



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2017 – 02:18 am GMT

PDB ID : 4APA  
Title : Crystal structure of Mycobacterium tuberculosis fumarase (Rv1098c) S318A in apo form  
Authors : Bellinzoni, M.; Haouz, A.; Mechaly, A.E.; Alzari, P.M.  
Deposited on : 2012-03-31  
Resolution : 2.04 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

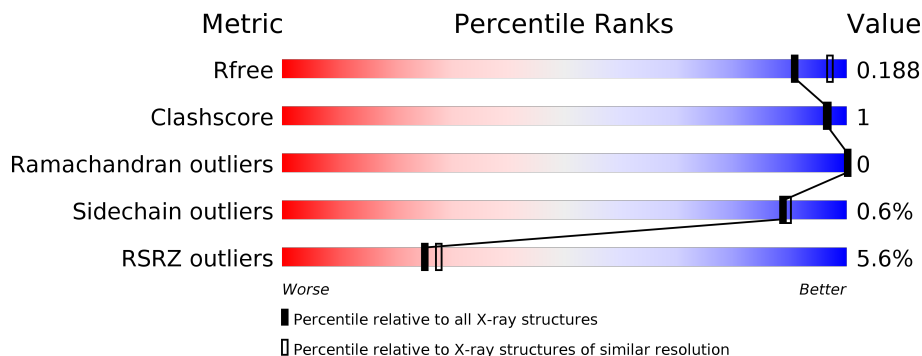
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.04 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))
$R_{free}$	100719	1316 (2.04-2.04)
Clashscore	112137	1394 (2.04-2.04)
Ramachandran outliers	110173	1383 (2.04-2.04)
Sidechain outliers	110143	1383 (2.04-2.04)
RSRZ outliers	101464	1319 (2.04-2.04)

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. 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	474	<div> <div>7%</div> <div> <div></div> <div>91%</div> <div>7%</div> </div> </div>
1	B	474	<div> <div>8%</div> <div> <div></div> <div>91%</div> <div>5%</div> </div> </div>
1	C	474	<div> <div>2%</div> <div> <div></div> <div>91%</div> <div>6%</div> </div> </div>
1	D	474	<div> <div>5%</div> <div> <div></div> <div>91%</div> <div>6%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13702 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 FUMARATE HYDRATASE CLASS II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	440	Total	C	N	O	S	0	1	0
			3175	1976	573	616	10			
1	B	449	Total	C	N	O	S	0	0	0
			3278	2046	589	632	11			
1	C	445	Total	C	N	O	S	0	0	0
			3244	2025	587	622	10			
1	D	444	Total	C	N	O	S	0	0	0
			3207	2003	575	619	10			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP O53446
A	318	ALA	SER	ENGINEERED MUTATION	UNP O53446
B	1	GLY	-	EXPRESSION TAG	UNP O53446
B	318	ALA	SER	ENGINEERED MUTATION	UNP O53446
C	1	GLY	-	EXPRESSION TAG	UNP O53446
C	318	ALA	SER	ENGINEERED MUTATION	UNP O53446
D	1	GLY	-	EXPRESSION TAG	UNP O53446
D	318	ALA	SER	ENGINEERED MUTATION	UNP O53446

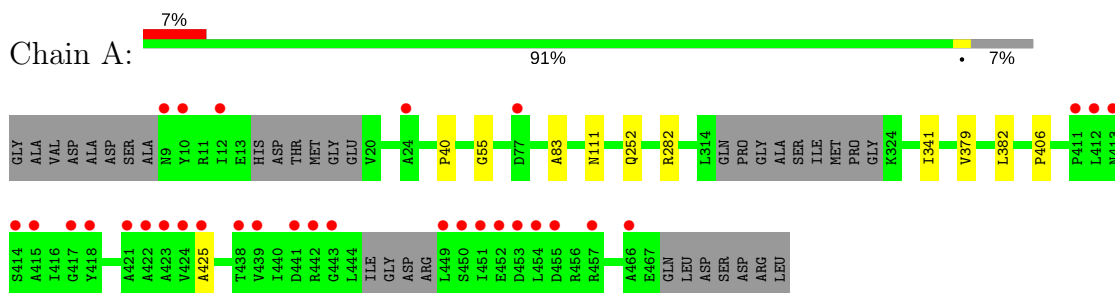
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	212	Total	O	0	0
			212	212		
2	B	156	Total	O	0	0
			156	156		
2	C	227	Total	O	0	0
			227	227		
2	D	203	Total	O	0	0
			203	203		

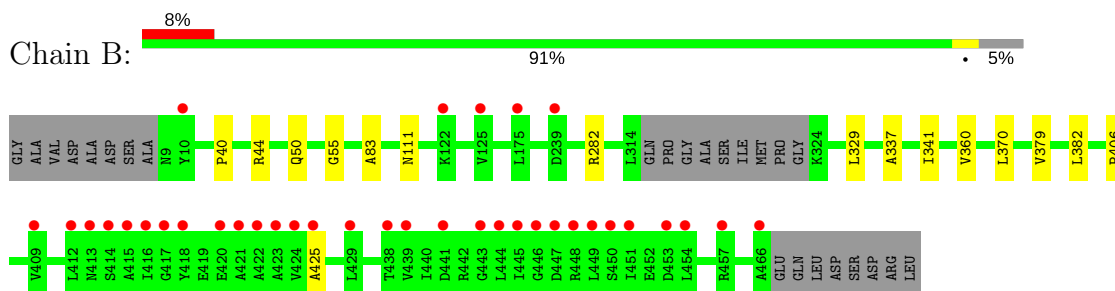
### 3 Residue-property plots [i](#)

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.

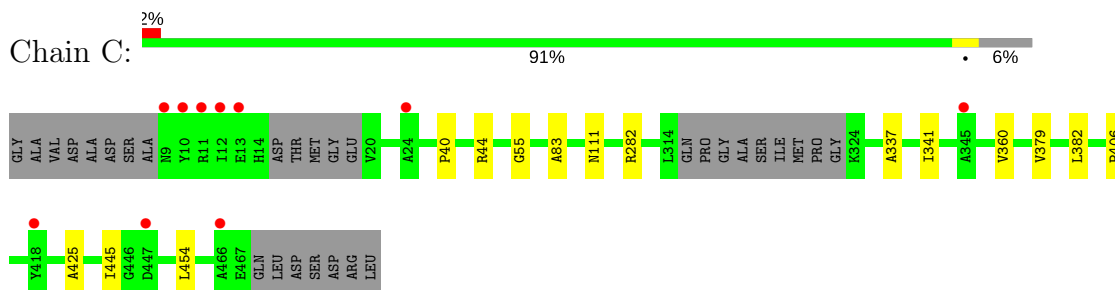
#### • Molecule 1: FUMARATE HYDRATASE CLASS II



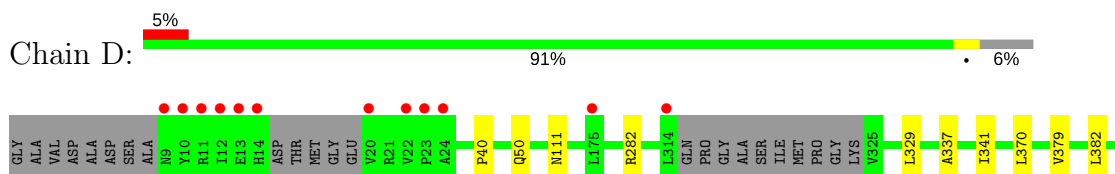
#### • Molecule 1: FUMARATE HYDRATASE CLASS II

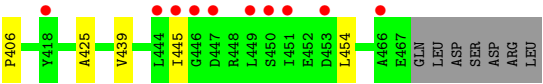


#### • Molecule 1: FUMARATE HYDRATASE CLASS II



#### • Molecule 1: FUMARATE HYDRATASE CLASS II





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	268.27Å 97.14Å 96.44Å 90.00° 104.54° 90.00°	Depositor
Resolution (Å)	43.41 – 2.04 43.41 – 2.04	Depositor EDS
% Data completeness (in resolution range)	97.2 (43.41-2.04) 97.8 (43.41-2.04)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.10 (at 2.05Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, $R_{free}$	0.165 , 0.179 0.171 , 0.188	Depositor DCC
$R_{free}$ test set	7424 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.9	Xtriage
Anisotropy	0.532	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 45.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13702	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.96% 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	A	0.48	0/3219	0.56	0/4389
1	B	0.48	0/3324	0.57	0/4529
1	C	0.49	0/3287	0.56	0/4475
1	D	0.50	0/3250	0.56	0/4431
All	All	0.49	0/13080	0.56	0/17824

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 [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	A	3175	0	3117	6	0
1	B	3278	0	3265	11	0
1	C	3244	0	3242	10	0
1	D	3207	0	3174	9	0
2	A	212	0	0	1	0
2	B	156	0	0	1	0
2	C	227	0	0	2	0
2	D	203	0	0	0	0
All	All	13702	0	12798	31	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 1.

All (31) 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:B:40:PRO:HD2	1:C:382:LEU:HD11	1.86	0.57
1:C:445:ILE:HD11	1:C:454:LEU:HD22	1.88	0.56
1:B:50:GLN:HG2	1:B:370:LEU:CD1	2.36	0.55
1:D:445:ILE:HD11	1:D:454:LEU:HD22	1.89	0.55
1:B:382:LEU:HD11	1:C:40:PRO:HD2	1.88	0.54
1:B:44:ARG:HG3	1:C:44:ARG:HH22	1.73	0.53
1:C:282:ARG:HG2	1:C:341:ILE:HD13	1.92	0.51
1:A:40:PRO:HD2	1:D:382:LEU:HD11	1.93	0.49
1:B:282:ARG:HG2	1:B:341:ILE:HD13	1.94	0.49
1:A:282:ARG:HG2	1:A:341:ILE:HD13	1.95	0.48
1:D:282:ARG:HG2	1:D:341:ILE:HD13	1.93	0.48
1:B:50:GLN:HG2	1:B:370:LEU:HD13	1.95	0.48
1:A:382:LEU:HD11	1:D:40:PRO:HD2	1.98	0.46
1:B:360:VAL:HG22	2:B:2138:HOH:O	2.17	0.44
1:C:44:ARG:HG2	2:C:2188:HOH:O	2.18	0.43
1:D:406:PRO:HB3	1:D:425:ALA:HB1	1.99	0.43
1:A:406:PRO:HB3	1:A:425:ALA:HB1	2.01	0.43
1:B:55:GLY:O	1:B:83:ALA:HB1	2.18	0.42
1:C:360:VAL:HG22	2:C:2180:HOH:O	2.19	0.42
1:A:252[B]:GLN:NE2	2:A:2030:HOH:O	2.52	0.42
1:C:337:ALA:O	1:C:341:ILE:HG12	2.20	0.42
1:C:406:PRO:HB3	1:C:425:ALA:HB1	2.02	0.42
1:C:55:GLY:O	1:C:83:ALA:HB1	2.20	0.41
1:B:406:PRO:HB3	1:B:425:ALA:HB1	2.02	0.41
1:A:55:GLY:O	1:A:83:ALA:HB1	2.20	0.41
1:D:337:ALA:O	1:D:341:ILE:HG12	2.21	0.41
1:D:50:GLN:HG2	1:D:370:LEU:CD1	2.50	0.41
1:B:329:LEU:HD23	1:B:329:LEU:HA	1.93	0.40
1:D:329:LEU:HA	1:D:329:LEU:HD23	1.97	0.40
1:B:337:ALA:O	1:B:341:ILE:HG12	2.21	0.40
1:D:439:VAL:HG11	1:D:454:LEU:HD13	2.03	0.40

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	433/474 (91%)	423 (98%)	10 (2%)	0	100	100
1	B	445/474 (94%)	435 (98%)	10 (2%)	0	100	100
1	C	439/474 (93%)	429 (98%)	10 (2%)	0	100	100
1	D	438/474 (92%)	428 (98%)	10 (2%)	0	100	100
All	All	1755/1896 (93%)	1715 (98%)	40 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	313/362 (86%)	311 (99%)	2 (1%)	89	89
1	B	330/362 (91%)	328 (99%)	2 (1%)	89	89
1	C	325/362 (90%)	323 (99%)	2 (1%)	89	89
1	D	318/362 (88%)	316 (99%)	2 (1%)	89	89
All	All	1286/1448 (89%)	1278 (99%)	8 (1%)	89	89

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	111	ASN
1	A	379	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	111	ASN
1	B	379	VAL
1	C	111	ASN
1	C	379	VAL
1	D	111	ASN
1	D	379	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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

### 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, 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	440/474 (92%)	0.09	31 (7%) 17 18	19, 31, 77, 93	0
1	B	449/474 (94%)	0.17	36 (8%) 13 14	20, 33, 69, 106	0
1	C	445/474 (93%)	-0.24	10 (2%) 62 67	18, 27, 52, 110	0
1	D	444/474 (93%)	-0.20	22 (4%) 30 32	19, 27, 59, 101	0
All	All	1778/1896 (93%)	-0.05	99 (5%) 25 28	18, 30, 67, 110	0

All (99) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	414	SER	6.3
1	B	418	TYR	6.0
1	D	12	ILE	5.9
1	C	12	ILE	5.8
1	A	418	TYR	5.0
1	A	421	ALA	5.0
1	B	446	GLY	4.9
1	B	414	SER	4.8
1	A	12	ILE	4.8
1	B	449	LEU	4.8
1	B	444	LEU	4.7
1	B	416	ILE	4.6
1	D	10	TYR	4.6
1	B	447	ASP	4.5
1	A	417	GLY	4.3
1	B	451	ILE	4.1
1	D	24	ALA	4.1
1	C	447	ASP	4.1
1	C	24	ALA	4.0
1	B	443	GLY	3.8
1	D	418	TYR	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	9	ASN	3.6
1	B	415	ALA	3.6
1	A	413	ASN	3.6
1	A	449	LEU	3.5
1	B	422	ALA	3.5
1	A	424	VAL	3.5
1	B	448	ARG	3.4
1	B	412	LEU	3.4
1	B	438	THR	3.4
1	D	466	ALA	3.4
1	A	454	LEU	3.4
1	A	439	VAL	3.4
1	D	22	VAL	3.4
1	A	453	ASP	3.4
1	B	466	ALA	3.3
1	D	11	ARG	3.3
1	A	415	ALA	3.3
1	D	445	ILE	3.3
1	C	10	TYR	3.2
1	B	445	ILE	3.2
1	D	14	HIS	3.2
1	B	175	LEU	3.2
1	A	466	ALA	3.2
1	A	423	ALA	3.1
1	B	439	VAL	3.1
1	A	425	ALA	3.1
1	A	450	SER	3.0
1	B	453	ASP	3.0
1	A	452	GLU	3.0
1	A	411	PRO	3.0
1	A	441	ASP	2.9
1	A	451	ILE	2.9
1	C	418	TYR	2.9
1	D	449	LEU	2.9
1	C	466	ALA	2.9
1	C	11	ARG	2.8
1	A	10	TYR	2.8
1	B	417	GLY	2.8
1	A	422	ALA	2.8
1	B	239	ASP	2.8
1	B	441	ASP	2.8
1	D	446	GLY	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	13	GLU	2.7
1	A	412	LEU	2.7
1	B	424	VAL	2.6
1	D	20	VAL	2.6
1	D	444	LEU	2.6
1	B	10	TYR	2.6
1	B	423	ALA	2.6
1	A	438	THR	2.6
1	A	443	GLY	2.6
1	B	429	LEU	2.5
1	A	77	ASP	2.4
1	D	447	ASP	2.4
1	A	442	ARG	2.4
1	B	125	VAL	2.3
1	D	314	LEU	2.3
1	B	413	ASN	2.3
1	B	450	SER	2.2
1	D	23	PRO	2.2
1	D	453	ASP	2.2
1	B	454	LEU	2.2
1	D	450	SER	2.2
1	A	457	ARG	2.2
1	D	451	ILE	2.2
1	A	455	ASP	2.1
1	C	345	ALA	2.1
1	B	420	GLU	2.1
1	D	9	ASN	2.1
1	B	425	ALA	2.1
1	C	9	ASN	2.1
1	C	13	GLU	2.0
1	A	24	ALA	2.0
1	D	175	LEU	2.0
1	B	421	ALA	2.0
1	B	409	VAL	2.0
1	B	122	LYS	2.0
1	B	457	ARG	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 [i](#)

There are no carbohydrates 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.