WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Jan 6, 2024 – 09:24 pm GMT

PDB ID	:	6GHX
Title	:	Alzheimer's Amyloid-Beta Peptide Fragment 31-35 in Complex with Cd-substituted Thermolysin
Authors	:	Leite, J.P.; Gales, L.
Deposited on	:	2018-05-09
Resolution	:	1.16 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
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)	:	2.36

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 1.16 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.
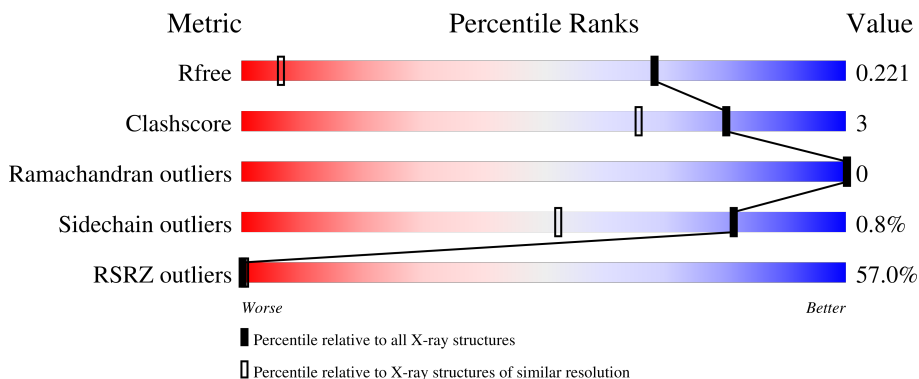
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1758 (1.20-1.12)
Clashscore	141614	1832 (1.20-1.12)
Ramachandran outliers	138981	1768 (1.20-1.12)
Sidechain outliers	138945	1768 (1.20-1.12)
RSRZ outliers	127900	1724 (1.20-1.12)

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 <=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	316	57% / 97%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	ILE	A	408	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 2841 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 Thermolysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	316	2449	1541	410	496	2	8	5	0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	A	4	4	4	0	0

- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cd		
3	A	1	1	1	0	0

- Molecule 4 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).
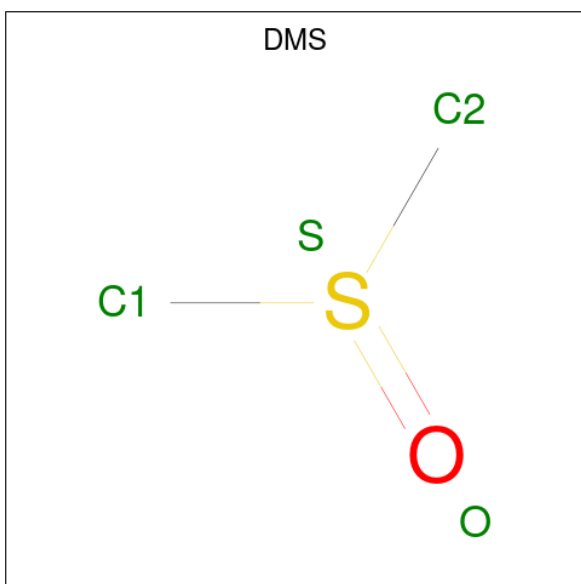
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 4	C 2	O 1	S 1	0	0

- Molecule 5 is ISOLEUCINE (three-letter code: ILE) (formula: $C_6H_{13}NO_2$).
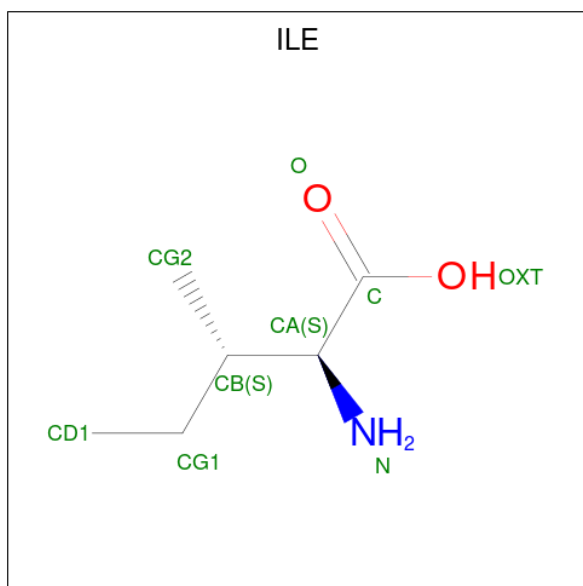
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 8	C 6	N 1	O 1	0	0
5	A	1	Total 8	C 6	N 1	O 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	367	Total 367	O 367	0	0

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: Thermolysin
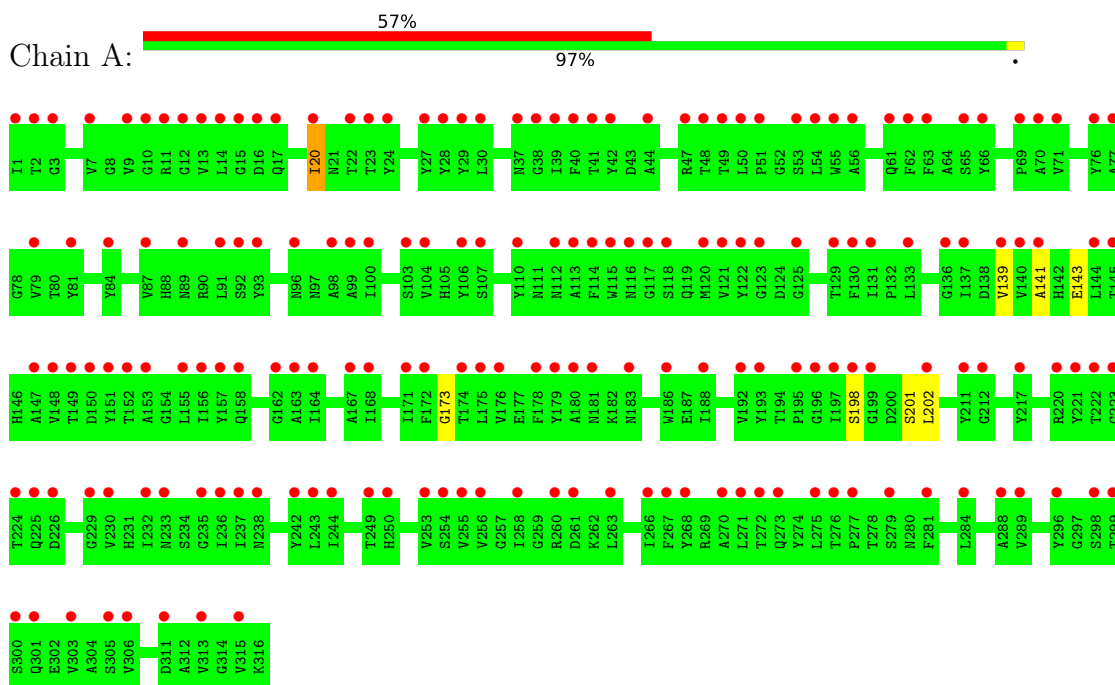
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	93.06Å 93.06Å 130.29Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.53 – 1.16 46.53 – 1.16	Depositor EDS
% Data completeness (in resolution range)	99.2 (46.53-1.16) 99.3 (46.53-1.16)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.97 (at 1.16Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.202 , 0.220 0.205 , 0.221	Depositor DCC
R_{free} test set	5749 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	18.0	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 54.7	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2841	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CA, CD, DMS

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	0.38	1/2523 (0.0%)	0.58	0/3434

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	201	SER	C-N	-6.44	1.19	1.34

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	2449	0	2300	12	0
2	A	4	0	0	0	0
3	A	1	0	0	0	0
4	A	4	0	6	0	0
5	A	16	0	20	10	0
6	A	367	0	0	1	0
All	All	2841	0	2326	13	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 (13) 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:202:LEU:HD21	5:A:408:ILE:CG2	1.71	1.19
1:A:202:LEU:HD21	5:A:408:ILE:HG23	1.38	1.02
1:A:202:LEU:HD21	5:A:408:ILE:HG22	1.50	0.92
1:A:202:LEU:CD2	5:A:408:ILE:CG2	2.54	0.85
1:A:202:LEU:CD2	5:A:408:ILE:HG23	2.13	0.78
1:A:202:LEU:HD11	5:A:408:ILE:HG21	1.75	0.68
1:A:202:LEU:HD11	5:A:408:ILE:CG2	2.28	0.63
1:A:20:ILE:HD13	1:A:20:ILE:N	2.26	0.50
1:A:139:VAL:O	1:A:143:GLU:HG2	2.13	0.49
1:A:202:LEU:CD1	5:A:408:ILE:HG21	2.41	0.48
1:A:202:LEU:CD1	5:A:408:ILE:CG2	2.96	0.43
5:A:407:ILE:N	6:A:509:HOH:O	2.53	0.42
1:A:141:ALA:HB3	1:A:173:GLY:HA2	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	319/316 (101%)	308 (97%)	11 (3%)	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	257/252 (102%)	255 (99%)	2 (1%)	81	52

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	20	ILE
1	A	198	SER

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

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 5 are monoatomic - leaving 3 for Mogul analysis.

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	$\#|Z| > 2$
4	DMS	A	406	-	3,3,3	0.64	0	3,3,3	0.88	0
5	ILE	A	407	-	6,7,8	0.59	0	5,8,10	0.93	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
5	ILE	A	408	-	6,7,8	0.81	0	5,8,10	1.25	1 (20%)

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
5	ILE	A	407	-	-	1/7/8/10	-
5	ILE	A	408	-	-	0/7/8/10	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	408	ILE	CB-CA-C	-2.22	109.44	112.83

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	407	ILE	O-C-CA-CB

There are no ring outliers.

2 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	407	ILE	1	0
5	A	408	ILE	9	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	201:SER	C	202:LEU	N	1.19

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(Å²)	Q<0.9
1	A	316/316 (100%)	2.18	180 (56%)	0	0	16, 21, 29, 36	2 (0%)

All (180) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	202	LEU	5.2
1	A	222	THR	4.5
1	A	198	SER	4.5
1	A	9	VAL	4.3
1	A	55	TRP	4.3
1	A	131	ILE	4.1
1	A	39	ILE	4.1
1	A	197	ILE	4.1
1	A	225	GLN	4.0
1	A	188	ILE	3.9
1	A	223	GLY	3.9
1	A	121	VAL	3.8
1	A	156[A]	ILE	3.7
1	A	168	ILE	3.6
1	A	15	GLY	3.6
1	A	96	ASN	3.6
1	A	144	LEU	3.6
1	A	56	ALA	3.6
1	A	20	ILE	3.5
1	A	249	THR	3.5
1	A	196	GLY	3.5
1	A	151	TYR	3.5
1	A	137	ILE	3.4
1	A	232	ILE	3.4
1	A	1	ILE	3.4
1	A	104	VAL	3.4
1	A	176	VAL	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	183	ASN	3.3
1	A	115	TRP	3.3
1	A	100	ILE	3.3
1	A	120	MET	3.3
1	A	171	ILE	3.3
1	A	179	TYR	3.3
1	A	276	THR	3.3
1	A	14	LEU	3.2
1	A	155	LEU	3.2
1	A	275	LEU	3.2
1	A	117	GLY	3.2
1	A	236	ILE	3.2
1	A	87	VAL	3.2
1	A	61	GLN	3.2
1	A	253	VAL	3.2
1	A	27	TYR	3.1
1	A	267	PHE	3.1
1	A	141	ALA	3.1
1	A	69	PRO	3.1
1	A	147	ALA	3.1
1	A	186	TRP	3.1
1	A	92	SER	3.0
1	A	110	TYR	3.0
1	A	193	TYR	3.0
1	A	149	THR	3.0
1	A	65[A]	SER	3.0
1	A	29	TYR	3.0
1	A	47	ARG	3.0
1	A	140	VAL	2.9
1	A	163	ALA	2.9
1	A	270	ALA	2.9
1	A	164	ILE	2.9
1	A	50	LEU	2.9
1	A	263	LEU	2.9
1	A	268	TYR	2.9
1	A	284	LEU	2.8
1	A	178	PHE	2.8
1	A	242	TYR	2.8
1	A	273	GLN	2.8
1	A	230	VAL	2.8
1	A	199	GLY	2.8
1	A	44	ALA	2.8

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	237	ILE	2.8
1	A	24	TYR	2.8
1	A	114	PHE	2.8
1	A	23	THR	2.7
1	A	243	LEU	2.7
1	A	311	ASP	2.7
1	A	221	TYR	2.7
1	A	7	VAL	2.7
1	A	254	SER	2.7
1	A	42	TYR	2.7
1	A	258	ILE	2.7
1	A	71	VAL	2.7
1	A	289	VAL	2.7
1	A	54	LEU	2.7
1	A	133	LEU	2.7
1	A	158	GLN	2.7
1	A	148	VAL	2.6
1	A	298	SER	2.6
1	A	123	GLY	2.6
1	A	181	ASN	2.6
1	A	305	SER	2.6
1	A	130	PHE	2.6
1	A	303	VAL	2.6
1	A	129	THR	2.6
1	A	76	TYR	2.6
1	A	93	TYR	2.6
1	A	122	TYR	2.6
1	A	211	TYR	2.6
1	A	217	TYR	2.6
1	A	40	PHE	2.6
1	A	38	GLY	2.5
1	A	277	PRO	2.5
1	A	30	LEU	2.5
1	A	116	ASN	2.5
1	A	229	GLY	2.5
1	A	79	VAL	2.5
1	A	153	ALA	2.5
1	A	175	LEU	2.5
1	A	157	TYR	2.5
1	A	107	SER	2.5
1	A	255	VAL	2.5
1	A	12	GLY	2.5

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	235	GLY	2.5
1	A	162	GLY	2.4
1	A	22	THR	2.4
1	A	152	THR	2.4
1	A	224	THR	2.4
1	A	139	VAL	2.4
1	A	313	VAL	2.4
1	A	174	THR	2.4
1	A	3	GLY	2.4
1	A	281	PHE	2.4
1	A	220	ARG	2.4
1	A	256	VAL	2.4
1	A	118	SER	2.4
1	A	172	PHE	2.4
1	A	272	THR	2.3
1	A	167	ALA	2.3
1	A	271	LEU	2.3
1	A	16	ASP	2.3
1	A	84	TYR	2.3
1	A	106	TYR	2.3
1	A	49	THR	2.3
1	A	53[A]	SER	2.3
1	A	13	VAL	2.3
1	A	315	VAL	2.3
1	A	99	ALA	2.3
1	A	81	TYR	2.3
1	A	91	LEU	2.3
1	A	266	ILE	2.3
1	A	113	ALA	2.3
1	A	288	ALA	2.3
1	A	125	GLY	2.3
1	A	63	PHE	2.3
1	A	11	ARG	2.2
1	A	261	ASP	2.2
1	A	10	GLY	2.2
1	A	212	GLY	2.2
1	A	233	ASN	2.2
1	A	28	TYR	2.2
1	A	70	ALA	2.2
1	A	51	PRO	2.2
1	A	150	ASP	2.2
1	A	250	HIS	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	98	ALA	2.2
1	A	17	GLN	2.2
1	A	301	GLN	2.2
1	A	260	ARG	2.2
1	A	279	SER	2.2
1	A	2	THR	2.2
1	A	145	THR	2.2
1	A	192	VAL	2.1
1	A	300	SER	2.1
1	A	306	VAL	2.1
1	A	180	ALA	2.1
1	A	62	PHE	2.1
1	A	41	THR	2.1
1	A	299	THR	2.1
1	A	103	SER	2.1
1	A	195	PRO	2.1
1	A	112	ASN	2.0
1	A	77	ALA	2.0
1	A	48	THR	2.0
1	A	66	TYR	2.0
1	A	296	TYR	2.0
1	A	226	ASP	2.0
1	A	89	ASN	2.0
1	A	136	GLY	2.0
1	A	244	ILE	2.0
1	A	37	ASN	2.0
1	A	238	ASN	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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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, 95th 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(Å^2)	Q<0.9
4	DMS	A	406	4/4	0.75	0.39	38,44,47,48	0
5	ILE	A	407	8/9	0.78	0.17	19,22,25,25	0
5	ILE	A	408	8/9	0.78	0.18	22,29,31,36	0
2	CA	A	404	1/1	0.99	0.12	23,23,23,23	0
3	CD	A	405	1/1	1.00	0.10	19,19,19,19	0
2	CA	A	402	1/1	1.00	0.29	12,12,12,12	0
2	CA	A	403	1/1	1.00	0.32	12,12,12,12	0
2	CA	A	401	1/1	1.00	0.39	11,11,11,11	0

6.5 Other polymers ⓘ

There are no such residues in this entry.