



Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 02:48 am BST

PDB ID : 2J7U
Title : Dengue virus NS5 RNA dependent RNA polymerase domain
Authors : Yap, T.L.; Xu, T.; Chen, Y.L.; Malet, H.; Egloff, M.P.; Canard, B.; Vasudevan, S.G.; Lescar, J.
Deposited on : 2006-10-17
Resolution : 1.85 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

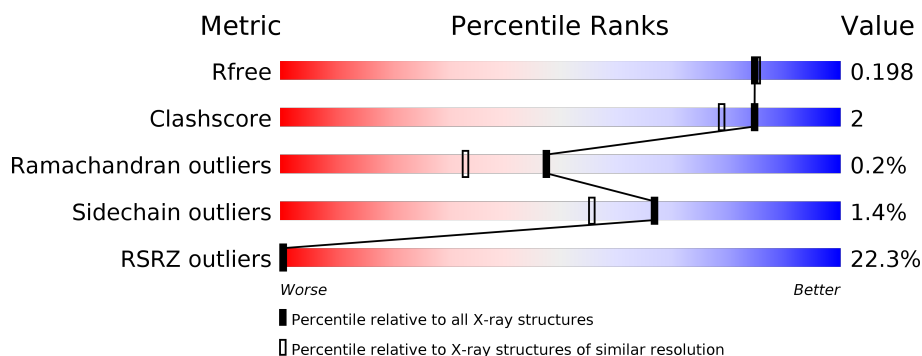
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 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	635	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 5220 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 RNA DEPENDENT RNA POLYMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	573	Total	C	N	O	S	0	11	0
			4701	2980	842	849	30			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	266	GLY	-	expression tag	UNP Q6DLV0
A	267	SER	-	expression tag	UNP Q6DLV0
A	268	HIS	-	expression tag	UNP Q6DLV0
A	269	MET	-	expression tag	UNP Q6DLV0
A	270	LEU	-	expression tag	UNP Q6DLV0
A	271	ASP	-	expression tag	UNP Q6DLV0
A	374	GLU	GLY	conflict	UNP Q6DLV0
A	480	ALA	VAL	conflict	UNP Q6DLV0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Zn	0	0
			2	2		

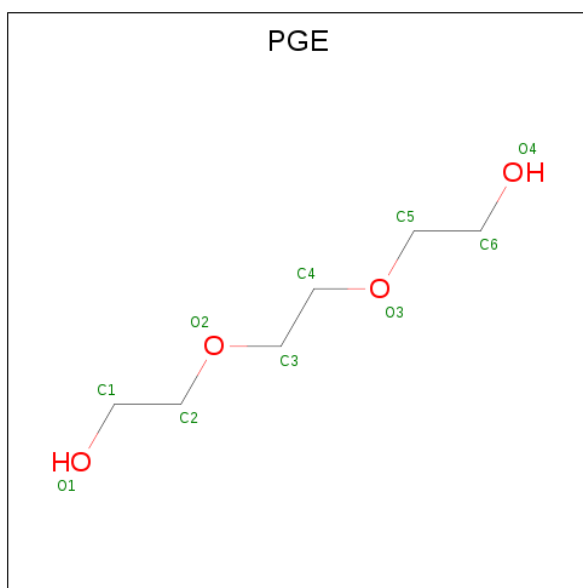
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Cl	0	0
			1	1		

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			10	6	4		

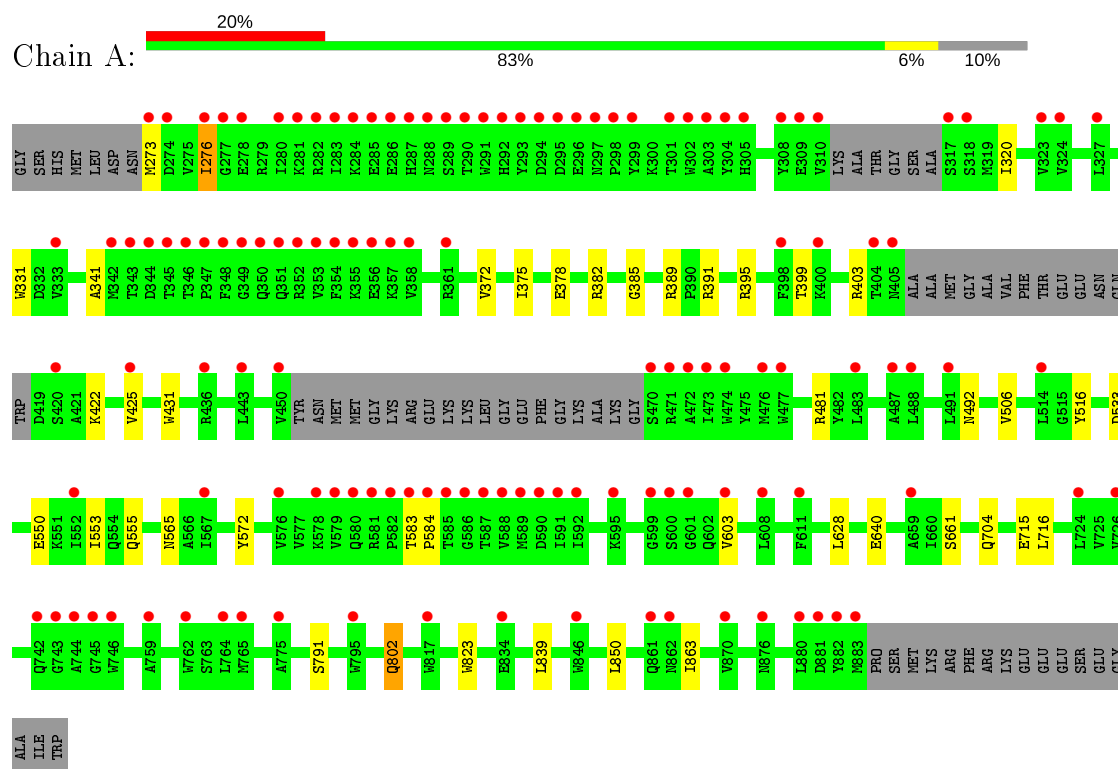
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	505	Total	O	0	0
			505	505		

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 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: RNA DEPENDENT RNA POLYMERASE



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	160.15Å 180.48Å 57.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.85 19.96 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.9 (20.00-1.85) 99.9 (19.96-1.85)	Depositor EDS
R_{merge}	0.01	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.66 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.200 , 0.234 0.200 , 0.198	Depositor DCC
R_{free} test set	3601 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	28.0	Xtriage
Anisotropy	0.516	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 43.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5220	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.61% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: ZN, PGE, MG, CL

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	0.35	0/4854	0.49	0/6576

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	4701	0	4598	21	0
2	A	2	0	0	0	0
3	A	1	0	0	0	0
4	A	1	0	0	1	0
5	A	10	0	14	0	0
6	A	505	0	0	0	0
All	All	5220	0	4612	21	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 2.

All (21) 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:716:LEU:HD21	1:A:839:LEU:HD23	1.60	0.83
1:A:802:GLN:HE21	1:A:802:GLN:H	1.30	0.77
1:A:385:GLY:HA3	1:A:555:GLN:HE22	1.64	0.61
1:A:704:GLN:NE2	1:A:715:GLU:H	2.05	0.55
1:A:403:ARG:HA	1:A:422:LYS:HD3	1.89	0.53
1:A:716:LEU:CD2	1:A:839:LEU:HD23	2.36	0.53
1:A:375:ILE:HD11	1:A:640:GLU:HG2	1.91	0.53
1:A:273:MET:HB3	1:A:276:ILE:HG22	1.96	0.47
1:A:572:TYR:OH	1:A:603:VAL:O	2.33	0.47
1:A:331:TRP:CE2	1:A:850:LEU:HD22	2.51	0.45
1:A:583:THR:HA	1:A:584:PRO:HD3	1.87	0.44
1:A:399:THR:HG23	1:A:425:VAL:CG1	2.47	0.44
1:A:395:ARG:HG3	1:A:431:TRP:CZ2	2.53	0.43
1:A:320:ILE:HD11	1:A:341:ALA:HB1	2.01	0.43
1:A:378:GLU:O	1:A:382:ARG:HG3	2.19	0.43
1:A:372[A]:VAL:HG11	1:A:628:LEU:HD11	2.00	0.43
1:A:492:ASN:ND2	4:A:1887:CL:CL	2.89	0.42
1:A:506:VAL:CG2	1:A:661[B]:SER:OG	2.68	0.41
1:A:550:GLU:O	1:A:553:ILE:HG12	2.21	0.41
1:A:516:TYR:HB3	1:A:823:TRP:CE2	2.57	0.40

There are no symmetry-related clashes.

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	576/635 (91%)	553 (96%)	22 (4%)	1 (0%)	47 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	791	SER

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	502/551 (91%)	495 (99%)	7 (1%)	67 55

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	276	ILE
1	A	389	ARG
1	A	391	ARG
1	A	481	ARG
1	A	533	ASP
1	A	802	GLN
1	A	863	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	339	GLN
1	A	387	ASN
1	A	548	ASN
1	A	555	GLN
1	A	645	GLN
1	A	701	HIS
1	A	704	GLN
1	A	760	GLN
1	A	768	HIS
1	A	802	GLN
1	A	835	ASN
1	A	869	GLN

5.3.3 RNA ⓘ

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 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	$\# Z > 2$
5	PGE	A	1888	-	9,9,9	0.43	0	8,8,8	0.61	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
5	PGE	A	1888	-	-	3/7/7/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1888	PGE	O2-C3-C4-O3
5	A	1888	PGE	O3-C5-C6-O4
5	A	1888	PGE	C6-C5-O3-C4

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, 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	573/635 (90%)	1.52	128 (22%) 0 0	28, 36, 42, 51	0

All (128) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	347	PRO	19.0
1	A	345	THR	17.2
1	A	583	THR	14.9
1	A	348	PHE	14.6
1	A	587	THR	14.2
1	A	353	VAL	13.3
1	A	293	TYR	12.9
1	A	346	THR	12.0
1	A	310	VAL	11.8
1	A	355	LYS	10.8
1	A	349	GLY	10.7
1	A	343	THR	9.9
1	A	354	PHE	9.3
1	A	582	PRO	9.2
1	A	291	TRP	8.8
1	A	344	ASP	8.6
1	A	580	GLN	8.3
1	A	304	TYR	8.2
1	A	317	SER	8.1
1	A	295	ASP	7.8
1	A	280	ILE	7.8
1	A	579	VAL	7.7
1	A	352	ARG	7.6
1	A	585	THR	7.4
1	A	289	SER	7.3
1	A	584	PRO	7.3
1	A	287	HIS	7.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	473	ILE	7.1
1	A	746	TRP	6.8
1	A	288	ASN	6.7
1	A	290	THR	6.7
1	A	305	HIS	6.5
1	A	586	GLY	6.5
1	A	358	VAL	6.4
1	A	292	HIS	6.4
1	A	273	MET	6.2
1	A	303	ALA	5.8
1	A	474	TRP	5.8
1	A	443	LEU	5.6
1	A	356	GLU	5.5
1	A	880	LEU	5.5
1	A	742	GLN	5.3
1	A	276	ILE	5.2
1	A	743	GLY	5.0
1	A	581	ARG	5.0
1	A	283	ILE	5.0
1	A	876	ASN	5.0
1	A	470	SER	4.9
1	A	285	GLU	4.9
1	A	342	MET	4.8
1	A	590	ASP	4.7
1	A	296	GLU	4.6
1	A	588	VAL	4.5
1	A	591	ILE	4.4
1	A	284	LYS	4.3
1	A	589	MET	4.3
1	A	350	GLN	4.3
1	A	881	ASP	4.3
1	A	333	VAL	4.2
1	A	795	TRP	4.2
1	A	302	TRP	4.1
1	A	277	GLY	4.1
1	A	318	SER	4.1
1	A	603	VAL	4.0
1	A	351	GLN	4.0
1	A	308	TYR	3.8
1	A	294	ASP	3.8
1	A	745	GLY	3.7
1	A	882	TYR	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	567	ILE	3.6
1	A	323	VAL	3.6
1	A	286	GLU	3.5
1	A	309	GLU	3.5
1	A	601	GLY	3.5
1	A	576	VAL	3.4
1	A	281	LYS	3.3
1	A	477	TRP	3.3
1	A	861	GLN	3.2
1	A	298	PRO	3.2
1	A	600	SER	3.2
1	A	299	TYR	3.1
1	A	301	THR	3.1
1	A	472	ALA	3.1
1	A	578	LYS	3.0
1	A	282	ARG	3.0
1	A	608	LEU	3.0
1	A	471	ARG	3.0
1	A	552	ILE	2.9
1	A	361	ARG	2.9
1	A	744	ALA	2.9
1	A	278	GLU	2.8
1	A	883	MET	2.8
1	A	297	ASN	2.8
1	A	357	LYS	2.8
1	A	592	ILE	2.8
1	A	405	ASN	2.7
1	A	491	LEU	2.7
1	A	450	VAL	2.7
1	A	595	LYS	2.7
1	A	327	LEU	2.7
1	A	324	VAL	2.7
1	A	599	GLY	2.6
1	A	483	LEU	2.6
1	A	476	MET	2.4
1	A	274	ASP	2.4
1	A	726	VAL	2.3
1	A	487	ALA	2.3
1	A	817	TRP	2.3
1	A	400	LYS	2.3
1	A	834	GLU	2.3
1	A	436	ARG	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	846	TRP	2.3
1	A	514	LEU	2.3
1	A	420	SER	2.2
1	A	775	ALA	2.2
1	A	659	ALA	2.2
1	A	425	VAL	2.2
1	A	488	LEU	2.2
1	A	611	PHE	2.2
1	A	862	ASN	2.1
1	A	765	MET	2.1
1	A	762	TRP	2.1
1	A	870	VAL	2.1
1	A	759	ALA	2.1
1	A	764	LEU	2.1
1	A	398	PHE	2.0
1	A	724	LEU	2.0
1	A	404	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 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. 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	RSCC	RSR	B-factors(\AA^2)	Q<0.9
5	PGE	A	1888	10/10	0.85	0.22	48,49,50,50	0
3	MG	A	1886	1/1	0.89	0.25	47,47,47,47	0
4	CL	A	1887	1/1	0.93	0.31	51,51,51,51	0
2	ZN	A	1885	1/1	0.97	0.08	31,31,31,31	1
2	ZN	A	1884	1/1	0.99	0.07	25,25,25,25	0

6.5 Other polymers [i](#)

There are no such residues in this entry.