Full wwPDB X-ray Structure Validation Report (i)

Jan 22, 2024 – 06:48 PM JST

PDB ID	:	8IPC
Title	:	The recombinant NZ-1 Fab complexed with the PDZ tandem fragment of A. aeolicus S2P homolog with the PA14 tag inserted between the residues 181 and 184
Authors	:	Adachi, Y.; Nogi, T.
Deposited on	:	2023-03-14
Resolution	:	2.20 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:

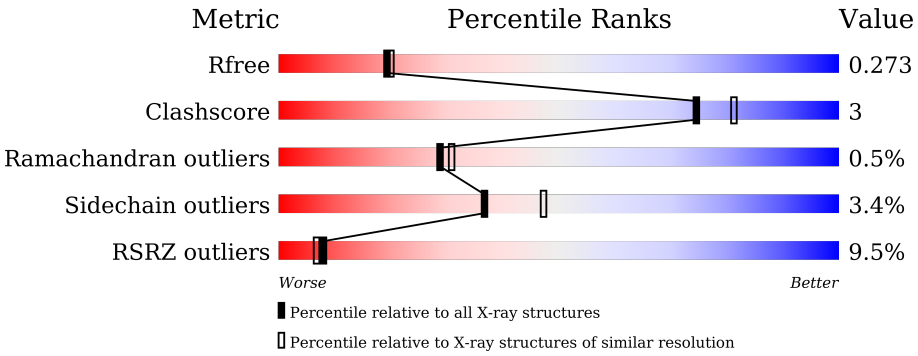
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	H	221	87% 9%
2	L	214	84% 14%
3	A	192	27% 88% 9%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4802 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 The recombinantly-expressed heavy chain of the monoclonal antibody NZ-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	213	Total	C	N	O	S	0	0	0
			1603	1012	266	316	9			

- Molecule 2 is a protein called The recombinantly-expressed light chain of the monoclonal antibody NZ-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	214	Total	C	N	O	S	0	0	0
			1641	1017	284	334	6			

- Molecule 3 is a protein called Putative zinc metalloprotease aq_1964.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	187	Total	C	N	O	S	0	0	0
			1452	940	245	265	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	113	GLY	-	expression tag	UNP O67776
A	114	SER	-	expression tag	UNP O67776
A	181A	GLU	-	insertion	UNP O67776
A	181B	GLY	-	insertion	UNP O67776
A	181C	GLY	-	insertion	UNP O67776
A	181D	VAL	-	insertion	UNP O67776
A	181E	ALA	-	insertion	UNP O67776
A	181F	MET	-	insertion	UNP O67776
A	181G	PRO	-	insertion	UNP O67776
A	181H	GLY	-	insertion	UNP O67776
A	181I	ALA	-	insertion	UNP O67776
A	181J	GLU	-	insertion	UNP O67776

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	181K	ASP	-	insertion	UNP O67776
A	181L	ASP	-	insertion	UNP O67776
A	181M	VAL	-	insertion	UNP O67776
A	181N	VAL	-	insertion	UNP O67776

- Molecule 4 is water.

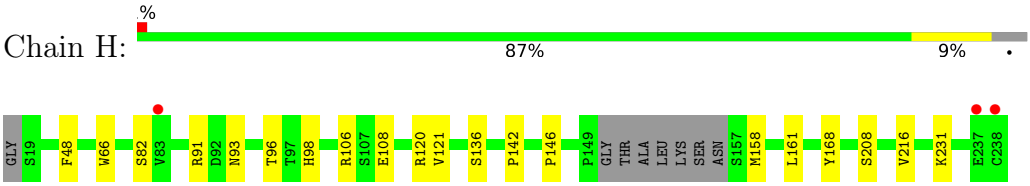
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	60	Total O 60 60	0	0
4	L	20	Total O 20 20	0	0
4	A	26	Total O 26 26	0	0

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.

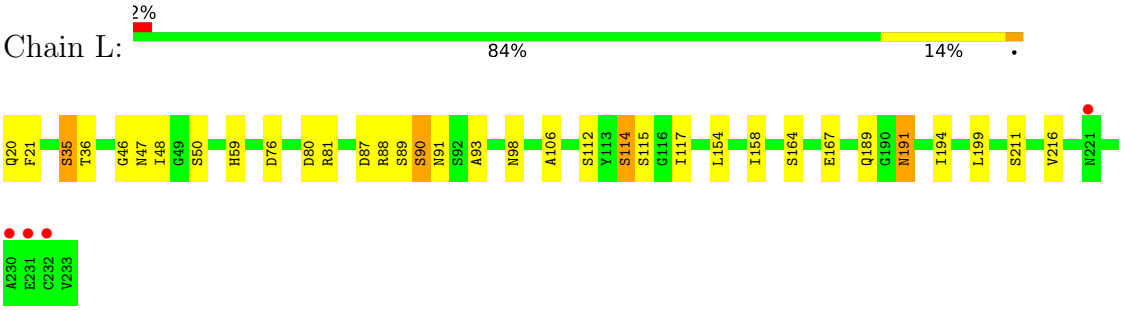
• Molecule 1: The recombinantly-expressed heavy chain of the monoclonal antibody NZ-1

Chain H:



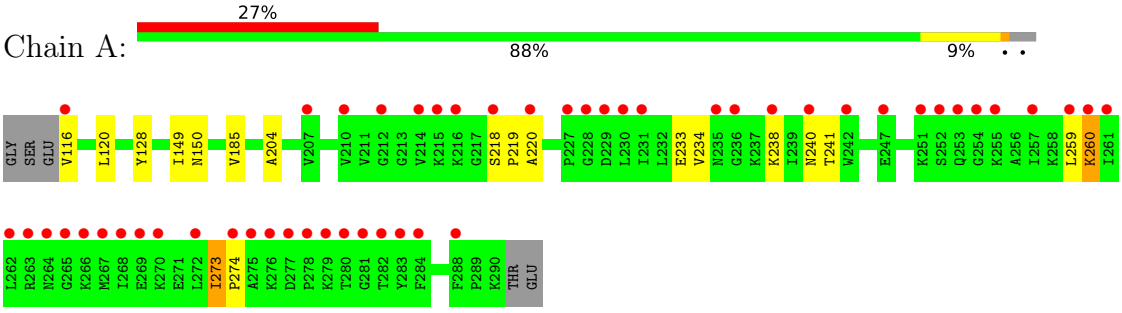
• Molecule 2: The recombinantly-expressed light chain of the monoclonal antibody NZ-1

Chain L:



• Molecule 3: Putative zinc metalloprotease aq_1964

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.94Å 75.09Å 173.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.32 – 2.20 38.32 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.32-2.20) 100.0 (38.32-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.90 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.243 , 0.270 0.246 , 0.273	Depositor DCC
R_{free} test set	1759 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	47.5	Xtriage
Anisotropy	0.425	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 37.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4802	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.14% 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: PCA, SNN

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	H	0.24	0/1642	0.49	0/2237
2	L	0.25	0/1669	0.48	0/2272
3	A	0.24	0/1467	0.48	0/1977
All	All	0.24	0/4778	0.48	0/6486

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
3	A	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	149	ILE	Mainchain,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	H	1603	0	1560	10	0
2	L	1641	0	1576	13	0
3	A	1452	0	1543	8	0
4	A	26	0	0	0	0
4	H	60	0	0	0	0
4	L	20	0	0	0	0
All	All	4802	0	4679	30	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 3.

All (30) 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)
3:A:233:GLU:HB2	3:A:260:LYS:HE2	1.74	0.70
2:L:114:SER:OG	2:L:115:SER:N	2.25	0.69
3:A:218:SER:O	3:A:220:ALA:N	2.26	0.67
1:H:142:PRO:HB3	1:H:168:TYR:HB3	1.79	0.62
2:L:191:ASN:OD1	2:L:191:ASN:N	2.22	0.62
2:L:46:GLY:O	2:L:91:ASN:ND2	2.41	0.53
3:A:234:VAL:HG12	3:A:259:LEU:HG	1.91	0.53
2:L:35:SER:OG	2:L:36:THR:N	2.42	0.52
1:H:146:PRO:HD3	1:H:231:LYS:HD2	1.93	0.51
3:A:128:TYR:HB2	3:A:241:THR:HG21	1.93	0.50
2:L:154:LEU:HB2	2:L:199:LEU:HB3	1.95	0.49
1:H:158:MET:HA	1:H:158:MET:HE2	1.95	0.49
2:L:158:ILE:HG12	2:L:216:VAL:HG21	1.95	0.48
1:H:120:ARG:HG3	1:H:121:VAL:HG13	1.96	0.48
2:L:47:ASN:O	2:L:50:SER:OG	2.31	0.47
2:L:48:ILE:HD11	2:L:93:ALA:HB2	1.95	0.47
1:H:91:ARG:HA	1:H:98:HIS:HA	1.95	0.47
3:A:128:TYR:HB3	3:A:204:ALA:HB3	1.97	0.47
3:A:233:GLU:HG2	3:A:238:LYS:HG2	1.97	0.45
2:L:81:ARG:HB3	2:L:98:ASN:O	2.18	0.43
1:H:106:ARG:NH2	1:H:108:GLU:OE1	2.46	0.42
2:L:87:ASP:OD2	2:L:90:SER:OG	2.38	0.41
1:H:161:LEU:HD12	1:H:216:VAL:HG11	2.01	0.41
1:H:66:TRP:CE2	2:L:117:ILE:HD12	2.55	0.41
3:A:116:VAL:HG13	3:A:120:LEU:HD12	2.02	0.41
2:L:189:GLN:HB2	2:L:194:ILE:HG12	2.01	0.41
2:L:59:HIS:CD2	2:L:106:ALA:HB2	2.55	0.41
1:H:91:ARG:NE	1:H:93:ASN:OD1	2.55	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:48:PHE:CD2	1:H:96:THR:HA	2.56	0.40
3:A:273:ILE:N	3:A:274:PRO:CD	2.85	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	H	209/221 (95%)	204 (98%)	5 (2%)	0	100	100
2	L	212/214 (99%)	204 (96%)	7 (3%)	1 (0%)	29	31
3	A	183/192 (95%)	169 (92%)	12 (7%)	2 (1%)	14	12
All	All	604/627 (96%)	577 (96%)	24 (4%)	3 (0%)	29	31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	114	SER
3	A	219	PRO
3	A	273	ILE

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	H	179/184 (97%)	176 (98%)	3 (2%)	60 74
2	L	187/187 (100%)	175 (94%)	12 (6%)	17 20
3	A	158/162 (98%)	155 (98%)	3 (2%)	57 71
All	All	524/533 (98%)	506 (97%)	18 (3%)	37 47

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	82	SER
1	H	136	SER
1	H	208	SER
2	L	21	PHE
2	L	35	SER
2	L	76	ASP
2	L	80	ASP
2	L	88	ARG
2	L	89	SER
2	L	90	SER
2	L	112	SER
2	L	164	SER
2	L	167	GLU
2	L	191	ASN
2	L	211	SER
3	A	185	VAL
3	A	240	ASN
3	A	260	LYS

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

Mol	Chain	Res	Type
1	H	128	GLN
1	H	187	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 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	PCA	L	20	2	7,8,9	1.81	1 (14%)	9,10,12	2.14	5 (55%)
3	SNN	A	150	3	7,8,8	5.32	5 (71%)	7,11,11	1.60	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	PCA	L	20	2	-	0/0/11/13	0/1/1/1
3	SNN	A	150	3	-	-	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	150	SNN	C-N1	9.45	1.49	1.37
3	A	150	SNN	C4-C5	9.08	1.63	1.51
2	L	20	PCA	CD-N	4.67	1.46	1.34
3	A	150	SNN	C5-N1	4.01	1.43	1.37
3	A	150	SNN	O-C	-2.24	1.18	1.23
3	A	150	SNN	O5-C5	-2.14	1.18	1.23

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	150	SNN	O5-C5-C4	3.36	130.83	126.39
2	L	20	PCA	OE-CD-CG	-3.04	121.46	126.76
2	L	20	PCA	CA-N-CD	-2.93	103.54	113.58
2	L	20	PCA	CB-CA-N	2.70	111.05	103.30
2	L	20	PCA	CG-CD-N	2.45	114.73	108.39
2	L	20	PCA	CB-CA-C	-2.41	109.39	112.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides 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	H	213/221 (96%)	0.06	3 (1%) 75 73	33, 46, 75, 123	0
2	L	213/214 (99%)	0.29	4 (1%) 66 65	37, 59, 80, 118	0
3	A	186/192 (96%)	1.23	51 (27%) 0 0	35, 66, 134, 149	0
All	All	612/627 (97%)	0.49	58 (9%) 8 7	33, 56, 119, 149	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	282	THR	10.1
3	A	216	LYS	9.7
3	A	278	PRO	8.9
3	A	254	GLY	7.5
3	A	279	LYS	7.1
2	L	232	CYS	6.9
1	H	238	CYS	6.5
3	A	214	VAL	5.8
3	A	252	SER	5.4
3	A	272	LEU	5.4
3	A	116	VAL	5.3
3	A	215	LYS	5.0
3	A	230	LEU	5.0
3	A	284	PHE	4.9
2	L	230	ALA	4.8
3	A	267	MET	4.8
3	A	262	LEU	4.7
2	L	231	GLU	4.6
3	A	210	VAL	4.6
3	A	264	ASN	4.3
3	A	212	GLY	4.2
3	A	277	ASP	4.0
3	A	268	ILE	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	A	288	PHE	3.9
1	H	237	GLU	3.9
3	A	261	ILE	3.8
3	A	207	VAL	3.5
3	A	280	THR	3.4
3	A	265	GLY	3.2
3	A	236	GLY	3.2
3	A	251	LYS	3.1
3	A	274	PRO	3.1
3	A	220	ALA	3.1
3	A	257	ILE	3.1
3	A	259	LEU	3.1
3	A	227	PRO	3.1
3	A	270	LYS	2.9
3	A	275	ALA	2.9
2	L	221	ASN	2.9
3	A	240	ASN	2.9
3	A	266	LYS	2.8
3	A	229	ASP	2.7
3	A	238	LYS	2.7
3	A	283	TYR	2.6
3	A	255	LYS	2.6
3	A	260	LYS	2.6
3	A	247	GLU	2.6
3	A	253	GLN	2.6
3	A	269	GLU	2.5
3	A	281	GLY	2.4
3	A	218	SER	2.3
3	A	228	GLY	2.3
3	A	263	ARG	2.2
3	A	242	TRP	2.1
3	A	276	LYS	2.1
1	H	83	VAL	2.1
3	A	235	ASN	2.0
3	A	231	ILE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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, 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	PCA	L	20	8/9	0.79	0.27	64,74,77,77	0
3	SNN	A	150	8/8	0.97	0.10	36,38,41,43	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.