Full wwPDB X-ray Structure Validation Report (i)

Feb 24, 2025 – 06:06 PM EST

PDB ID	:	9DEM
Title	:	USP7 in complex with macrocycle MC04
Authors	:	Rouge, L.; Ultsch, M.; Dueber, E.C.; Harris, S.F.
Deposited on	:	2024-08-29
Resolution	:	1.77 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

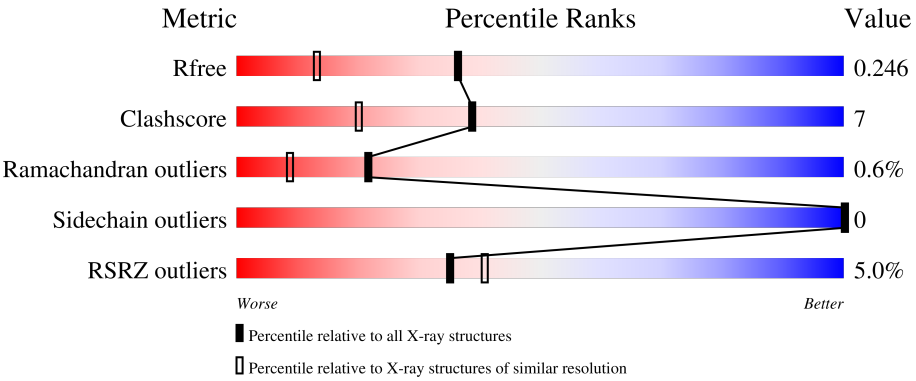
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

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 1.77 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}	164625	1191 (1.78-1.78)
Clashscore	180529	1282 (1.78-1.78)
Ramachandran outliers	177936	1270 (1.78-1.78)
Sidechain outliers	177891	1270 (1.78-1.78)
RSRZ outliers	164620	1191 (1.78-1.78)

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 of 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 <=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	368	5%, 74%, 14%, 10%
2	G	16	75%, 12%, 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2969 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	330	Total	C	N	O	S	0	0	0
			2680	1699	456	509	16			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	187	MET	-	initiating methionine	UNP Q93009
A	188	GLY	-	expression tag	UNP Q93009
A	189	SER	-	expression tag	UNP Q93009
A	190	SER	-	expression tag	UNP Q93009
A	191	HIS	-	expression tag	UNP Q93009
A	192	HIS	-	expression tag	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	SER	-	expression tag	UNP Q93009
A	198	SER	-	expression tag	UNP Q93009
A	199	GLY	-	expression tag	UNP Q93009
A	200	LEU	-	expression tag	UNP Q93009
A	201	VAL	-	expression tag	UNP Q93009
A	202	PRO	-	expression tag	UNP Q93009
A	203	ARG	-	expression tag	UNP Q93009
A	204	GLY	-	expression tag	UNP Q93009
A	205	SER	-	expression tag	UNP Q93009
A	206	HIS	-	expression tag	UNP Q93009
A	207	MET	-	expression tag	UNP Q93009

- Molecule 2 is a protein called Macrocycle peptide MC04.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	16	Total	C	N	O	S	0	0	0
			126	82	23	18	3			

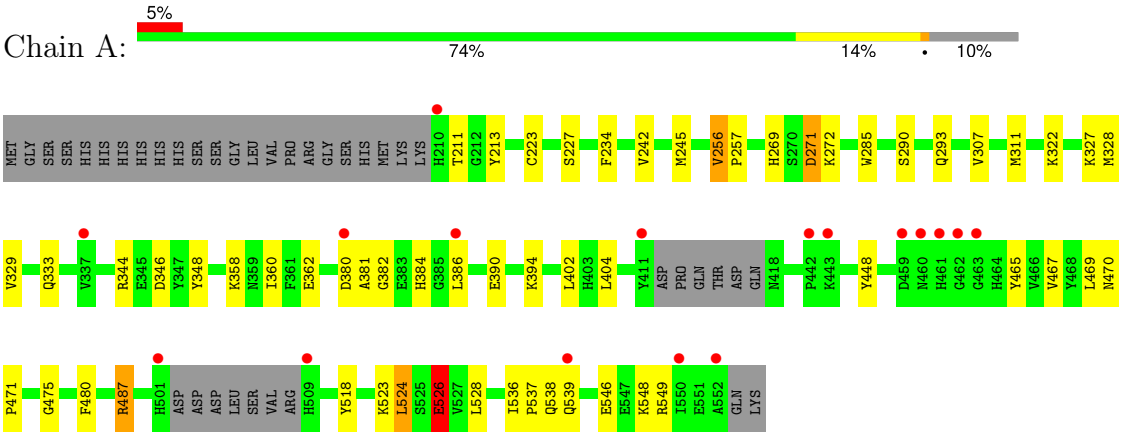
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	153	Total	O	0	0
			153	153		
3	G	10	Total	O	0	0
			10	10		

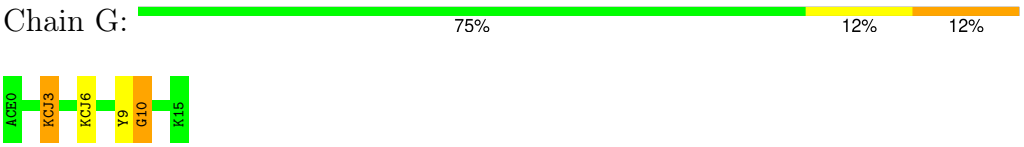
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Ubiquitin carboxyl-terminal hydrolase 7



• Molecule 2: Macrocycle peptide MC04



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	113.47Å 67.71Å 71.62Å 90.00° 125.22° 90.00°	Depositor
Resolution (Å)	35.73 – 1.77 35.73 – 1.77	Depositor EDS
% Data completeness (in resolution range)	65.9 (35.73-1.77) 65.9 (35.73-1.77)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.45 (at 1.77Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.187 , 0.237 0.196 , 0.246	Depositor DCC
R_{free} test set	40947 reflections (7.10%)	wwPDB-VP
Wilson B-factor (Å^2)	35.5	Xtriage
Anisotropy	0.013	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 35.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2969	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.01% 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: KCJ, DAL, NMC, ACE

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	0.72	3/2737 (0.1%)	1.47	20/3690 (0.5%)
2	G	0.63	0/89	0.89	0/111
All	All	0.72	3/2826 (0.1%)	1.45	20/3801 (0.5%)

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
1	A	0	4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	256	VAL	CB-CG1	-8.72	1.34	1.52
1	A	487	ARG	CB-CG	-7.09	1.33	1.52
1	A	526	GLU	CB-CG	5.36	1.62	1.52

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	271	ASP	CB-CG-OD1	45.31	159.08	118.30
1	A	271	ASP	CB-CG-OD2	-35.81	86.07	118.30
1	A	271	ASP	OD1-CG-OD2	-19.73	85.82	123.30
1	A	526	GLU	OE1-CD-OE2	-18.66	100.91	123.30
1	A	524	LEU	CB-CG-CD2	-15.86	84.04	111.00
1	A	526	GLU	CG-CD-OE1	15.13	148.56	118.30
1	A	526	GLU	CG-CD-OE2	-14.38	89.53	118.30
1	A	487	ARG	CG-CD-NE	-12.59	85.37	111.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	524	LEU	CB-CG-CD1	8.26	125.03	111.00
1	A	526	GLU	CA-CB-CG	7.59	130.10	113.40
1	A	524	LEU	CD1-CG-CD2	-6.78	90.16	110.50
1	A	272	LYS	CD-CE-NZ	6.41	126.45	111.70
1	A	487	ARG	NE-CZ-NH1	-6.25	117.17	120.30
1	A	487	ARG	N-CA-CB	-6.12	99.58	110.60
1	A	348	TYR	CB-CG-CD2	-6.04	117.37	121.00
1	A	256	VAL	CG1-CB-CG2	-6.03	101.25	110.90
1	A	402	LEU	CB-CG-CD2	-5.47	101.70	111.00
1	A	322	LYS	CD-CE-NZ	5.38	124.06	111.70
1	A	272	LYS	CB-CG-CD	-5.21	98.04	111.60
1	A	469	LEU	CB-CG-CD2	-5.12	102.29	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	271	ASP	Sidechain
1	A	293	GLN	Sidechain
1	A	526	GLU	Sidechain
1	A	537	PRO	Peptide

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	2680	0	2617	36	0
2	G	126	0	112	2	0
3	A	153	0	0	7	0
3	G	10	0	0	0	0
All	All	2969	0	2729	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 7.

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:329:VAL:HG22	1:A:344:ARG:HG2	1.64	0.80
1:A:390:GLU:O	3:A:601:HOH:O	1.99	0.79
1:A:470:ASN:HD22	1:A:475:GLY:H	1.41	0.67
1:A:381:ALA:HB3	1:A:384:HIS:HB2	1.79	0.62
1:A:394:LYS:NZ	3:A:604:HOH:O	2.26	0.60
1:A:333:GLN:HB3	3:A:601:HOH:O	2.03	0.59
1:A:211:THR:HG22	1:A:213:TYR:H	1.67	0.59
1:A:546:GLU:OE2	1:A:549:ARG:NH1	2.37	0.57
1:A:269:HIS:HE1	1:A:536:ILE:HD12	1.70	0.57
1:A:327:LYS:HE2	1:A:346:ASP:OD1	2.07	0.54
1:A:524:LEU:CD1	1:A:528:LEU:HD12	2.38	0.53
1:A:242:VAL:O	1:A:245:MET:HE2	2.09	0.53
1:A:448:TYR:HB3	1:A:518:TYR:HB3	1.91	0.52
1:A:539:GLN:HG2	3:A:748:HOH:O	2.10	0.52
1:A:358:LYS:HE3	1:A:362:GLU:OE2	2.10	0.51
1:A:329:VAL:HG22	1:A:344:ARG:CG	2.40	0.51
1:A:380:ASP:O	1:A:380:ASP:OD1	2.29	0.50
1:A:256:VAL:HG13	1:A:257:PRO:HD3	1.94	0.50
1:A:470:ASN:ND2	1:A:475:GLY:H	2.09	0.48
1:A:307:VAL:O	1:A:311:MET:HG3	2.14	0.47
1:A:256:VAL:HG12	3:A:716:HOH:O	2.14	0.47
1:A:360:ILE:HD11	1:A:404:LEU:HD13	1.97	0.46
1:A:227:SER:HB3	1:A:467:VAL:HB	1.97	0.46
1:A:487:ARG:NH2	1:A:487:ARG:HG3	2.31	0.45
1:A:524:LEU:HD12	1:A:528:LEU:HD12	1.99	0.45
2:G:9:TYR:HA	2:G:10:NMC:HCN1	1.82	0.45
1:A:256:VAL:CG1	1:A:257:PRO:HD3	2.47	0.45
1:A:523:LYS:O	1:A:526:GLU:HB3	2.18	0.44
1:A:223:CYS:HB3	1:A:465:TYR:CZ	2.53	0.43
1:A:328:MET:SD	2:G:3:KCJ:C15	3.07	0.43
1:A:548:LYS:NZ	3:A:608:HOH:O	2.50	0.43
1:A:390:GLU:HB2	3:A:601:HOH:O	2.19	0.43
1:A:285:TRP:HB3	1:A:290:SER:HB3	2.00	0.42
1:A:467:VAL:HG13	1:A:480:PHE:HB2	2.01	0.42
1:A:234:PHE:CD1	1:A:471:PRO:HB3	2.56	0.41
1:A:329:VAL:CG2	1:A:344:ARG:HG2	2.43	0.41
1:A:380:ASP:HB3	1:A:386:LEU:HD23	2.02	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	324/368 (88%)	304 (94%)	18 (6%)	2 (1%)	22 9
2	G	10/16 (62%)	10 (100%)	0	0	100 100
All	All	334/384 (87%)	314 (94%)	18 (5%)	2 (1%)	22 9

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	538	GLN
1	A	382	GLY

5.3.2 Protein sidechains (i)

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	297/332 (90%)	297 (100%)	0	100 100
2	G	9/9 (100%)	9 (100%)	0	100 100
All	All	306/341 (90%)	306 (100%)	0	100 100

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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	293	GLN
1	A	470	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

4 non-standard protein/DNA/RNA residues 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	KCJ	G	3	2	5,10,11	1.89	1 (20%)	2,12,14	3.67	2 (100%)
2	DAL	G	5	2	3,4,5	0.86	0	2,4,6	0.84	0
2	KCJ	G	6	2	5,10,11	3.46	1 (20%)	2,12,14	3.38	1 (50%)
2	NMC	G	10	2	7,8,9	1.19	0	7,9,11	1.32	1 (14%)

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	KCJ	G	3	2	-	1/5/6/8	0/1/1/1
2	DAL	G	5	2	-	1/1/2/4	-
2	KCJ	G	6	2	-	0/5/6/8	0/1/1/1
2	NMC	G	10	2	-	1/4/7/8	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	6	KCJ	C13-S14	-7.66	1.61	1.70
2	G	3	KCJ	C13-S14	-4.12	1.65	1.70

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	6	KCJ	C13-S14-C15	4.52	101.58	92.37
2	G	3	KCJ	C13-S14-C15	4.03	100.59	92.37
2	G	3	KCJ	C11-C12-C13	3.27	136.12	130.18
2	G	10	NMC	CX2-CX1-CN	-3.11	115.18	119.17

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	3	KCJ	O-C-CA-C11
2	G	5	DAL	O-C-CA-CB
2	G	10	NMC	C-CA-N-CN

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	3	KCJ	1	0
2	G	10	NMC	1	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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^{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	330/368 (89%)	0.27	17 (5%) 34 39	26, 38, 69, 106	0
2	G	11/16 (68%)	-0.28	0 100 100	26, 32, 45, 46	0
All	All	341/384 (88%)	0.26	17 (4%) 35 41	26, 37, 69, 106	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	411	TYR	5.8
1	A	550	ILE	5.0
1	A	210	HIS	4.1
1	A	462	GLY	3.9
1	A	501	HIS	3.6
1	A	461	HIS	3.4
1	A	460	ASN	3.3
1	A	380	ASP	3.1
1	A	552	ALA	3.1
1	A	509	HIS	2.7
1	A	442	PRO	2.7
1	A	539	GLN	2.7
1	A	443	LYS	2.4
1	A	386	LEU	2.1
1	A	459	ASP	2.1
1	A	337	VAL	2.1
1	A	463	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

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, 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	B-factors(Å^2)	Q<0.9
2	KCJ	G	3	10/11	0.93	0.09	28,38,44,45	0
2	DAL	G	5	5/6	0.95	0.07	29,29,38,38	0
2	NMC	G	10	8/9	0.97	0.06	24,27,33,35	0
2	KCJ	G	6	10/11	0.98	0.05	24,29,32,33	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.