



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 09:04 am BST

PDB ID : 1MF2  
Title : ANTI HIV1 PROTEASE FAB COMPLEX  
Authors : Lescar, J.; Bentley, G.A.  
Deposited on : 1996-12-27  
Resolution : 2.60 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  
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

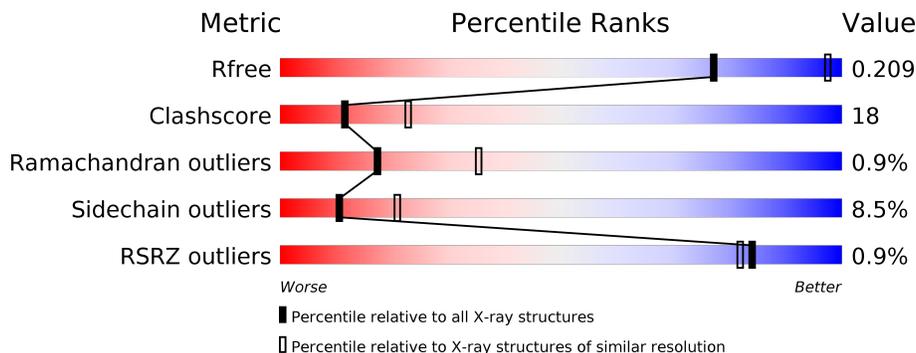
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	L	215	 68% 29%
1	M	215	 68% 29%
2	H	226	 2% 61% 31% 5%
2	N	226	 1% 62% 29% 7%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6857 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 MONOCLONAL ANTIBODY F11.2.32.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	215	1647	1024	281	334	8	0	0	0
1	M	215	1647	1024	281	334	8	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	4	LEU	MET	CONFLICT	GB 600718
L	18	ARG	SER	CONFLICT	GB 600718
L	19	ALA	VAL	CONFLICT	GB 600718
L	27C	ASP	GLU	CONFLICT	GB 600718
L	30	LYS	THR	CONFLICT	GB 600718
L	32	PHE	LEU	CONFLICT	GB 600718
L	34	ASN	GLN	CONFLICT	GB 600718
L	36	PHE	TYR	CONFLICT	GB 600718
L	50	ALA	GLY	CONFLICT	GB 600718
L	54	GLN	VAL	CONFLICT	GB 600718
L	55	GLY	GLU	CONFLICT	GB 600718
L	74	HIS	ASN	CONFLICT	GB 600718
L	78	MET	VAL	CONFLICT	GB 600718
L	83	SER	ILE	CONFLICT	GB 600718
L	85	MET	ILE	CONFLICT	GB 600718
L	92	LYS	ARG	CONFLICT	GB 600718
L	93	GLU	LYS	CONFLICT	GB 600718
L	96	TRP	ALA	CONFLICT	GB 600718
L	100	GLY	SER	CONFLICT	GB 600718
M	4	LEU	MET	CONFLICT	GB 600718
M	18	ARG	SER	CONFLICT	GB 600718
M	19	ALA	VAL	CONFLICT	GB 600718
M	27C	ASP	GLU	CONFLICT	GB 600718
M	30	LYS	THR	CONFLICT	GB 600718
M	32	PHE	LEU	CONFLICT	GB 600718

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
M	34	ASN	GLN	CONFLICT	GB 600718
M	36	PHE	TYR	CONFLICT	GB 600718
M	50	ALA	GLY	CONFLICT	GB 600718
M	54	GLN	VAL	CONFLICT	GB 600718
M	55	GLY	GLU	CONFLICT	GB 600718
M	74	HIS	ASN	CONFLICT	GB 600718
M	78	MET	VAL	CONFLICT	GB 600718
M	83	SER	ILE	CONFLICT	GB 600718
M	85	MET	ILE	CONFLICT	GB 600718
M	92	LYS	ARG	CONFLICT	GB 600718
M	93	GLU	LYS	CONFLICT	GB 600718
M	96	TRP	ALA	CONFLICT	GB 600718
M	100	GLY	SER	CONFLICT	GB 600718

- Molecule 2 is a protein called MONOCLONAL ANTIBODY F11.2.32.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	221	1664	1051	281	323	9	0	0	0
2	N	221	1664	1051	281	323	9	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	5	VAL	LEU	CONFLICT	GB 600716
H	13	GLN	LYS	CONFLICT	GB 600716
H	18	ARG	LEU	CONFLICT	GB 600716
H	30	MET	SER	CONFLICT	GB 600716
H	31	ARG	ASP	CONFLICT	GB 600716
H	32	PHE	TYR	CONFLICT	GB 600716
H	74	PRO	ALA	CONFLICT	GB 600716
H	89	LEU	MET	CONFLICT	GB 600716
H	95	SER	-	INSERTION	GB 600716
H	96	GLY	-	INSERTION	GB 600716
H	97	GLY	TRP	CONFLICT	GB 600716
H	98	ILE	ASP	CONFLICT	GB 600716
H	99	GLU	THR	CONFLICT	GB 600716
H	100	ARG	THR	CONFLICT	GB 600716
H	100A	TYR	VAL	CONFLICT	GB 600716
H	100B	ASP	SER	CONFLICT	GB 600716
H	100D	THR	HIS	CONFLICT	GB 600716

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	182	PRO	THR	CONFLICT	GB 600716
H	183	ARG	TRP	CONFLICT	GB 600716
H	186	GLU	GLN	CONFLICT	GB 600716
N	5	VAL	LEU	CONFLICT	GB 600716
N	13	GLN	LYS	CONFLICT	GB 600716
N	18	ARG	LEU	CONFLICT	GB 600716
N	30	MET	SER	CONFLICT	GB 600716
N	31	ARG	ASP	CONFLICT	GB 600716
N	32	PHE	TYR	CONFLICT	GB 600716
N	74	PRO	ALA	CONFLICT	GB 600716
N	89	LEU	MET	CONFLICT	GB 600716
N	95	SER	-	INSERTION	GB 600716
N	96	GLY	-	INSERTION	GB 600716
N	97	GLY	TRP	CONFLICT	GB 600716
N	98	ILE	ASP	CONFLICT	GB 600716
N	99	GLU	THR	CONFLICT	GB 600716
N	100	ARG	THR	CONFLICT	GB 600716
N	100A	TYR	VAL	CONFLICT	GB 600716
N	100B	ASP	SER	CONFLICT	GB 600716
N	100D	THR	HIS	CONFLICT	GB 600716
N	182	PRO	THR	CONFLICT	GB 600716
N	183	ARG	TRP	CONFLICT	GB 600716
N	186	GLU	GLN	CONFLICT	GB 600716

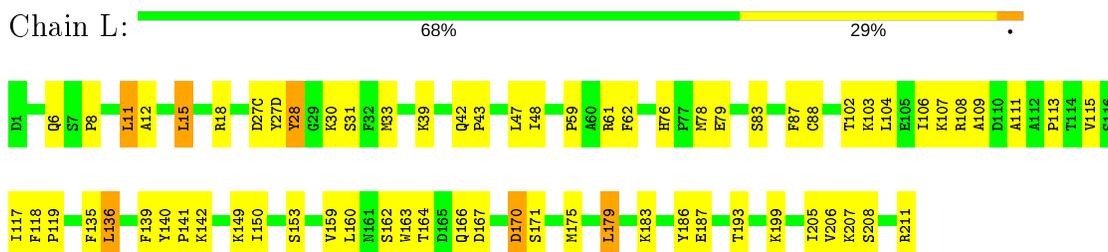
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	L	48	Total O 48 48	0	0
3	H	78	Total O 78 78	0	0
3	M	45	Total O 45 45	0	0
3	N	64	Total O 64 64	0	0

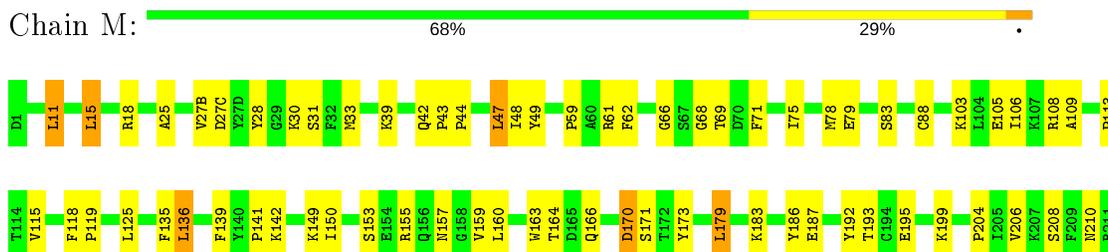
### 3 Residue-property plots [i](#)

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.

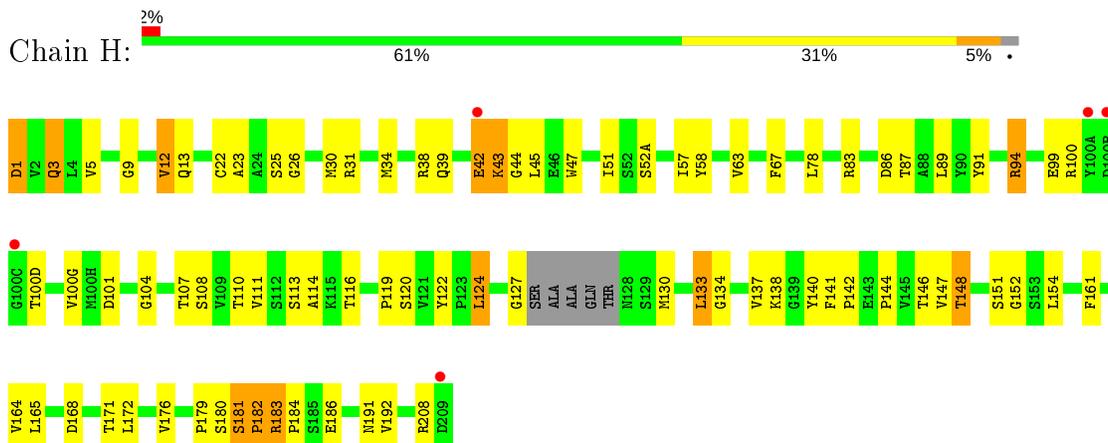
- Molecule 1: MONOCLONAL ANTIBODY F11.2.32



- Molecule 1: MONOCLONAL ANTIBODY F11.2.32

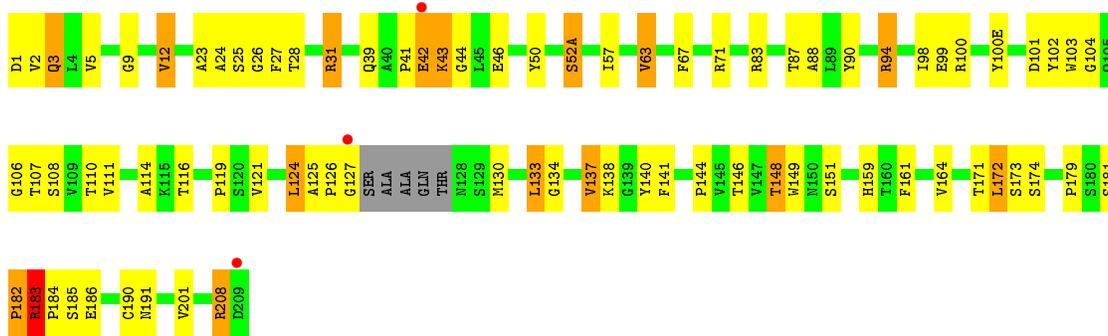


- Molecule 2: MONOCLONAL ANTIBODY F11.2.32



- Molecule 2: MONOCLONAL ANTIBODY F11.2.32





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.60 Å 94.70 Å 70.30 Å 90.00° 105.80° 90.00°	Depositor
Resolution (Å)	8.00 – 2.60 7.99 – 2.60	Depositor EDS
% Data completeness (in resolution range)	96.0 (8.00-2.60) 96.7 (7.99-2.60)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.74 (at 2.58 Å)	Xtrriage
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.206 , 0.307 0.200 , 0.209	Depositor DCC
$R_{free}$ test set	1194 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.3	Xtrriage
Anisotropy	0.040	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 65.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.144 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6857	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.35% 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](#)

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	L	0.50	0/1686	0.76	0/2285
1	M	0.51	0/1686	0.77	0/2285
2	H	0.59	0/1706	0.83	1/2323 (0.0%)
2	N	0.55	0/1706	0.82	1/2323 (0.0%)
All	All	0.54	0/6784	0.80	2/9216 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	N	183	ARG	NE-CZ-NH1	6.27	123.44	120.30
2	H	42	GLU	N-CA-CB	-5.03	101.55	110.60

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	L	1647	0	1567	57	0
1	M	1647	0	1567	58	0
2	H	1664	0	1625	67	0
2	N	1664	0	1625	72	0
3	H	78	0	0	4	0
3	L	48	0	0	1	0
3	M	45	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	N	64	0	0	1	0
All	All	6857	0	6384	229	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 18.

The worst 5 of 229 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:N:9:GLY:H	2:N:107:THR:HG21	1.29	0.97
2:H:183:ARG:HD2	2:H:184:PRO:HA	1.46	0.96
2:N:183:ARG:HD2	2:N:184:PRO:HA	1.49	0.94
2:H:9:GLY:H	2:H:107:THR:HG21	1.32	0.93
2:H:183:ARG:HB3	2:H:183:ARG:HH11	1.32	0.92

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	L	213/215 (99%)	204 (96%)	7 (3%)	2 (1%)	17	35
1	M	213/215 (99%)	202 (95%)	9 (4%)	2 (1%)	17	35
2	H	217/226 (96%)	203 (94%)	12 (6%)	2 (1%)	17	35
2	N	217/226 (96%)	203 (94%)	12 (6%)	2 (1%)	17	35
All	All	860/882 (98%)	812 (94%)	40 (5%)	8 (1%)	17	35

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	12	ALA

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	L	28	TYR
2	H	43	LYS
1	M	28	TYR
2	N	43	LYS

### 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	L	185/187 (99%)	176 (95%)	9 (5%)	25	48
1	M	185/187 (99%)	175 (95%)	10 (5%)	22	44
2	H	186/192 (97%)	162 (87%)	24 (13%)	4	7
2	N	186/192 (97%)	166 (89%)	20 (11%)	6	12
All	All	742/758 (98%)	679 (92%)	63 (8%)	10	21

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	180	SER
1	M	47	LEU
2	N	172	LEU
2	H	181	SER
2	H	183	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	13	GLN
1	M	53	ASN
1	M	137	ASN
2	H	3	GLN
1	M	124	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](#)

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 [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	L	215/215 (100%)	-0.46	0 <a href="#">100</a> <a href="#">100</a>	3, 16, 36, 50	0
1	M	215/215 (100%)	-0.49	0 <a href="#">100</a> <a href="#">100</a>	3, 14, 37, 53	0
2	H	221/226 (97%)	-0.53	5 (2%) <a href="#">60</a> <a href="#">54</a>	3, 12, 34, 70	0
2	N	221/226 (97%)	-0.50	3 (1%) <a href="#">75</a> <a href="#">71</a>	3, 14, 35, 67	0
All	All	872/882 (98%)	-0.50	8 (0%) <a href="#">84</a> <a href="#">82</a>	3, 14, 37, 70	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	N	209	ASP	4.7
2	H	209	ASP	4.5
2	H	100(A)	TYR	3.2
2	H	42	GLU	3.1
2	H	100(B)	ASP	2.9

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.