



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 07:54 AM UTC

PDB ID : 5GND / pdb_00005gnd
Title : Structure of Deg protease HhoA from Synechocystis sp. PCC 6803
Authors : Dong, W.; Wang, J.; Liu, L.
Deposited on : 2016-07-20
Resolution : 2.50 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

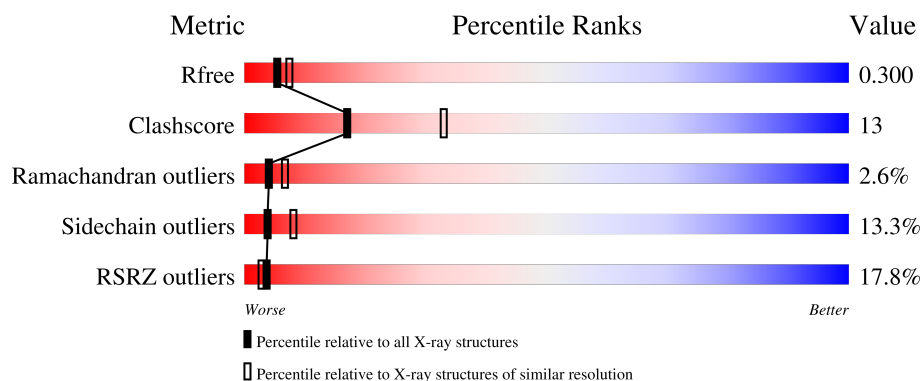
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	180053	5829 (2.50-2.50)
Clashscore	190562	6492 (2.50-2.50)
Ramachandran outliers	187476	6378 (2.50-2.50)
Sidechain outliers	187428	6380 (2.50-2.50)
RSRZ outliers	180081	5833 (2.50-2.50)

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 $\leq 5\%$. 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	348	
2	U	5	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2143 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 Putative serine protease HhoA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	280	Total	C	N	O	S	0	2	0
			2047	1280	369	395	3			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	55	MET	-	expression tag	UNP P72780
A	395	LEU	-	expression tag	UNP P72780
A	396	GLU	-	expression tag	UNP P72780
A	397	HIS	-	expression tag	UNP P72780
A	398	HIS	-	expression tag	UNP P72780
A	399	HIS	-	expression tag	UNP P72780
A	400	HIS	-	expression tag	UNP P72780
A	401	HIS	-	expression tag	UNP P72780
A	402	HIS	-	expression tag	UNP P72780

- Molecule 2 is a protein called UNK-UNK-UNK-UNK-TRP.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	U	5	Total	C	N	O	0	0	0
			30	19	6	5			

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

- Molecule 4 is SODIUM ION (CCD ID: NA) (formula: Na).

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

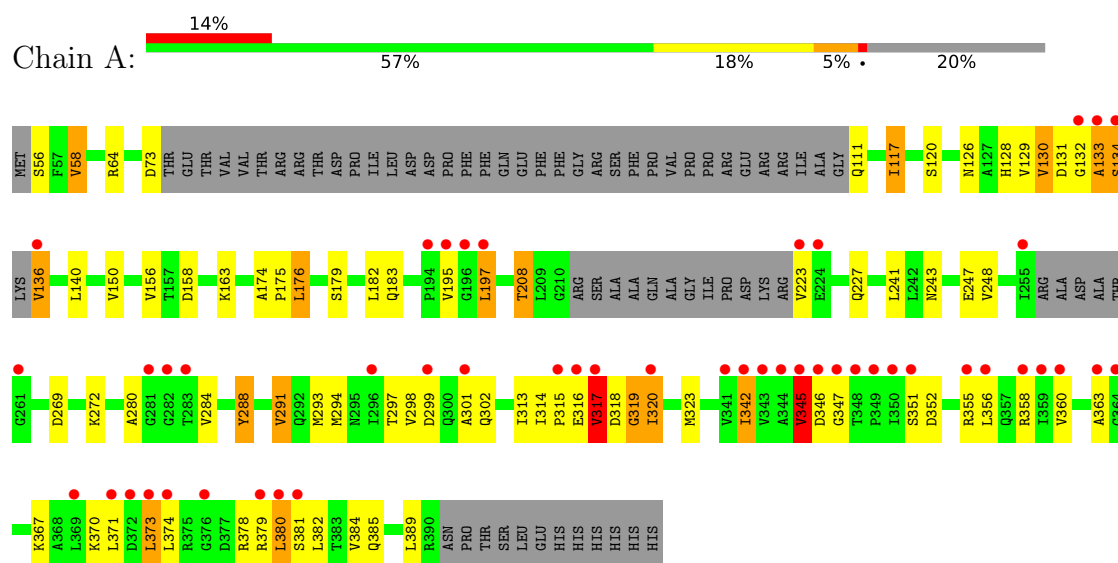
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	63	Total 63	O 63	0	0
5	U	1	Total 1	O 1	0	0

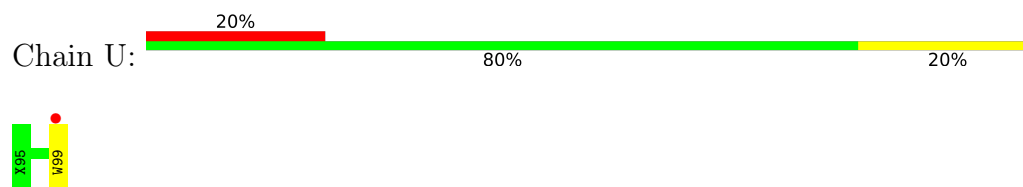
3 Residue-property plots [i](#)

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: Putative serine protease HhoA



- Molecule 2: UNK-UNK-UNK-UNK-TRP



4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	135.22Å 135.22Å 113.75Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.92 – 2.50 37.92 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.0 (37.92-2.50) 96.9 (37.92-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.25 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.247 , 0.305 0.253 , 0.300	Depositor DCC
R_{free} test set	676 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å ²)	32.7	Xtriage
Anisotropy	0.037	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 51.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	2143	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.03% 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, NA

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	0.47	2/2073 (0.1%)	0.95	7/2823 (0.2%)
2	U	0.25	0/15	0.22	0/20
All	All	0.47	2/2088 (0.1%)	0.95	7/2843 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	130	VAL	C-N	7.97	1.44	1.33
1	A	136	VAL	C-N	6.97	1.42	1.33

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	133	ALA	N-CA-C	11.90	136.14	110.80
1	A	134	SER	N-CA-CB	11.12	129.41	110.50
1	A	136	VAL	O-C-N	9.42	138.07	123.00
1	A	136	VAL	CA-C-N	-7.82	111.41	122.75
1	A	136	VAL	C-N-CA	-7.82	111.41	122.75
1	A	130	VAL	O-C-N	7.00	131.32	122.57
1	A	319	GLY	N-CA-C	5.36	118.08	110.74

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	2047	0	2087	53	0
2	U	30	0	12	5	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	63	0	0	10	1
5	U	1	0	0	0	0
All	All	2143	0	2099	53	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 13.

All (53) 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:342:ILE:O	5:A:601:HOH:O	2.04	0.74
1:A:320:ILE:O	5:A:602:HOH:O	2.11	0.69
1:A:130:VAL:C	1:A:132:GLY:H	2.05	0.65
1:A:111:GLN:N	5:A:609:HOH:O	2.31	0.64
1:A:297:THR:HG22	1:A:299:ASP:H	1.63	0.64
1:A:243:ASN:HD21	1:A:247:GLU:HG3	1.67	0.60
1:A:363:ALA:HB1	1:A:367:LYS:HG2	1.84	0.59
1:A:297:THR:O	1:A:301:ALA:N	2.35	0.58
1:A:130:VAL:HG11	1:A:150:VAL:HG21	1.86	0.58
1:A:130:VAL:C	1:A:132:GLY:N	2.62	0.58
1:A:208:THR:HG23	1:A:227:GLN:HB2	1.86	0.57
1:A:130:VAL:O	1:A:132:GLY:N	2.31	0.57
1:A:288:TYR:HD1	2:U:99:TRP:HB3	1.69	0.57
1:A:117:ILE:O	5:A:603:HOH:O	2.18	0.56
1:A:293:MET:HE1	1:A:356:LEU:HB3	1.90	0.53
1:A:130:VAL:CG1	1:A:150:VAL:HG21	2.38	0.53
1:A:126:ASN:HB2	1:A:129:VAL:HG23	1.91	0.53
1:A:117:ILE:HD11	1:A:176:LEU:HD11	1.91	0.51
1:A:378:ARG:HD2	1:A:380:LEU:HD22	1.91	0.51
1:A:342:ILE:N	5:A:602:HOH:O	2.42	0.51
1:A:288:TYR:HB2	2:U:99:TRP:HE3	1.75	0.51
1:A:163:LYS:NZ	5:A:615:HOH:O	2.44	0.50
1:A:360:VAL:HG21	2:U:99:TRP:HH2	1.76	0.50
1:A:73:ASP:OD2	1:A:111:GLN:NE2	2.45	0.49
1:A:319:GLY:HA2	1:A:351:SER:HA	1.94	0.48
1:A:371:LEU:N	1:A:382:LEU:O	2.46	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:288:TYR:HB2	2:U:99:TRP:CE3	2.48	0.48
1:A:379:ARG:HG2	5:A:627:HOH:O	2.12	0.48
1:A:297:THR:HG22	1:A:299:ASP:N	2.27	0.48
1:A:298:VAL:O	1:A:302:GLN:HG2	2.13	0.48
1:A:346[A]:ASP:HA	1:A:370:LYS:HD3	1.96	0.47
1:A:64:ARG:O	5:A:604:HOH:O	2.20	0.47
1:A:294:MET:HB2	1:A:323:MET:HG3	1.97	0.47
1:A:293:MET:HB2	1:A:320:ILE:HG22	1.97	0.47
1:A:371:LEU:O	1:A:381:SER:HA	2.15	0.47
1:A:345:VAL:HG23	1:A:346[B]:ASP:H	1.80	0.46
1:A:352:ASP:HB3	1:A:355:ARG:HB3	1.99	0.45
1:A:293:MET:SD	1:A:356:LEU:HD23	2.57	0.45
1:A:302:GLN:HB3	1:A:316:GLU:OE1	2.17	0.45
1:A:291:VAL:O	2:U:99:TRP:CD1	2.70	0.44
1:A:298:VAL:N	1:A:318:ASP:OD2	2.38	0.43
1:A:269:ASP:HA	1:A:272:LYS:HD2	2.00	0.43
1:A:356:LEU:C	1:A:358:ARG:H	2.26	0.43
1:A:342:ILE:HA	1:A:373:LEU:HB3	2.00	0.42
1:A:342:ILE:O	5:A:602:HOH:O	2.20	0.42
1:A:320:ILE:HG13	5:A:602:HOH:O	2.19	0.42
1:A:243:ASN:ND2	1:A:247:GLU:HG3	2.34	0.42
1:A:314:ILE:HA	1:A:315:PRO:HD3	1.95	0.42
1:A:56:SER:OG	1:A:58:VAL:HG13	2.20	0.41
1:A:128:HIS:ND1	1:A:158:ASP:OD2	2.46	0.41
1:A:174:ALA:HA	1:A:175:PRO:HD2	1.97	0.41
1:A:179:SER:HA	1:A:182:LEU:HG	2.02	0.41
1:A:315:PRO:O	1:A:317:VAL:HG13	2.22	0.40

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)
5:A:653:HOH:O	5:A:653:HOH:O[18_445]	1.58	0.62

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	272/348 (78%)	250 (92%)	15 (6%)	7 (3%)	4 7

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	131	ASP
1	A	133	ALA
1	A	317	VAL
1	A	345	VAL
1	A	347	GLY
1	A	280	ALA
1	A	197	LEU

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	219/278 (79%)	190 (87%)	29 (13%)	4 8
2	U	1/1 (100%)	1 (100%)	0	100 100
All	All	220/279 (79%)	191 (87%)	29 (13%)	4 8

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	VAL
1	A	117	ILE
1	A	120	SER
1	A	134	SER
1	A	136	VAL
1	A	140	LEU
1	A	156	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	176	LEU
1	A	183	GLN
1	A	195	VAL
1	A	197	LEU
1	A	208	THR
1	A	223	VAL
1	A	241	LEU
1	A	248	VAL
1	A	284	VAL
1	A	288	TYR
1	A	291	VAL
1	A	313	ILE
1	A	317	VAL
1	A	320	ILE
1	A	342	ILE
1	A	345	VAL
1	A	373	LEU
1	A	374	LEU
1	A	380	LEU
1	A	384	VAL
1	A	385	GLN
1	A	389	LEU

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

Mol	Chain	Res	Type
1	A	126	ASN
1	A	233	ASN
1	A	252	ASN
1	A	362	GLN
1	A	385	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 [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are 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 [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, 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	280/348 (80%)	0.84	49 (17%) 4 3	11, 34, 64, 83	2 (0%)
2	U	1/5 (20%)	4.61	1 (100%) 0 0	81, 81, 81, 81	0
All	All	281/353 (79%)	0.86	50 (17%) 4 3	11, 34, 64, 83	2 (0%)

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	346[A]	ASP	5.8
1	A	344	ALA	5.6
1	A	316	GLU	5.5
1	A	345	VAL	5.2
1	A	223	VAL	5.1
1	A	343	VAL	4.9
1	A	347	GLY	4.8
2	U	99	TRP	4.6
1	A	134	SER	4.6
1	A	356	LEU	4.1
1	A	136	VAL	3.9
1	A	350	ILE	3.9
1	A	282	GLY	3.8
1	A	132	GLY	3.8
1	A	360	VAL	3.7
1	A	320	ILE	3.7
1	A	373	LEU	3.6
1	A	299	ASP	3.5
1	A	196	GLY	3.3
1	A	261	GLY	3.2
1	A	349	PRO	3.2
1	A	133	ALA	3.0
1	A	369	LEU	3.0
1	A	379	ARG	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	355	ARG	2.9
1	A	374	LEU	2.9
1	A	342	ILE	2.9
1	A	371	LEU	2.8
1	A	296	ILE	2.8
1	A	317	VAL	2.7
1	A	341	VAL	2.7
1	A	359	ILE	2.6
1	A	348	THR	2.6
1	A	358	ARG	2.5
1	A	255	ILE	2.4
1	A	195	VAL	2.4
1	A	363	ALA	2.4
1	A	315	PRO	2.3
1	A	197	LEU	2.3
1	A	380	LEU	2.2
1	A	381	SER	2.2
1	A	364	GLY	2.2
1	A	351	SER	2.2
1	A	281	GLY	2.1
1	A	376	GLY	2.1
1	A	194	PRO	2.1
1	A	224	GLU	2.1
1	A	301	ALA	2.0
1	A	283	THR	2.0
1	A	372	ASP	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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
4	NA	A	502	1/1	0.88	0.10	21,21,21,21	1
3	ZN	A	501	1/1	0.94	0.06	33,33,33,33	1

6.5 Other polymers [i](#)

There are no such residues in this entry.