



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 04:21 PM UTC

PDB ID : 5H8P / pdb_00005h8p
Title : Crystal structure of Mycobacterium tuberculosis malate synthase in apo form
Authors : Krieger, I.V.; Huang, H.-L.; Sacchettini, J.C.
Deposited on : 2015-12-23
Resolution : 2.10 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

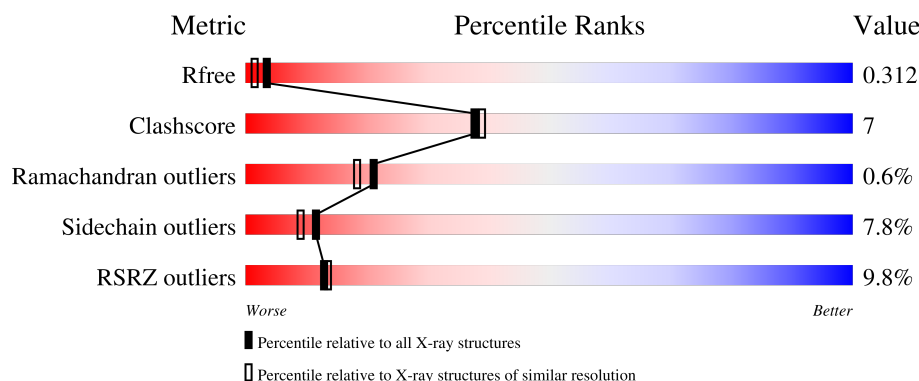
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	6658 (2.10-2.10)
Clashscore	190562	7164 (2.10-2.10)
Ramachandran outliers	187476	7099 (2.10-2.10)
Sidechain outliers	187428	7100 (2.10-2.10)
RSRZ outliers	180081	6662 (2.10-2.10)

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	741	9% 76% 16% 5%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5489 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 Malate synthase G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	701	Total	C	N	O	S	0	3	0
			5387	3387	950	1027	23			

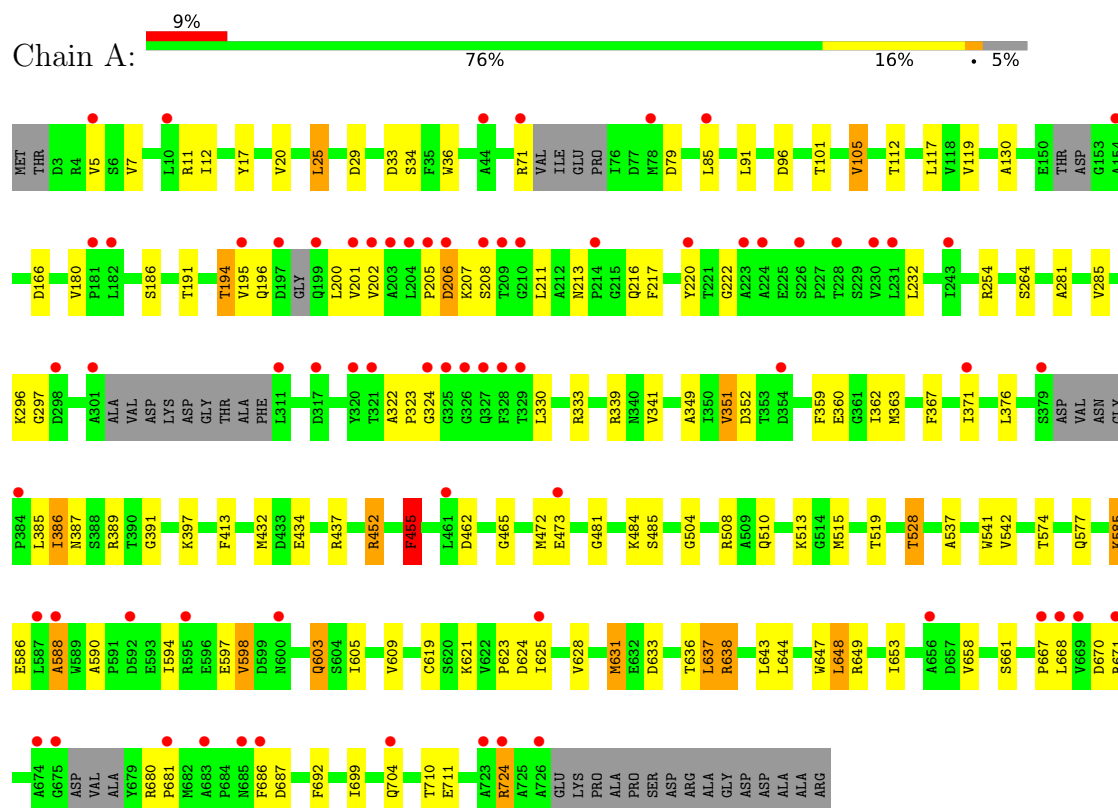
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	102	Total	O	0	0
			102	102		

3 Residue-property plots

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: Malate synthase G



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	81.79Å 81.79Å 227.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.96 – 2.10 45.96 – 2.10	Depositor EDS
% Data completeness (in resolution range)	88.2 (45.96-2.10) 88.2 (45.96-2.10)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.44 (at 2.10Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.251 , 0.312 0.252 , 0.312	Depositor DCC
R_{free} test set	2052 reflections (4.49%)	wwPDB-VP
Wilson B-factor (Å ²)	34.3	Xtriage
Anisotropy	0.433	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 26.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5489	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% 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 [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	0.54	1/5494 (0.0%)	0.91	8/7459 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	264	SER	CA-C	5.06	1.54	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	455	PHE	N-CA-C	6.39	118.26	108.42
1	A	397	LYS	CA-C-N	5.91	127.22	119.84
1	A	397	LYS	C-N-CA	5.91	127.22	119.84
1	A	590	ALA	CA-C-N	5.32	126.49	119.84
1	A	590	ALA	C-N-CA	5.32	126.49	119.84
1	A	724	ARG	N-CA-C	-5.13	107.19	113.50
1	A	213	ASN	CA-C-N	5.06	126.16	119.84
1	A	213	ASN	C-N-CA	5.06	126.16	119.84

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	5387	0	5351	75	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	102	0	0	8	0
All	All	5489	0	5351	75	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 7.

All (75) 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:297:GLY:HA3	1:A:385:LEU:HB2	1.64	0.78
1:A:515:MET:HE2	1:A:541:TRP:HB3	1.68	0.73
1:A:194:THR:HG23	1:A:201:VAL:HG13	1.71	0.70
1:A:359:PHE:HB2	1:A:362:ILE:HD12	1.76	0.68
1:A:434:GLU:OE1	2:A:801:HOH:O	2.12	0.68
1:A:619:CYS:HB2	1:A:631:MET:HE1	1.76	0.67
1:A:333:ARG:HH21	1:A:333:ARG:HB2	1.62	0.65
1:A:322:ALA:O	1:A:324:GLY:N	2.27	0.64
1:A:462:ASP:OD2	2:A:802:HOH:O	2.14	0.64
1:A:633:ASP:O	1:A:636:THR:HG22	1.99	0.62
1:A:7:VAL:HG11	1:A:36:TRP:HB3	1.84	0.59
1:A:473:GLU:OE1	1:A:649:ARG:NE	2.34	0.59
1:A:605:ILE:HD13	1:A:699:ILE:HD11	1.85	0.59
1:A:711:GLU:OE2	2:A:803:HOH:O	2.17	0.58
1:A:333:ARG:HE	1:A:385:LEU:HD13	1.69	0.57
1:A:603:GLN:NE2	1:A:623:PRO:O	2.36	0.57
1:A:339:ARG:NH1	1:A:432[B]:MET:SD	2.77	0.57
1:A:341:VAL:HG11	1:A:360:GLU:HG2	1.87	0.56
1:A:455:PHE:C	1:A:455:PHE:CD1	2.85	0.55
1:A:413:PHE:CB	1:A:452:ARG:HH11	2.20	0.55
1:A:386:ILE:HD11	1:A:389:ARG:HA	1.90	0.54
1:A:222:GLY:N	2:A:805:HOH:O	2.28	0.53
1:A:658:VAL:HG11	1:A:699:ILE:HG21	1.91	0.53
1:A:117:LEU:HD23	1:A:542:VAL:HG23	1.91	0.53
1:A:200:LEU:HB2	1:A:217:PHE:CD1	2.43	0.53
1:A:609:VAL:HG23	1:A:637:LEU:HD11	1.91	0.52
1:A:594:ILE:O	1:A:598:VAL:HG12	2.08	0.52
1:A:281:ALA:HB2	1:A:349:ALA:HA	1.91	0.52
1:A:413:PHE:HB3	1:A:452:ARG:HH11	1.75	0.51
1:A:11:ARG:HB2	1:A:351:VAL:HG12	1.93	0.50
1:A:281:ALA:O	1:A:285:VAL:HG23	2.12	0.49
1:A:481:GLY:HA2	1:A:484:LYS:HE3	1.93	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:296[B]:LYS:HA	1:A:296[B]:LYS:HD2	1.48	0.48
1:A:367:PHE:O	1:A:371:ILE:HG12	2.13	0.48
1:A:513:LYS:HD2	1:A:537:ALA:HB2	1.96	0.48
1:A:360:GLU:HA	1:A:363:MET:HE3	1.94	0.47
1:A:603:GLN:HE21	1:A:624:ASP:HA	1.80	0.47
1:A:206:ASP:OD2	1:A:206:ASP:N	2.48	0.47
1:A:504:GLY:O	1:A:508:ARG:HD2	2.15	0.46
1:A:648:LEU:HD13	1:A:653:ILE:HG13	1.97	0.46
1:A:455:PHE:C	1:A:455:PHE:HD1	2.23	0.46
1:A:623:PRO:HA	1:A:628:VAL:O	2.15	0.46
1:A:585:LYS:O	1:A:588:ALA:HB2	2.16	0.46
1:A:597:GLU:HG2	1:A:647:TRP:CH2	2.51	0.46
1:A:637:LEU:HB3	1:A:710:THR:HG21	1.99	0.45
1:A:119:VAL:HG11	1:A:130:ALA:HB2	1.98	0.45
1:A:195:VAL:HG11	1:A:220:TYR:CE2	2.51	0.45
1:A:296[B]:LYS:NZ	2:A:817:HOH:O	2.49	0.45
1:A:455:PHE:HB2	1:A:510:GLN:HB2	1.98	0.44
1:A:465:GLY:HA3	1:A:638:ARG:HG2	1.99	0.44
1:A:191:THR:HA	1:A:254:ARG:O	2.18	0.44
1:A:20:VAL:HG12	1:A:25:LEU:HD22	1.99	0.44
1:A:71:ARG:HA	1:A:71:ARG:HD3	1.86	0.44
1:A:680:ARG:HA	1:A:681:PRO:HD2	1.84	0.43
1:A:472:MET:HE2	1:A:472:MET:HB2	1.85	0.43
1:A:12:ILE:HD13	1:A:36:TRP:CZ3	2.53	0.43
1:A:105:VAL:HG21	2:A:879:HOH:O	2.19	0.43
1:A:333:ARG:HB2	1:A:333:ARG:NH2	2.32	0.43
1:A:574:THR:OG1	1:A:577:GLN:HG3	2.19	0.43
1:A:452:ARG:HD2	2:A:896:HOH:O	2.18	0.42
1:A:5:VAL:HG21	1:A:17:TYR:CD2	2.55	0.42
1:A:648:LEU:HD12	1:A:648:LEU:HA	1.90	0.42
1:A:519:THR:HG21	1:A:621:LYS:NZ	2.35	0.41
1:A:644:LEU:HD23	1:A:644:LEU:HA	1.90	0.41
1:A:386:ILE:HD12	2:A:897:HOH:O	2.21	0.41
1:A:637:LEU:HD12	1:A:637:LEU:HA	1.79	0.41
1:A:686:PHE:HB3	1:A:692:PHE:CD2	2.56	0.41
1:A:376:LEU:HD22	1:A:391:GLY:HA2	2.03	0.41
1:A:455:PHE:CB	1:A:510:GLN:HB2	2.51	0.41
1:A:605:ILE:CD1	1:A:699:ILE:HD11	2.48	0.41
1:A:180:VAL:HG11	1:A:232:LEU:HD13	2.04	0.40
1:A:211:LEU:HD13	1:A:216:GLN:HB2	2.02	0.40
1:A:341:VAL:HG11	1:A:360:GLU:CG	2.51	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:667:PRO:O	1:A:670:ASP:HB2	2.21	0.40

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	690/741 (93%)	654 (95%)	32 (5%)	4 (1%)	21 18

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	588	ALA
1	A	387	ASN
1	A	205	PRO
1	A	323	PRO

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	A	568/594 (96%)	524 (92%)	44 (8%)	12 9

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	LEU
1	A	29	ASP
1	A	33	ASP
1	A	34	SER
1	A	79	ASP
1	A	85	LEU
1	A	91	LEU
1	A	96	ASP
1	A	101	THR
1	A	105	VAL
1	A	112	THR
1	A	166	ASP
1	A	186	SER
1	A	194	THR
1	A	196	GLN
1	A	202	VAL
1	A	206	ASP
1	A	207	LYS
1	A	208	SER
1	A	330	LEU
1	A	351	VAL
1	A	352	ASP
1	A	386	ILE
1	A	437	ARG
1	A	452	ARG
1	A	455	PHE
1	A	485	SER
1	A	528	THR
1	A	585	LYS
1	A	586	GLU
1	A	598	VAL
1	A	603	GLN
1	A	625	ILE
1	A	631	MET
1	A	637	LEU
1	A	638	ARG
1	A	643	LEU
1	A	648	LEU
1	A	661	SER
1	A	668	LEU
1	A	671	ARG
1	A	687	ASP
1	A	704	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	724	ARG

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	51	GLN
1	A	213	ASN
1	A	486	GLN
1	A	603	GLN
1	A	707	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 ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	701/741 (94%)	0.74	69 (9%) 13 13	13, 31, 57, 79	3 (0%)

All (69) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	321	THR	5.2
1	A	675	GLY	4.6
1	A	209	THR	4.3
1	A	324	GLY	4.3
1	A	588	ALA	3.8
1	A	197	ASP	3.7
1	A	354	ASP	3.7
1	A	214	PRO	3.7
1	A	220	TYR	3.5
1	A	205	PRO	3.4
1	A	668	LEU	3.1
1	A	327	GLN	3.1
1	A	329	THR	3.0
1	A	224	ALA	3.0
1	A	674	ALA	3.0
1	A	587	LEU	3.0
1	A	231	LEU	2.9
1	A	326	GLY	2.9
1	A	195	VAL	2.9
1	A	681	PRO	2.9
1	A	10	LEU	2.8
1	A	384	PRO	2.8
1	A	379	SER	2.8
1	A	667	PRO	2.8
1	A	311	LEU	2.7
1	A	669	VAL	2.7
1	A	210	GLY	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	201	VAL	2.7
1	A	154	ALA	2.6
1	A	202	VAL	2.6
1	A	723	ALA	2.6
1	A	473	GLU	2.6
1	A	44	ALA	2.5
1	A	683	ALA	2.5
1	A	461	LEU	2.5
1	A	301	ALA	2.5
1	A	208	SER	2.4
1	A	724	ARG	2.4
1	A	686	PHE	2.4
1	A	228	THR	2.4
1	A	182	LEU	2.3
1	A	595	ARG	2.3
1	A	671	ARG	2.3
1	A	600	ASN	2.3
1	A	704	GLN	2.2
1	A	625	ILE	2.2
1	A	181	PRO	2.2
1	A	230	VAL	2.2
1	A	298	ASP	2.2
1	A	78	MET	2.2
1	A	204	LEU	2.2
1	A	203	ALA	2.2
1	A	71	ARG	2.2
1	A	243	ILE	2.2
1	A	371	ILE	2.2
1	A	592	ASP	2.2
1	A	199	GLN	2.2
1	A	223	ALA	2.2
1	A	726	ALA	2.2
1	A	5	VAL	2.1
1	A	85	LEU	2.1
1	A	317	ASP	2.1
1	A	328	PHE	2.1
1	A	325	GLY	2.1
1	A	656	ALA	2.1
1	A	206	ASP	2.1
1	A	320	TYR	2.1
1	A	685	ASN	2.1
1	A	226	SER	2.1

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.