



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 11:55 am BST

PDB ID : 1W62  
Title : proline racemase in complex with one molecule of pyrrole-2-carboxylic acid (hemi form)  
Authors : Buschiazzo, A.; Alzari, P.  
Deposited on : 2004-08-12  
Resolution : 2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 following versions of software and data (see [references ⓘ](#)) 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	:	2.11
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)	:	2.11

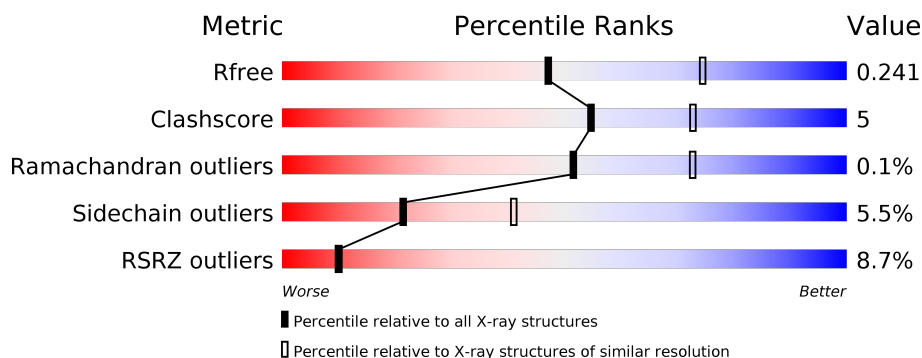
# 1 Overall quality at a glance ⓘ

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))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	414	<div> <div>11%</div> <div> <div></div> <div>73%</div> <div>13%</div> <div>•</div> <div>13%</div> </div> </div>
1	B	414	<div> <div>4%</div> <div> <div></div> <div>70%</div> <div>13%</div> <div>•</div> <div>15%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5554 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 Proline racemase A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	0	0
			2739	1736	465	522	16			
1	B	351	Total	C	N	O	S	0	0	0
			2656	1684	450	509	13			

There are 46 discrepancies between the modelled and reference sequences:

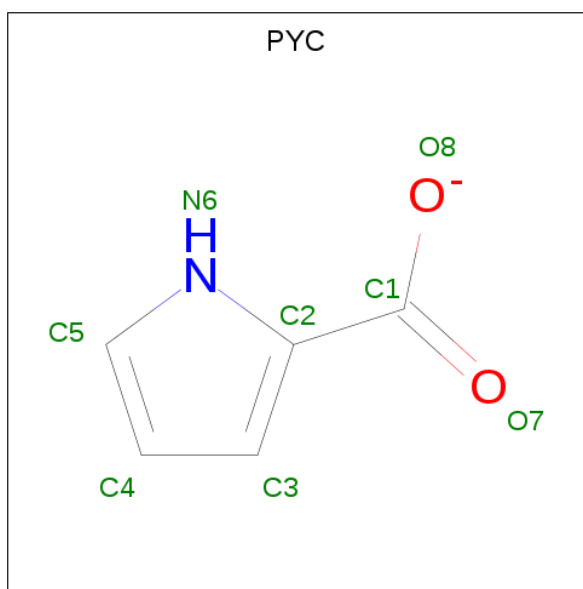
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q4DA80
A	118	ILE	MET	conflict	UNP Q4DA80
A	394	TYR	-	expression tag	UNP Q4DA80
A	395	ILE	-	expression tag	UNP Q4DA80
A	396	TRP	-	expression tag	UNP Q4DA80
A	397	SER	-	expression tag	UNP Q4DA80
A	398	SER	-	expression tag	UNP Q4DA80
A	399	SER	-	expression tag	UNP Q4DA80
A	400	VAL	-	expression tag	UNP Q4DA80
A	401	ASP	-	expression tag	UNP Q4DA80
A	402	LYS	-	expression tag	UNP Q4DA80
A	403	LEU	-	expression tag	UNP Q4DA80
A	404	ALA	-	expression tag	UNP Q4DA80
A	405	ALA	-	expression tag	UNP Q4DA80
A	406	ALA	-	expression tag	UNP Q4DA80
A	407	LEU	-	expression tag	UNP Q4DA80
A	408	GLU	-	expression tag	UNP Q4DA80
A	409	HIS	-	expression tag	UNP Q4DA80
A	410	HIS	-	expression tag	UNP Q4DA80
A	411	HIS	-	expression tag	UNP Q4DA80
A	412	HIS	-	expression tag	UNP Q4DA80
A	413	HIS	-	expression tag	UNP Q4DA80
A	414	HIS	-	expression tag	UNP Q4DA80
B	1	MET	-	initiating methionine	UNP Q4DA80
B	118	ILE	MET	conflict	UNP Q4DA80

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	394	TYR	-	expression tag	UNP Q4DA80
B	395	ILE	-	expression tag	UNP Q4DA80
B	396	TRP	-	expression tag	UNP Q4DA80
B	397	SER	-	expression tag	UNP Q4DA80
B	398	SER	-	expression tag	UNP Q4DA80
B	399	SER	-	expression tag	UNP Q4DA80
B	400	VAL	-	expression tag	UNP Q4DA80
B	401	ASP	-	expression tag	UNP Q4DA80
B	402	LYS	-	expression tag	UNP Q4DA80
B	403	LEU	-	expression tag	UNP Q4DA80
B	404	ALA	-	expression tag	UNP Q4DA80
B	405	ALA	-	expression tag	UNP Q4DA80
B	406	ALA	-	expression tag	UNP Q4DA80
B	407	LEU	-	expression tag	UNP Q4DA80
B	408	GLU	-	expression tag	UNP Q4DA80
B	409	HIS	-	expression tag	UNP Q4DA80
B	410	HIS	-	expression tag	UNP Q4DA80
B	411	HIS	-	expression tag	UNP Q4DA80
B	412	HIS	-	expression tag	UNP Q4DA80
B	413	HIS	-	expression tag	UNP Q4DA80
B	414	HIS	-	expression tag	UNP Q4DA80

- Molecule 2 is PYRROLE-2-CARBOXYLATE (three-letter code: PYC) (formula:  $C_5H_4NO_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			8	5	1	2		

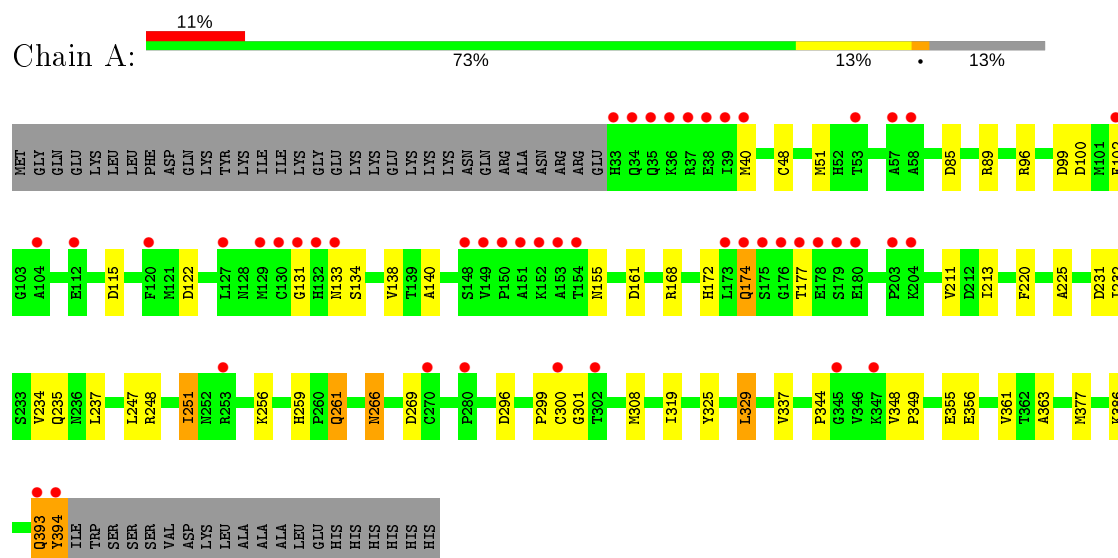
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	73	Total 73	O 73	0	0
3	B	78	Total 78	O 78	0	0

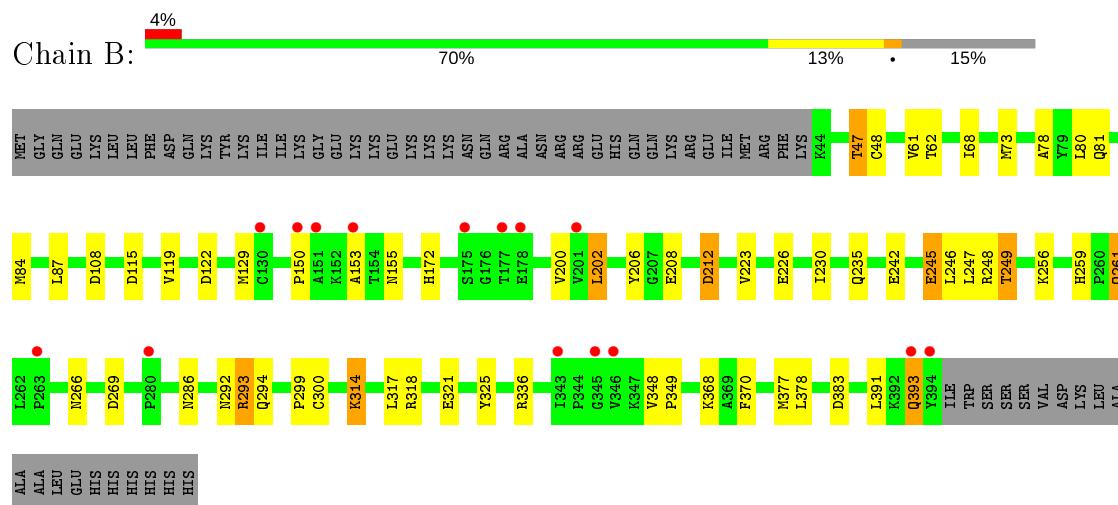
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Proline racemase A



#### • Molecule 1: Proline racemase A



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.02Å 89.38Å 84.46Å 90.00° 125.98° 90.00°	Depositor
Resolution (Å)	30.00 – 2.50 28.76 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.3 (30.00-2.50) 97.2 (28.76-2.50)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.46 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, $R_{free}$	0.170 , 0.236 0.179 , 0.241	Depositor DCC
$R_{free}$ test set	2778 reflections (10.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.5	Xtriage
Anisotropy	0.410	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 48.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.032 for -h-2*k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5554	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PYC

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	0.80	0/2794	0.95	12/3791 (0.3%)
1	B	0.88	0/2710	0.95	7/3682 (0.2%)
All	All	0.84	0/5504	0.95	19/7473 (0.3%)

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	1

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	122	ASP	CB-CG-OD2	9.87	127.18	118.30
1	B	108	ASP	CB-CG-OD2	8.36	125.83	118.30
1	A	115	ASP	CB-CG-OD2	8.04	125.54	118.30
1	A	161	ASP	CB-CG-OD2	7.38	124.94	118.30
1	A	122	ASP	CB-CG-OD2	6.81	124.43	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	174	GLN	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	A	2739	0	2706	29	0
1	B	2656	0	2627	36	0
2	A	8	0	4	0	0
3	A	73	0	0	0	0
3	B	78	0	0	0	0
All	All	5554	0	5337	56	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 5.

The worst 5 of 56 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:LEU:CD1	1:B:391:LEU:HD21	2.15	0.77
1:A:394:TYR:N	1:B:235:GLN:HE22	1.92	0.65
1:B:48:CYS:HB3	1:B:61:VAL:HA	1.78	0.65
1:A:155:ASN:ND2	1:A:172:HIS:ND1	2.48	0.62
1:B:84:MET:HE1	1:B:87:LEU:HD11	1.82	0.61

There are no symmetry-related clashes.

## 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	360/414 (87%)	344 (96%)	16 (4%)	0	100 100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	349/414 (84%)	337 (97%)	11 (3%)	1 (0%)	41	61
All	All	709/828 (86%)	681 (96%)	27 (4%)	1 (0%)	51	73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	129	MET

### 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	294/347 (85%)	277 (94%)	17 (6%)	20	38
1	B	286/347 (82%)	271 (95%)	15 (5%)	23	44
All	All	580/694 (84%)	548 (94%)	32 (6%)	21	41

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	355	GLU
1	B	47	THR
1	B	325	TYR
1	A	394	TYR
1	B	202	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	195	GLN
1	B	236	ASN
1	B	294	GLN
1	B	155	ASN
1	B	316	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is 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	PYC	A	700	-	5,8,8	0.57	0	4,10,10	1.47	0

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	PYC	A	700	-	-	0/0/4/4	0/1/1/1

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	362/414 (87%)	0.54	47 (12%) <b>3</b> <b>3</b>	26, 42, 55, 71	0
1	B	351/414 (84%)	0.16	15 (4%) 35 38	27, 38, 50, 64	0
All	All	713/828 (86%)	0.35	62 (8%) <b>10</b> <b>10</b>	26, 40, 53, 71	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	151	ALA	6.2
1	A	394	TYR	6.2
1	A	178	GLU	5.6
1	A	175	SER	5.6
1	A	204	LYS	4.7

### 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 carbohydrates 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, 95<sup>th</sup> 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
2	PYC	A	700	8/8	0.97	0.22	42,43,44,44	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.