carboxyl-terminal hydrolase 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	356	2883	1825	491	551	16	3	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	192	MET	-	initiating methionine	UNP Q93009
A	193	HIS	-	expression tag	UNP Q93009
A	194	HIS	-	expression tag	UNP Q93009
A	195	HIS	-	expression tag	UNP Q93009
A	196	HIS	-	expression tag	UNP Q93009
A	197	HIS	-	expression tag	UNP Q93009
A	198	HIS	-	expression tag	UNP Q93009
A	199	GLY	-	expression tag	UNP Q93009
A	200	GLU	-	expression tag	UNP Q93009
A	201	ASN	-	expression tag	UNP Q93009
A	202	LEU	-	expression tag	UNP Q93009
A	203	TYR	-	expression tag	UNP Q93009
A	204	PHE	-	expression tag	UNP Q93009
A	205	GLN	-	expression tag	UNP Q93009
A	206	GLY	-	expression tag	UNP Q93009
A	207	SER	-	expression tag	UNP Q93009
A	208	LYS	-	expression tag	UNP Q93009
A	1079	GLY	-	linker	UNP Q93009
A	1080	GLY	-	linker	UNP Q93009
A	1081	SER	-	linker	UNP Q93009
A	1082	GLY	-	linker	UNP Q93009
A	1083	GLY	-	linker	UNP Q93009

- Molecule 2 is a protein called Polyubiquitin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	76	Total	C	N	O	S	0	0	0
			596	375	104	116	1			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

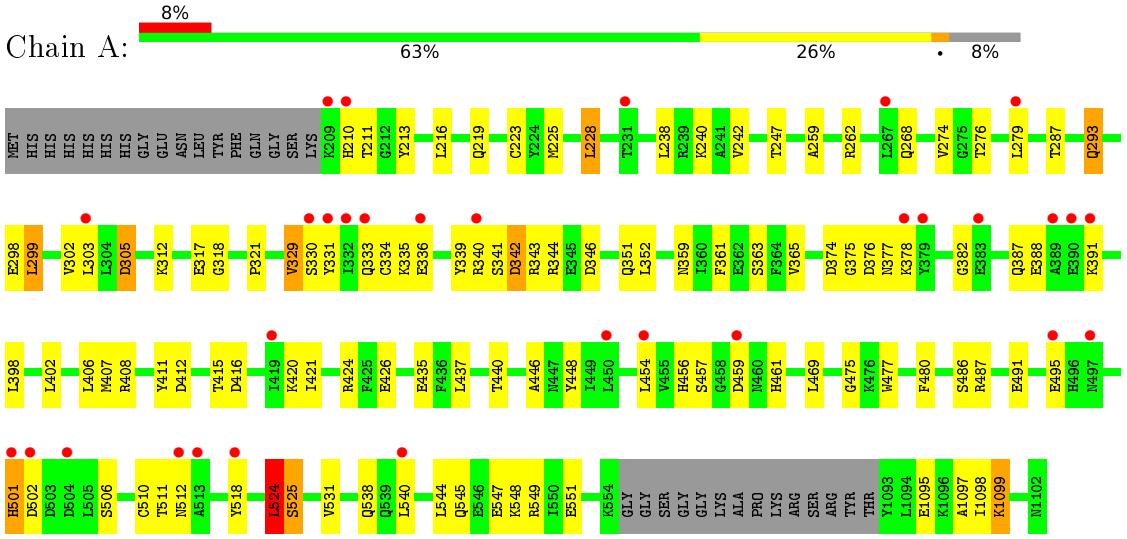
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		

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.

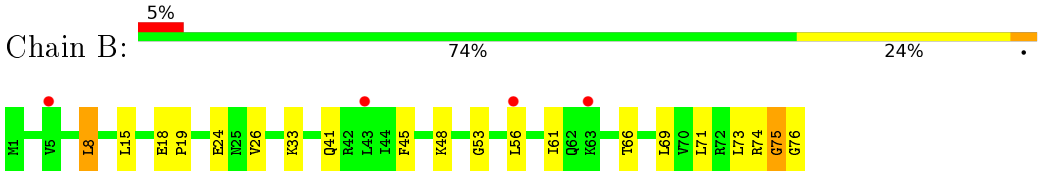
• Molecule 1: Ubiquitin carboxyl-terminal hydrolase 7,Ubiquitin carboxyl-terminal hydrolase 7

Chain A:



• Molecule 2: Polyubiquitin-B

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α, β, γ	148.06Å 148.06Å 179.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	74.03 – 3.32 128.22 – 3.32	Depositor EDS
% Data completeness (in resolution range)	99.9 (74.03-3.32) 100.0 (128.22-3.32)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.18 (at 3.33Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.191 , 0.209 0.204 , 0.218	Depositor DCC
R_{free} test set	939 reflections (5.28%)	DCC
Wilson B-factor (Å^2)	107.0	Xtriage
Anisotropy	0.065	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 99.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	3480	wwPDB-VP
Average B, all atoms (Å^2)	111.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: CA

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	0.44	0/2944	0.60	1/3970 (0.0%)
2	B	0.43	0/602	0.62	1/811 (0.1%)
All	All	0.44	0/3546	0.60	2/4781 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	75	GLY	C-N-CA	5.91	134.70	122.30
1	A	524	LEU	CA-CB-CG	5.16	127.16	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	75	GLY	Peptide

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	2883	0	2810	58	0
2	B	596	0	618	15	0
3	A	1	0	0	0	0
All	All	3480	0	3428	68	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 10.

All (68) 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:1095:GLU:HG2	1:A:1099:LYS:HE2	1.55	0.89
2:B:26:VAL:HG21	2:B:56:LEU:HD21	1.67	0.77
1:A:312:LYS:HA	1:A:317:GLU:HG3	1.67	0.76
1:A:495:GLU:O	1:A:501:HIS:HE1	1.68	0.75
2:B:24:GLU:OE2	2:B:53:GLY:N	2.22	0.72
1:A:424:ARG:NH1	1:A:426:GLU:OE2	2.24	0.70
1:A:547:GLU:O	1:A:551:GLU:HG3	1.93	0.67
1:A:491:GLU:HG3	1:A:495:GLU:HG3	1.77	0.67
1:A:216:LEU:HD23	1:A:274:VAL:HB	1.78	0.65
1:A:305:ASP:OD1	2:B:48:LYS:HE2	2.01	0.60
1:A:1098:ILE:O	1:A:1098:ILE:HG12	2.02	0.58
1:A:545:GLN:HA	1:A:548:LYS:HD2	1.89	0.55
1:A:375:GLY:O	1:A:378:LYS:HG3	2.07	0.55
1:A:359:ASN:HB3	1:A:426:GLU:HB2	1.89	0.54
1:A:545:GLN:O	1:A:549:ARG:HG3	2.08	0.54
1:A:448:TYR:HB3	1:A:518:TYR:HB3	1.89	0.53
1:A:495:GLU:O	1:A:501:HIS:CE1	2.58	0.53
1:A:1095:GLU:HG3	1:A:1097:ALA:H	1.72	0.53
1:A:223:CYS:H	2:B:76:GLY:C	2.12	0.53
1:A:333:GLN:HG2	1:A:340:ARG:HG3	1.91	0.53
1:A:524:LEU:HD13	1:A:525:SER:N	2.24	0.53
1:A:219:GLN:NE2	1:A:276:THR:O	2.43	0.51
2:B:18:GLU:HG3	2:B:19:PRO:HD2	1.91	0.51
1:A:268:GLN:HE22	1:A:531:VAL:HG12	1.76	0.50
1:A:1098:ILE:HG23	1:A:1098:ILE:O	2.12	0.48
1:A:210:HIS:O	1:A:487:ARG:NH1	2.46	0.48
1:A:259:ALA:HA	1:A:262:ARG:NH1	2.28	0.48
1:A:469:LEU:HB2	1:A:480:PHE:HE1	1.78	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:8:LEU:HD11	2:B:71:LEU:HD13	1.95	0.48
1:A:240:LYS:NZ	1:A:531:VAL:HA	2.29	0.48
1:A:412:ASP:OD1	1:A:415:THR:HG23	2.14	0.47
2:B:45:PHE:CE2	2:B:61:ILE:HG12	2.48	0.47
1:A:238:LEU:O	1:A:242:VAL:HG23	2.14	0.47
1:A:334:CYS:HB3	1:A:387:GLN:HG2	1.96	0.47
1:A:293:GLN:O	2:B:74:ARG:HD2	2.15	0.47
1:A:342:ASP:O	1:A:343:ARG:HD2	2.15	0.47
1:A:406:LEU:O	1:A:408:ARG:N	2.48	0.47
1:A:361:PHE:O	1:A:365:VAL:HG23	2.15	0.47
1:A:298:GLU:O	1:A:302:VAL:HG23	2.16	0.46
1:A:411:TYR:OH	1:A:416:ASP:OD1	2.32	0.46
2:B:41:GLN:HB2	2:B:69:LEU:HD11	1.97	0.46
1:A:228:LEU:HD12	1:A:299:LEU:HB3	1.96	0.46
1:A:408:ARG:NH1	1:A:421:ILE:O	2.44	0.46
1:A:475:GLY:O	1:A:477:TRP:HD1	1.98	0.46
2:B:8:LEU:HD22	2:B:8:LEU:H	1.80	0.46
1:A:225:MET:SD	1:A:279:LEU:HD12	2.57	0.45
1:A:293:GLN:HB2	2:B:74:ARG:HB3	1.98	0.45
1:A:259:ALA:HA	1:A:262:ARG:HH11	1.82	0.44
1:A:491:GLU:HA	1:A:495:GLU:HG2	1.99	0.44
1:A:398:LEU:HD12	1:A:446:ALA:HB1	1.98	0.44
1:A:352:LEU:HB3	1:A:363:SER:OG	2.18	0.43
2:B:45:PHE:CD2	2:B:61:ILE:HG12	2.54	0.43
1:A:211:THR:C	1:A:213:TYR:H	2.23	0.42
1:A:544:LEU:O	1:A:548:LYS:HG3	2.19	0.42
1:A:330:SER:OG	1:A:391:LYS:HE2	2.19	0.42
1:A:318:GLY:C	1:A:321:PRO:HD2	2.41	0.41
1:A:240:LYS:HZ1	1:A:531:VAL:HA	1.85	0.41
1:A:377:ASN:OD1	2:B:33:LYS:NZ	2.50	0.41
2:B:15:LEU:HA	2:B:15:LEU:HD23	1.89	0.41
1:A:335:LYS:HD3	1:A:388:GLU:HB3	2.02	0.41
1:A:331:TYR:O	1:A:391:LYS:HA	2.21	0.41
1:A:437:LEU:HD23	1:A:437:LEU:HA	1.92	0.41
1:A:262:ARG:NH2	1:A:547:GLU:OE2	2.54	0.41
1:A:457:SER:HB2	1:A:511:THR:OG1	2.21	0.41
1:A:329:VAL:HG13	1:A:344:ARG:HG2	2.03	0.40
1:A:420:LYS:HD3	1:A:459:ASP:HA	2.03	0.40
2:B:73:LEU:HD23	2:B:73:LEU:HA	1.64	0.40
1:A:420:LYS:HE3	1:A:456:HIS:ND1	2.36	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	352/387 (91%)	327 (93%)	23 (6%)	2 (1%)	30 69
2	B	74/76 (97%)	72 (97%)	2 (3%)	0	100 100
All	All	426/463 (92%)	399 (94%)	25 (6%)	2 (0%)	34 72

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	382	GLY
1	A	407	MET

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	318/346 (92%)	286 (90%)	32 (10%)	9 35
2	B	67/68 (98%)	65 (97%)	2 (3%)	48 79
All	All	385/414 (93%)	351 (91%)	34 (9%)	12 43

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	228	LEU
1	A	247	THR
1	A	287	THR
1	A	293	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	299	LEU
1	A	303	LEU
1	A	305	ASP
1	A	329	VAL
1	A	336	GLU
1	A	339	TYR
1	A	341	SER
1	A	342	ASP
1	A	346	ASP
1	A	351	GLN
1	A	374	ASP
1	A	376	ASP
1	A	402	LEU
1	A	435	GLU
1	A	440	THR
1	A	454	LEU
1	A	461	HIS
1	A	486	SER
1	A	501	HIS
1	A	502	ASP
1	A	506	SER
1	A	510	CYS
1	A	512	ASN
1	A	524	LEU
1	A	525	SER
1	A	538	GLN
1	A	540	LEU
1	A	1099	LYS
2	B	8	LEU
2	B	66	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	293	GLN
1	A	306	ASN
1	A	501	HIS
1	A	534	HIS
1	A	538	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

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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

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^{th} 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	356/387 (91%)	0.91	31 (8%)	13	12	76, 110, 158, 232	1 (0%)
2	B	76/76 (100%)	0.89	4 (5%)	30	29	70, 99, 115, 119	0
All	All	432/463 (93%)	0.91	35 (8%)	15	14	70, 106, 156, 232	1 (0%)

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	391	LYS	3.8
1	A	209	LYS	3.7
1	A	332	ILE	3.2
1	A	459	ASP	3.2
1	A	333	GLN	3.2
1	A	330	SER	3.1
1	A	390	GLU	2.8
1	A	340	ARG	2.7
1	A	495	GLU	2.6
1	A	497	ASN	2.6
1	A	379	TYR	2.5
1	A	454	LEU	2.4
2	B	56	LEU	2.4
1	A	267	LEU	2.3
1	A	383	GLU	2.3
2	B	43	LEU	2.3
1	A	210	HIS	2.3
1	A	518	TYR	2.3
1	A	336	GLU	2.2
1	A	501	HIS	2.2
1	A	303	LEU	2.2
1	A	450	LEU	2.2
2	B	63	LYS	2.1
1	A	502	ASP	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	512	ASN	2.1
1	A	389	ALA	2.1
1	A	419	ILE	2.1
2	B	5	VAL	2.1
1	A	331	TYR	2.1
1	A	231	THR	2.1
1	A	504	ASP	2.0
1	A	513	ALA	2.0
1	A	540	LEU	2.0
1	A	378	LYS	2.0
1	A	279	LEU	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^{th} 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(Å^2)	Q<0.9
3	CA	A	1201	1/1	0.89	0.41	-	148,148,148,148	1

6.5 Other polymers (i)

There are no such residues in this entry.