Full wwPDB X-ray Structure Validation Report (i)

Dec 30, 2023 – 09:00 am GMT

PDB ID	:	7NZF
Title	:	Crystal structure of HLA-DR4 in complex with a mutated human collagen type II peptide
Authors	:	Ge, C.; Dobritzsch, D.; Holmdahl, R.
Deposited on	:	2021-03-24
Resolution	:	1.90 Å(reported)

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

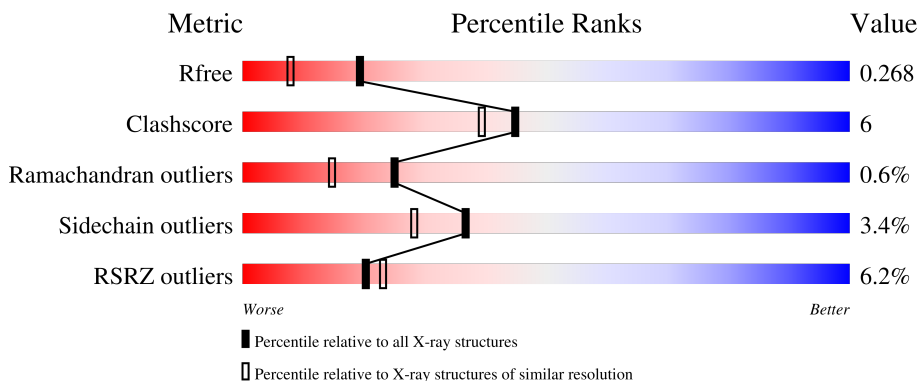|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

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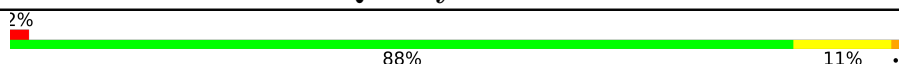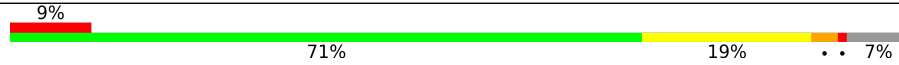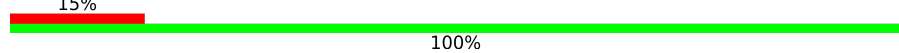*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	Percentile Ranks	Value
Rfree		0.268
Clashscore		6
Ramachandran outliers		0.6%
Sidechain outliers		3.4%
RSRZ outliers		6.2%

Worse — *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	AAA	180	2% / 88% / 11% / ·
2	BBB	191	9% / 71% / 19% / · · 7%
3	CCC	13	15% / 100%

2   Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 6129 atoms, of which 2908 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	180	2906	956	1430	241	274	5	36	0	0

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DR beta chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	BBB	177	2841	928	1378	255	275	5	46	0	0

- Molecule 3 is a protein called mutant human collagen type II,259-273.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
3	CCC	13	155	51	72	14	18	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	AAA	1	Total	C	H	N	O	3	0
			28	8	14	1	5		
4	AAA	1	Total	C	H	N	O	3	0
			28	8	14	1	5		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	94	Total	O	0	0
			94	94		
5	BBB	74	Total	O	0	0
			74	74		
5	CCC	3	Total	O	0	0
			3	3		

3 Residue-property plots ⓘ

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.
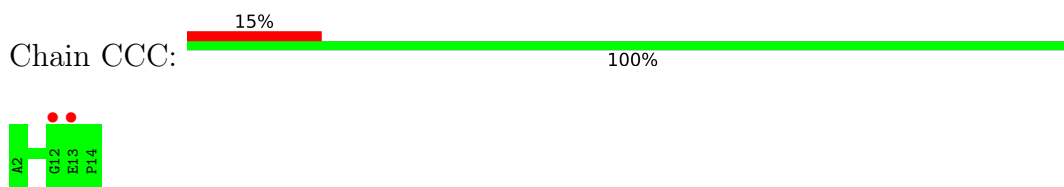
• Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain AAA:
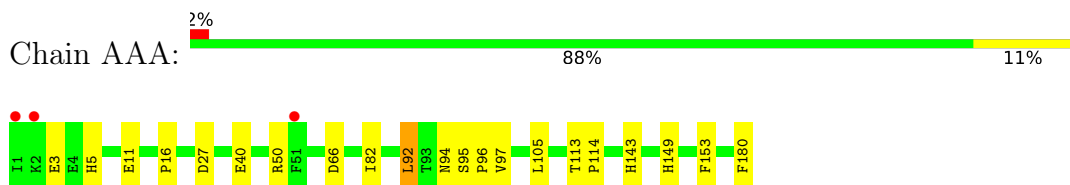
• Molecule 2: HLA class II histocompatibility antigen, DR beta chain

Chain BBB:
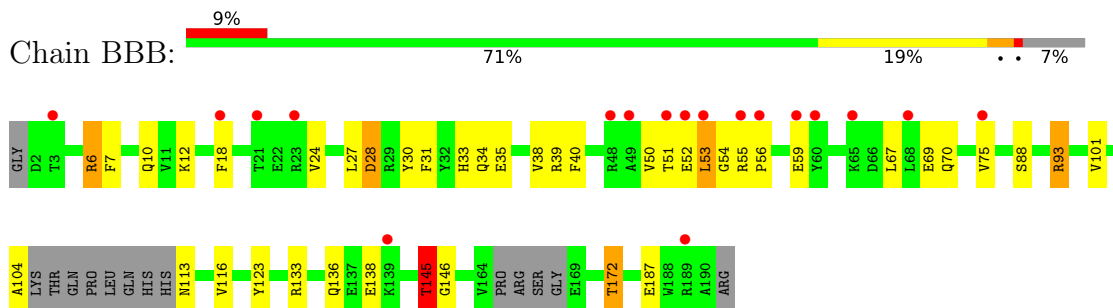
• Molecule 3: mutant human collagen type II,259-273

Chain CCC:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	71.44Å 71.44Å 138.05Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.95 – 1.90 36.92 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.4 (36.95-1.90) 98.5 (36.92-1.90)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.87 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.221 , 0.267 0.228 , 0.268	Depositor DCC
R_{free} test set	1564 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å^2)	27.1	Xtriage
Anisotropy	0.113	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 40.5	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.037 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6129	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.53% 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: NAG

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	AAA	0.71	1/1520 (0.1%)	0.85	1/2072 (0.0%)
2	BBB	0.70	0/1501	0.92	4/2038 (0.2%)
3	CCC	0.66	0/85	0.79	0/114
All	All	0.71	1/3106 (0.0%)	0.89	5/4224 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	40	GLU	CD-OE2	-5.08	1.20	1.25

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	BBB	6	ARG	NE-CZ-NH1	7.05	123.83	120.30
2	BBB	93	ARG	CG-CD-NE	-6.25	98.67	111.80
2	BBB	145	THR	N-CA-CB	-6.11	98.70	110.30
2	BBB	6	ARG	NE-CZ-NH2	-6.06	117.27	120.30
1	AAA	180	PHE	CA-C-O	-5.34	108.88	120.10

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	AAA	1476	1430	1424	13	1
2	BBB	1463	1378	1365	30	1
3	CCC	83	72	69	0	0
4	AAA	28	28	26	0	0
5	AAA	94	0	0	2	0
5	BBB	74	0	0	5	0
5	CCC	3	0	0	0	0
All	All	3221	2908	2884	38	1

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 6.

All (38) 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)
2:BBB:104:ALA:O	5:BBB:201:HOH:O	1.87	0.92
2:BBB:38:VAL:HG13	2:BBB:54:GLY:HA3	1.61	0.81
2:BBB:113:ASN:CB	5:BBB:257:HOH:O	2.43	0.66
2:BBB:133:ARG:NH2	5:BBB:203:HOH:O	2.29	0.64
2:BBB:138:GLU:HA	2:BBB:138:GLU:OE1	2.02	0.59
2:BBB:10:GLN:HB2	2:BBB:31:PHE:HB2	1.84	0.59
2:BBB:24:VAL:HB	2:BBB:75:VAL:HG13	1.83	0.59
1:AAA:5:HIS:HD2	1:AAA:27:ASP:OD2	1.86	0.58
2:BBB:145:THR:CG2	2:BBB:146:GLY:O	2.51	0.58
1:AAA:96:PRO:O	5:AAA:301:HOH:O	2.17	0.57
2:BBB:59:GLU:HB2	5:BBB:260:HOH:O	2.03	0.57
1:AAA:92:LEU:HD12	1:AAA:92:LEU:N	2.21	0.56
2:BBB:35:GLU:HG3	2:BBB:51:THR:HG21	1.90	0.54
2:BBB:145:THR:HG22	2:BBB:146:GLY:O	2.08	0.54
2:BBB:51:THR:O	2:BBB:53:LEU:N	2.40	0.53
2:BBB:7:PHE:HA	2:BBB:33:HIS:HE1	1.74	0.53
2:BBB:172:THR:HG23	5:BBB:246:HOH:O	2.08	0.53
2:BBB:35:GLU:CG	2:BBB:51:THR:HG21	2.40	0.52
1:AAA:92:LEU:HD12	1:AAA:92:LEU:H	1.76	0.51
2:BBB:172:THR:HB	2:BBB:187:GLU:HG3	1.93	0.50
2:BBB:55:ARG:HB2	2:BBB:56:PRO:HD3	1.93	0.49
1:AAA:3:GLU:HA	2:BBB:18:PHE:CD2	2.48	0.48
1:AAA:16:PRO:HD2	2:BBB:6:ARG:HD3	1.95	0.48
1:AAA:143:HIS:HD2	2:BBB:12:LYS:NZ	2.11	0.47
1:AAA:113:THR:OG1	1:AAA:114:PRO:HA	2.14	0.47
2:BBB:51:THR:O	2:BBB:51:THR:OG1	2.29	0.46
2:BBB:145:THR:HG23	2:BBB:146:GLY:O	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:105:LEU:HG	1:AAA:153:PHE:CE1	2.50	0.46
2:BBB:30:TYR:HB2	2:BBB:38:VAL:HG23	1.98	0.46
2:BBB:93:ARG:HG2	2:BBB:123:TYR:CD1	2.52	0.44
1:AAA:82:ILE:HG13	2:BBB:33:HIS:HB3	1.99	0.44
2:BBB:27:LEU:HA	2:BBB:40:PHE:O	2.17	0.44
2:BBB:28:ASP:O	2:BBB:39:ARG:HA	2.18	0.43
1:AAA:143:HIS:HD2	2:BBB:12:LYS:HZ2	1.66	0.43
1:AAA:149:HIS:HD2	5:AAA:389:HOH:O	2.02	0.43
1:AAA:11:GLU:OE1	1:AAA:66:ASP:OD2	2.38	0.42
2:BBB:101:VAL:HA	2:BBB:116:VAL:O	2.21	0.41
2:BBB:67:LEU:O	2:BBB:70:GLN:HB2	2.21	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:50:ARG:HE	2:BBB:69:GLU:OE1[6_654]	1.57	0.03

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	AAA	178/180 (99%)	177 (99%)	1 (1%)	0	100	100
2	BBB	171/191 (90%)	161 (94%)	8 (5%)	2 (1%)	13	4
3	CCC	11/13 (85%)	11 (100%)	0	0	100	100
All	All	360/384 (94%)	349 (97%)	9 (2%)	2 (1%)	25	15

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	BBB	52	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	BBB	50	VAL

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	AAA	164/164 (100%)	160 (98%)	4 (2%)	49	43
2	BBB	157/172 (91%)	150 (96%)	7 (4%)	27	18
3	CCC	6/6 (100%)	6 (100%)	0	100	100
All	All	327/342 (96%)	316 (97%)	11 (3%)	37	28

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	92	LEU
1	AAA	94	ASN
1	AAA	95	SER
1	AAA	97	VAL
2	BBB	28	ASP
2	BBB	34	GLN
2	BBB	53	LEU
2	BBB	88	SER
2	BBB	136	GLN
2	BBB	145	THR
2	BBB	172	THR

Sometimes 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 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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 ligands 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	#$|Z| > 2$
4	NAG	AAA	202	1	14,14,15	0.57	0	17,19,21	1.80	3 (17%)
4	NAG	AAA	201	1	14,14,15	0.67	0	17,19,21	2.27	5 (29%)

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
4	NAG	AAA	202	1	-	0/6/23/26	0/1/1/1
4	NAG	AAA	201	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	AAA	202	NAG	C1-O5-C5	5.57	119.74	112.19
4	AAA	201	NAG	C2-N2-C7	4.98	129.99	122.90
4	AAA	201	NAG	C1-O5-C5	4.61	118.44	112.19
4	AAA	201	NAG	O5-C5-C6	4.08	113.60	107.20

Continued on next page...

WORLDWIDE PDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	AAA	201	NAG	C1-C2-N2	3.59	116.62	110.49
4	AAA	202	NAG	C6-C5-C4	-2.30	107.61	113.00
4	AAA	201	NAG	O7-C7-N2	2.10	125.81	121.95
4	AAA	202	NAG	C3-C4-C5	2.04	113.88	110.24

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	AAA	201	NAG	O5-C5-C6-O6
4	AAA	201	NAG	C4-C5-C6-O6
4	AAA	201	NAG	C3-C2-N2-C7

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	AAA	180/180 (100%)	0.27	3 (1%)	70	72	17, 29, 55, 102	0
2	BBB	177/191 (92%)	0.44	18 (10%)	6	8	18, 38, 75, 88	4 (2%)
3	CCC	13/13 (100%)	0.72	2 (15%)	2	2	36, 44, 69, 72	0
All	All	370/384 (96%)	0.37	23 (6%)	20	23	17, 33, 72, 102	4 (1%)

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	1	ILE	18.2
1	AAA	2	LYS	5.9
2	BBB	51	THR	4.5
2	BBB	21	THR	4.3
2	BBB	75	VAL	3.8
3	CCC	13	GLU	3.0
2	BBB	68	LEU	3.0
2	BBB	48	ARG	2.9
2	BBB	139	LYS	2.8
2	BBB	55	ARG	2.8
2	BBB	65	LYS	2.8
1	AAA	51	PHE	2.8
2	BBB	3	THR	2.7
2	BBB	189	ARG	2.5
2	BBB	53	LEU	2.5
2	BBB	52	GLU	2.5
2	BBB	56	PRO	2.4
2	BBB	49	ALA	2.2
2	BBB	23	ARG	2.2
2	BBB	18	PHE	2.1
3	CCC	12	GLY	2.1
2	BBB	59	GLU	2.1
2	BBB	60	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 monosaccharides 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. 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(Å^2)	Q<0.9
4	NAG	AAA	201	14/15	0.79	0.16	30,80,83,88	3
4	NAG	AAA	202	14/15	0.85	0.15	30,42,49,50	3

6.5 Other polymers ⓘ

There are no such residues in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK