



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 10:02 AM UTC

PDB ID : 8CWS / pdb_00008cws
Title : Accurate computational design of genetically encoded 3D protein crystals
Authors : Bera, A.K.; Li, Z.; Baker, D.
Deposited on : 2022-05-19
Resolution : 4.40 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

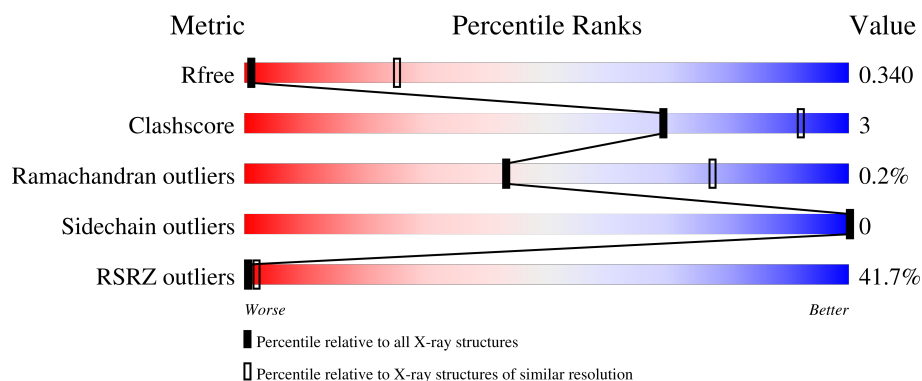
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.40 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	1088 (4.90-3.90)
Clashscore	190562	1135 (4.90-3.90)
Ramachandran outliers	187476	1026 (4.90-3.90)
Sidechain outliers	187428	1010 (4.90-3.90)
RSRZ outliers	180081	1084 (4.90-3.90)

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 $\leq 5\%$. 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	466	42% 88% 9% .
2	B	69	25% 75% . 20%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3920 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 F4132-2 Chain A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	451	Total	C	N	O	S	0	0	0
			3458	2163	599	692	4			

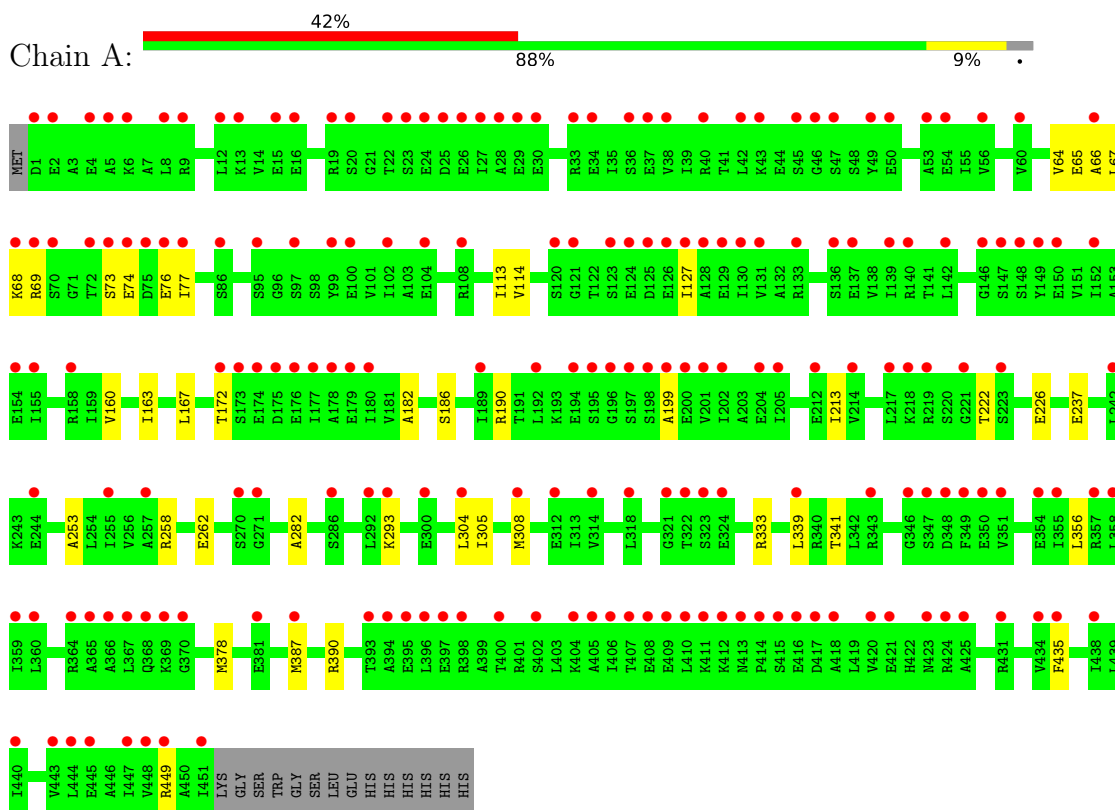
- Molecule 2 is a protein called F4132-2 Chain B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	55	Total	C	N	O	S	0	0	0
			462	286	85	88	3			

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: F4132-2 Chain A



4 Data and refinement statistics

Property	Value	Source
Space group	F 41 3 2	Depositor
Cell constants a, b, c, α , β , γ	405.41 Å 405.41 Å 405.41 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.16 – 4.40 49.16 – 4.40	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.16-4.40) 99.5 (49.16-4.40)	Depositor EDS
R_{merge}	2.57	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.08 (at 4.45 Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.323 , 0.342 0.321 , 0.340	Depositor DCC
R_{free} test set	1865 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å ²)	104.3	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 999.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.19$, $\langle L^2 \rangle = 0.07$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.68	EDS
Total number of atoms	3920	wwPDB-VP
Average B, all atoms (Å ²)	207.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.01% 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

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	0.08	0/3466	0.20	0/4674
2	B	0.08	0/461	0.22	0/611
All	All	0.08	0/3927	0.21	0/5285

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	3458	0	3678	27	0
2	B	462	0	496	2	0
All	All	3920	0	4174	28	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 (28) 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:73:SER:HB3	1:A:76:GLU:HB2	1.68	0.74
1:A:73:SER:HB2	1:A:77:ILE:HG13	1.75	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:199:ALA:N	2:B:52:GLU:OE2	2.30	0.65
1:A:64:VAL:HG11	1:A:113:ILE:HG23	1.84	0.58
1:A:387:MET:SD	1:A:390:ARG:NH1	2.79	0.56
1:A:339:LEU:HD13	1:A:435:PHE:HZ	1.70	0.55
1:A:186:SER:OG	1:A:190:ARG:NH1	2.40	0.54
1:A:66:ALA:HA	1:A:69:ARG:HE	1.72	0.54
1:A:167:LEU:HD22	1:A:172:THR:HG21	1.94	0.49
1:A:67:LEU:HB3	1:A:77:ILE:HG12	1.95	0.48
1:A:282:ALA:HB2	1:A:333:ARG:HG3	1.95	0.47
1:A:160:VAL:HG13	1:A:213:ILE:HD11	1.95	0.47
1:A:182:ALA:HB1	1:A:237:GLU:HG3	1.97	0.47
1:A:258:ARG:HH12	1:A:262:GLU:HB2	1.81	0.46
1:A:68:LYS:HA	1:A:68:LYS:HD2	1.77	0.45
1:A:339:LEU:HD11	1:A:356:LEU:HD21	1.99	0.45
1:A:114:VAL:HG22	1:A:127:ILE:HG23	1.98	0.44
1:A:378:MET:HE1	1:A:449:ARG:NH1	2.33	0.44
1:A:114:VAL:HG21	1:A:163:ILE:HD12	1.99	0.43
1:A:293:LYS:NZ	1:A:341:THR:OG1	2.43	0.43
2:B:23:LEU:HD11	2:B:50:ILE:HG23	2.00	0.43
1:A:222:THR:HB	1:A:226:GLU:HB3	2.01	0.43
1:A:253:ALA:HB2	1:A:305:ILE:HG12	2.00	0.43
1:A:339:LEU:HD21	1:A:356:LEU:HD21	2.02	0.42
1:A:339:LEU:HD13	1:A:435:PHE:CZ	2.52	0.42
1:A:167:LEU:HB3	1:A:172:THR:HB	2.01	0.41
1:A:65:GLU:O	1:A:69:ARG:HG3	2.21	0.41
1:A:304:LEU:HG	1:A:308:MET:HE2	2.02	0.41

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	449/466 (96%)	446 (99%)	2 (0%)	1 (0%)	43	77
2	B	53/69 (77%)	53 (100%)	0	0	100	100
All	All	502/535 (94%)	499 (99%)	2 (0%)	1 (0%)	43	77

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	74	GLU

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	380/393 (97%)	380 (100%)	0	100	100
2	B	52/64 (81%)	52 (100%)	0	100	100
All	All	432/457 (94%)	432 (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 (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	436	ASN

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](#)

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, 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	451/466 (96%)	2.08	194 (43%) 0 2	105, 202, 305, 370	0
2	B	55/69 (79%)	1.38	17 (30%) 1 4	110, 183, 296, 337	0
All	All	506/535 (94%)	2.01	211 (41%) 0 2	105, 199, 305, 370	0

All (211) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	408	GLU	11.7
1	A	124	GLU	11.4
1	A	15	GLU	9.8
1	A	396	LEU	9.1
1	A	45	SER	8.8
2	B	12	SER	8.6
1	A	30	GLU	8.4
1	A	126	GLU	8.2
1	A	397	GLU	8.2
1	A	46	GLY	7.9
1	A	421	GLU	7.7
2	B	11	ARG	7.6
1	A	176	GLU	7.4
1	A	19	ARG	7.1
1	A	411	LYS	6.8
1	A	404	LYS	6.8
1	A	180	ILE	6.6
1	A	128	ALA	6.6
1	A	175	ASP	6.1
1	A	152	ILE	6.1
1	A	177	ILE	6.0
2	B	63	GLU	6.0
1	A	381	GLU	5.8
1	A	395	GLU	5.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	366	ALA	5.7
1	A	394	ALA	5.6
1	A	24	GLU	5.6
1	A	76	GLU	5.6
1	A	149	TYR	5.5
1	A	300	GLU	5.5
1	A	34	GLU	5.5
1	A	400	THR	5.4
1	A	72	THR	5.4
1	A	99	TYR	5.4
1	A	200	GLU	5.3
1	A	398	ARG	5.3
1	A	33	ARG	5.3
1	A	27	ILE	5.2
1	A	47	SER	5.2
1	A	434	VAL	5.1
1	A	70	SER	5.1
1	A	412	LYS	5.0
1	A	444	LEU	5.0
1	A	26	GLU	5.0
1	A	104	GLU	4.9
1	A	407	THR	4.9
1	A	43	LYS	4.8
1	A	6	LYS	4.8
1	A	74	GLU	4.7
1	A	417	ASP	4.6
1	A	174	GLU	4.6
1	A	198	SER	4.5
1	A	120	SER	4.4
1	A	12	LEU	4.4
1	A	97	SER	4.4
1	A	121	GLY	4.4
1	A	322	THR	4.3
1	A	125	ASP	4.3
1	A	69	ARG	4.3
2	B	64	GLN	4.3
1	A	286	SER	4.3
1	A	179	GLU	4.3
1	A	202	ILE	4.3
2	B	13	LEU	4.3
1	A	367	LEU	4.3
2	B	65	LYS	4.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	414	PRO	4.3
1	A	196	GLY	4.3
2	B	56	GLU	4.2
1	A	364	ARG	4.2
1	A	25	ASP	4.1
1	A	29	GLU	4.1
1	A	199	ALA	4.1
1	A	150	GLU	4.0
1	A	158	ARG	4.0
1	A	95	SER	4.0
1	A	346	GLY	4.0
1	A	321	GLY	4.0
1	A	409	GLU	3.9
1	A	270	SER	3.9
1	A	415	SER	3.9
2	B	15	GLU	3.9
1	A	323	SER	3.8
1	A	447	ILE	3.7
1	A	42	LEU	3.7
1	A	127	ILE	3.7
1	A	357	ARG	3.7
1	A	142	LEU	3.7
1	A	244	GLU	3.7
1	A	49	TYR	3.7
1	A	418	ALA	3.7
1	A	130	ILE	3.6
1	A	368	GLN	3.6
1	A	354	GLU	3.6
1	A	131	VAL	3.6
1	A	192	LEU	3.6
1	A	154	GLU	3.6
1	A	1	ASP	3.5
1	A	440	ILE	3.5
1	A	16	GLU	3.5
1	A	20	SER	3.5
1	A	8	LEU	3.5
1	A	148	SER	3.5
1	A	349	PHE	3.4
2	B	14	ARG	3.4
1	A	197	SER	3.4
1	A	9	ARG	3.4
1	A	431	ARG	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	36	SER	3.3
1	A	370	GLY	3.3
1	A	173	SER	3.3
1	A	68	LYS	3.2
1	A	402	SER	3.2
1	A	292	LEU	3.2
2	B	18	GLU	3.2
1	A	242	LEU	3.2
1	A	304	LEU	3.1
1	A	2	GLU	3.1
1	A	355	ILE	3.1
1	A	449	ARG	3.1
1	A	405	ALA	3.1
2	B	59	GLU	3.1
1	A	123	SER	3.0
1	A	100	GLU	3.0
1	A	60	VAL	3.0
1	A	73	SER	3.0
1	A	416	GLU	3.0
1	A	38	VAL	3.0
1	A	314	VAL	3.0
1	A	194	GLU	2.9
1	A	54	GLU	2.9
1	A	13	LYS	2.8
1	A	4	GLU	2.8
1	A	37	GLU	2.8
1	A	221	GLY	2.8
1	A	28	ALA	2.8
1	A	324	GLU	2.8
1	A	448	VAL	2.8
1	A	410	LEU	2.8
1	A	50	GLU	2.8
1	A	445	GLU	2.8
1	A	147	SER	2.7
1	A	293	LYS	2.7
1	A	339	LEU	2.7
1	A	438	ILE	2.7
1	A	133	ARG	2.7
1	A	5	ALA	2.6
1	A	360	LEU	2.6
1	A	393	THR	2.6
1	A	86	SER	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	312	GLU	2.5
1	A	318	LEU	2.5
1	A	102	ILE	2.5
1	A	204	GLU	2.5
1	A	178	ALA	2.5
2	B	40	ASP	2.5
1	A	351	VAL	2.5
1	A	406	ILE	2.5
1	A	108	ARG	2.5
1	A	308	MET	2.5
1	A	257	ALA	2.5
1	A	219	ARG	2.4
1	A	189	ILE	2.4
1	A	195	SER	2.4
2	B	44	ARG	2.4
1	A	214	VAL	2.4
1	A	223	SER	2.4
1	A	205	ILE	2.4
1	A	413	ASN	2.4
1	A	218	LYS	2.4
1	A	40	ARG	2.4
1	A	271	GLY	2.3
1	A	136	SER	2.3
1	A	423	ASN	2.3
1	A	155	ILE	2.3
1	A	137	GLU	2.3
1	A	435	PHE	2.3
1	A	129	GLU	2.3
2	B	16	GLN	2.3
2	B	52	GLU	2.3
1	A	359	ILE	2.3
1	A	369	LYS	2.2
1	A	56	VAL	2.2
1	A	443	VAL	2.2
1	A	424	ARG	2.2
1	A	77	ILE	2.2
1	A	146	GLY	2.2
2	B	24	MET	2.2
1	A	217	LEU	2.2
1	A	425	ALA	2.2
1	A	172	THR	2.2
1	A	343	ARG	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	58	GLU	2.2
1	A	365	ALA	2.2
1	A	53	ALA	2.1
1	A	139	ILE	2.1
1	A	23	SER	2.1
1	A	75	ASP	2.1
1	A	358	LEU	2.1
1	A	387	MET	2.1
1	A	66	ALA	2.1
1	A	348	ASP	2.1
1	A	22	THR	2.1
1	A	255	ILE	2.1
1	A	140	ARG	2.1
1	A	420	VAL	2.1
1	A	451	ILE	2.0
1	A	201	VAL	2.0
1	A	212	GLU	2.0
1	A	350	GLU	2.0
1	A	347	SER	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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.