WORLDWIDE
wwPDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 22, 2024 – 01:27 AM EST

PDB ID	:	4RQZ
Title	:	re-refinement of 1soz, Crystal Structure of DegS protease in complex with an activating peptide
Authors	:	Sauer, R.T.; Grant, R.A.
Deposited on	:	2014-11-05
Resolution	:	2.40 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:

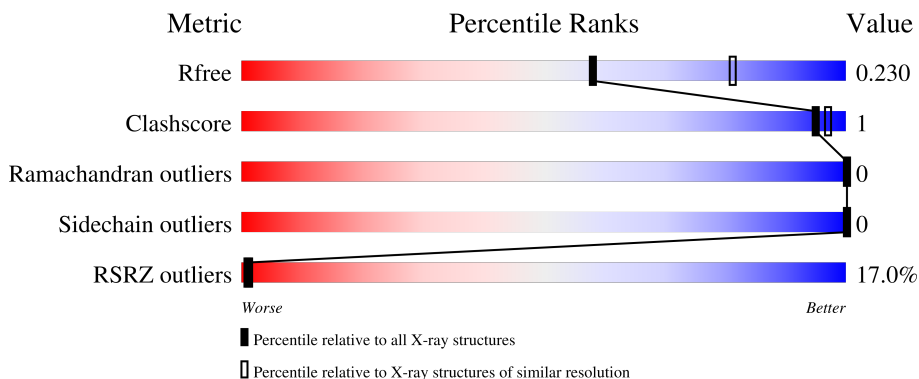|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.40 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	Percentile Ranks	Value
Rfree		0.230
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		17.0%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	314	12% / 85% • 13%
1	B	314	16% / 85% • 13%
1	C	314	15% / 82% • 14%
2	D	10	30% / 30% 70%
2	E	10	30% / 30% 10% 60%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	10	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 12205 atoms, of which 5952 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 Protease degS.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	272	Total	C	H	N	O	S	0	1	0
			3904	1232	1937	352	378	5			
1	B	274	Total	C	H	N	O	S	0	0	0
			3952	1247	1980	346	374	5			
1	C	271	Total	C	H	N	O	S	0	0	0
			3918	1235	1958	348	372	5			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	MET	-	expression tag	UNP H9UXC8
B	42	MET	-	expression tag	UNP H9UXC8
C	42	MET	-	expression tag	UNP H9UXC8

- Molecule 2 is a protein called activating peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	3	Total	C	H	N	O	0	0	0
			58	23	25	4	6			
2	E	4	Total	C	H	N	O	0	0	0
			74	28	34	5	7			
2	F	3	Total	C	H	N	O	0	0	0
			44	17	18	4	5			

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	117	Total	O	0	0
			117	117		
3	B	64	Total	O	0	0
			64	64		

Continued on next page...

Continued from previous page...

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	74	Total O 74 74		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.
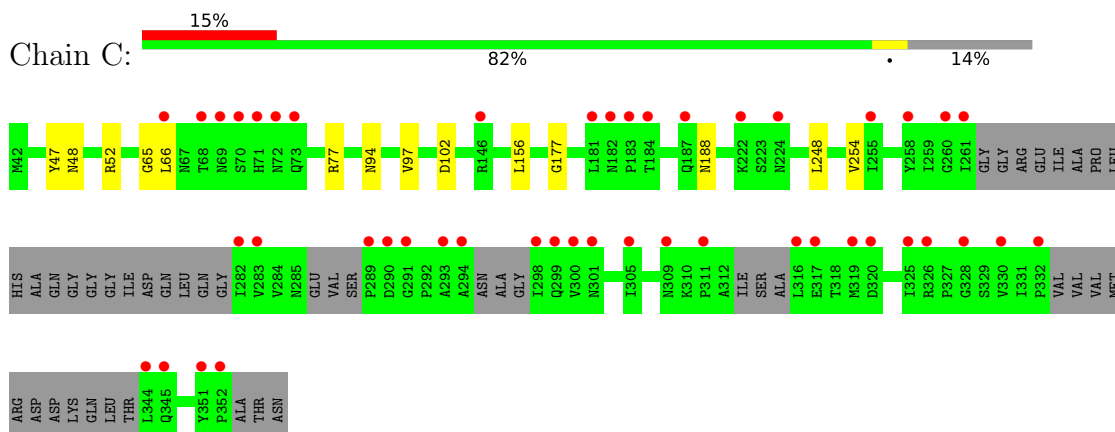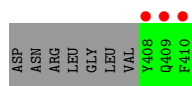
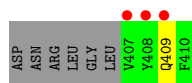- Molecule 1: Protease degS
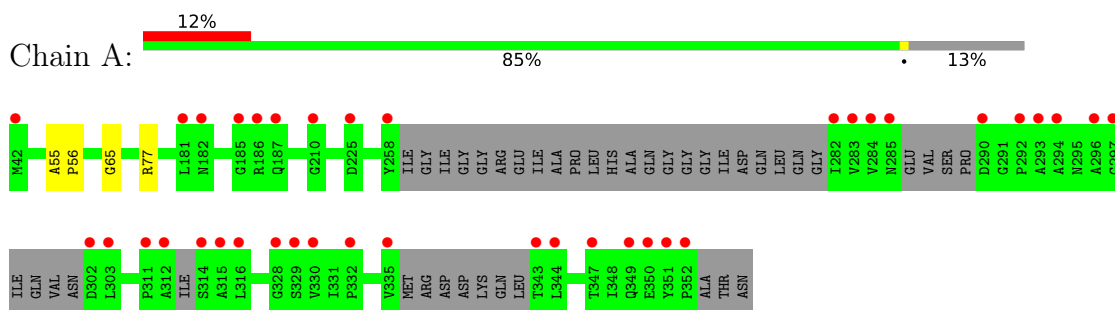
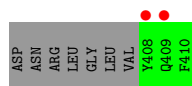- Molecule 1: Protease degS
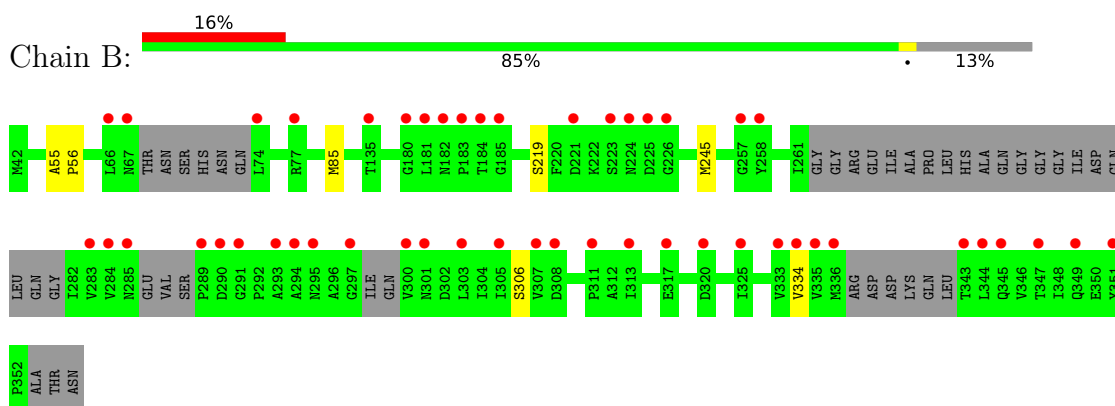
- Molecule 1: Protease degS

• Molecule 2: activating peptide

Chain D:

• Molecule 2: activating peptide

Chain E:

• Molecule 2: activating peptide

Chain F:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	205.98Å 142.71Å 41.17Å 90.00° 89.24° 90.00°	Depositor
Resolution (Å)	14.94 – 2.40 14.94 – 2.40	Depositor EDS
% Data completeness (in resolution range)	90.9 (14.94-2.40) 91.0 (14.94-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 2.39Å)	Xtriage
Refinement program	PHENIX dev_1760	Depositor
R, R_{free}	0.196 , 0.226 0.200 , 0.230	Depositor DCC
R_{free} test set	2135 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	59.9	Xtriage
Anisotropy	0.347	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 76.3	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.039 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12205	wwPDB-VP
Average B, all atoms (Å^2)	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.12% 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 ⓘ

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	0.50	0/1993	0.63	0/2716
1	B	0.44	0/1995	0.57	0/2718
1	C	0.45	0/1982	0.58	0/2698
2	D	0.31	0/34	0.52	0/43
2	E	0.39	0/41	0.54	0/53
2	F	0.30	0/26	0.30	0/32
All	All	0.46	0/6071	0.59	0/8260

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 ⓘ

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	1967	1937	1937	2	0
1	B	1972	1980	1980	4	0
1	C	1960	1958	1958	7	0
2	D	33	25	25	0	0
2	E	40	34	34	1	0
2	F	26	18	18	0	0
3	A	117	0	0	0	0
3	B	64	0	0	1	0
3	C	74	0	0	0	0
All	All	6253	5952	5952	14	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 (14) 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:219:SER:OG	3:B:463:HOH:O	2.04	0.74
1:B:85:MET:HG3	1:B:245:MET:SD	2.45	0.56
1:C:177:GLY:HA2	1:C:188:ASN:HB2	1.93	0.51
2:E:409:GLN:N	2:E:409:GLN:OE1	2.44	0.50
1:C:94:ASN:HB2	1:C:97:VAL:HG23	1.94	0.49
1:C:47:TYR:HB3	1:C:156:LEU:HD11	1.96	0.48
1:A:55:ALA:N	1:A:56:PRO:CD	2.77	0.47
1:C:66:LEU:N	1:C:102:ASP:OD2	2.48	0.47
1:B:55:ALA:N	1:B:56:PRO:CD	2.79	0.45
1:C:48:ASN:OD1	1:C:52:ARG:NH1	2.50	0.45
1:C:65:GLY:HA3	1:C:77:ARG:HD2	1.99	0.43
1:A:65:GLY:HA3	1:A:77:ARG:HD3	2.02	0.42
1:C:248:LEU:HD21	1:C:254:VAL:HG23	2.02	0.42
1:B:306:SER:HB3	1:B:334:VAL:HB	2.03	0.41

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	261/314 (83%)	257 (98%)	4 (2%)	0	100	100
1	B	262/314 (83%)	259 (99%)	3 (1%)	0	100	100
1	C	259/314 (82%)	254 (98%)	5 (2%)	0	100	100
2	D	1/10 (10%)	0	1 (100%)	0	100	100
2	E	2/10 (20%)	2 (100%)	0	0	100	100
2	F	1/10 (10%)	1 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	786/972 (81%)	773 (98%)	13 (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	202/253 (80%)	202 (100%)	0	100	100
1	B	204/253 (81%)	204 (100%)	0	100	100
1	C	202/253 (80%)	202 (100%)	0	100	100
2	D	3/9 (33%)	3 (100%)	0	100	100
2	E	4/9 (44%)	4 (100%)	0	100	100
2	F	2/9 (22%)	2 (100%)	0	100	100
All	All	617/786 (78%)	617 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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 ⓘ

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(Å²)	Q<0.9
1	A	272/314 (86%)	0.52	38 (13%)	2	2	39, 60, 162, 193	0
1	B	274/314 (87%)	0.75	49 (17%)	1	1	41, 74, 166, 195	0
1	C	271/314 (86%)	0.73	46 (16%)	1	1	44, 73, 169, 196	0
2	D	3/10 (30%)	3.48	3 (100%)	0	0	141, 141, 151, 151	0
2	E	4/10 (40%)	5.58	3 (75%)	0	0	136, 150, 150, 154	0
2	F	3/10 (30%)	5.19	2 (66%)	0	0	158, 158, 158, 166	0
All	All	827/972 (85%)	0.72	141 (17%)	1	1	39, 70, 166, 196	0

All (141) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	408	TYR	10.4
2	F	408	TYR	9.7
1	C	289	PRO	9.5
1	B	301	ASN	9.1
2	E	407	VAL	8.5
1	B	300	VAL	8.2
1	A	290	ASP	7.5
1	B	66	LEU	7.3
1	A	294	ALA	7.2
1	C	181	LEU	7.0
1	B	290	ASP	7.0
1	C	283	VAL	6.7
1	A	349	GLN	6.6
1	B	313	ILE	6.6
1	A	328	GLY	6.4
1	C	309	ASN	6.3
1	C	68	THR	6.1
1	B	224	ASN	6.0
1	C	300	VAL	6.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	351	TYR	6.0
1	C	316	LEU	5.8
1	B	283	VAL	5.8
1	C	344	LEU	5.8
1	C	328	GLY	5.8
1	C	298	ILE	5.7
1	B	294	ALA	5.7
1	C	311	PRO	5.6
1	C	182	ASN	5.6
1	B	67	ASN	5.4
1	A	315	ALA	5.3
1	A	314	SER	5.3
1	A	258	TYR	5.3
1	A	42	MET	5.2
1	B	223	SER	5.1
1	C	261	ILE	5.0
1	B	311	PRO	5.0
1	A	329	SER	5.0
1	B	289	PRO	4.9
1	A	296	ALA	4.8
1	A	285	ASN	4.8
1	C	332	PRO	4.8
1	B	291	GLY	4.8
1	C	183	PRO	4.8
1	C	345	GLN	4.8
1	C	282	ILE	4.8
1	B	285	ASN	4.6
1	B	221	ASP	4.6
1	A	283	VAL	4.5
1	A	182	ASN	4.5
2	D	408	TYR	4.5
1	A	282	ILE	4.4
1	A	344	LEU	4.4
1	A	343	THR	4.4
1	B	135	THR	4.4
1	A	225	ASP	4.4
1	A	311	PRO	4.3
2	F	409	GLN	4.3
1	B	183	PRO	4.3
1	C	305	ILE	4.3
1	A	293	ALA	4.2
1	C	70	SER	4.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	352	PRO	4.1
1	C	317	GLU	4.1
1	A	303	LEU	4.1
1	B	349	GLN	4.0
1	C	352	PRO	4.0
1	A	181	LEU	4.0
1	B	334	VAL	4.0
1	C	290	ASP	4.0
1	A	351	TYR	4.0
1	C	294	ALA	3.9
1	A	330	VAL	3.8
1	B	226	GLY	3.8
1	C	260	GLY	3.8
1	B	317	GLU	3.8
1	C	72	ASN	3.7
1	A	297	GLY	3.7
1	B	303	LEU	3.5
1	C	330	VAL	3.5
1	C	184	THR	3.5
1	A	284	VAL	3.4
1	A	347	THR	3.4
1	C	255	ILE	3.4
1	B	284	VAL	3.4
1	B	293	ALA	3.3
1	C	258	TYR	3.3
1	C	66	LEU	3.3
1	B	184	THR	3.3
2	D	409	GLN	3.2
1	B	225	ASP	3.2
1	C	325	ILE	3.2
1	B	258	TYR	3.2
1	B	297	GLY	3.2
1	A	187	GLN	3.2
2	E	409	GLN	3.2
1	B	335	VAL	3.2
1	A	335	VAL	3.1
1	B	308	ASP	3.1
1	B	345	GLN	3.1
1	B	182	ASN	3.1
1	C	301	ASN	3.0
1	B	77	ARG	2.9
1	C	69	ASN	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	299	GLN	2.9
1	B	336	MET	2.8
1	C	320	ASP	2.8
2	D	410	PHE	2.7
1	A	332	PRO	2.7
1	C	326	ARG	2.6
1	C	71	HIS	2.6
1	C	222	LYS	2.6
1	B	185	GLY	2.6
1	B	307	VAL	2.5
1	A	316	LEU	2.5
1	B	181	LEU	2.5
1	A	350	GLU	2.5
1	C	351	TYR	2.4
1	C	73	GLN	2.4
1	A	302	ASP	2.4
1	C	146	ARG	2.4
1	A	185	GLY	2.4
1	C	187	GLN	2.4
1	B	343	THR	2.3
1	A	312	ALA	2.3
1	C	291	GLY	2.3
1	A	210	GLY	2.3
1	B	347	THR	2.3
1	B	74	LEU	2.3
1	A	292	PRO	2.2
1	B	257	GLY	2.2
1	B	333	VAL	2.2
1	B	325	ILE	2.2
1	B	344	LEU	2.2
1	A	186	ARG	2.2
1	B	295	ASN	2.2
1	B	305	ILE	2.2
1	C	224	ASN	2.2
1	C	319	MET	2.1
1	C	293	ALA	2.0
1	B	180	GLY	2.0
1	B	320	ASP	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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.