W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

May 3, 2025 – 10:37 AM EDT

PDB ID	:	2FSF / pdb_00002fsf
Title	:	Escherichia coli SecA, the preprotein translocase dimeric ATPase
Authors	:	Papanikolau, Y.; Petratos, K.; Economou, A.
Deposited on	:	2006-01-23
Resolution	:	2.00 Å(reported)

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

|		
MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

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 2.00 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		0.297
Clashscore		12
Ramachandran outliers		3.1%
Sidechain outliers		10.8%
RSRZ outliers		28.4%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
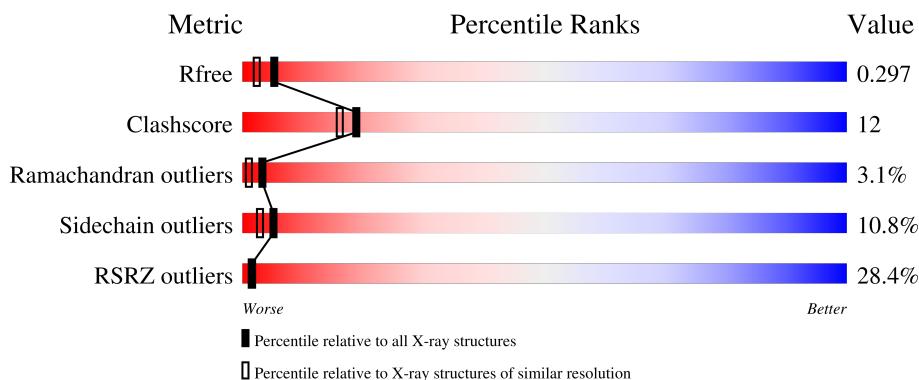
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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	853	18% — 59% — 17% · · 20%
1	B	853	29% — 51% — 23% — 9% · 15%
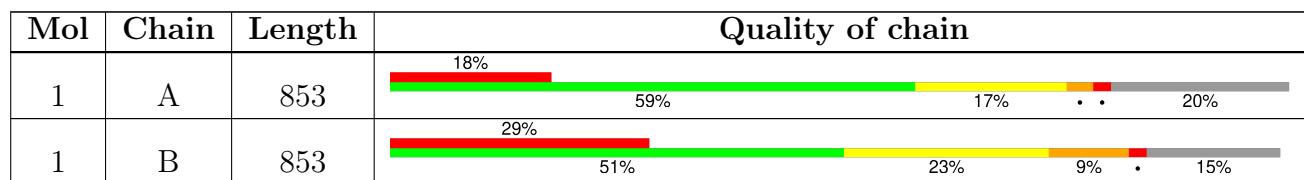
2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 11637 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 Preprotein translocase secA subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	686	Total	C	N	O	S	0	0	0
			5462	3426	966	1044	26			
1	B	723	Total	C	N	O	S	0	0	0
			5741	3600	1009	1104	28			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	177	Total	O	0	0
			177	177		
2	B	257	Total	O	0	0
			257	257		

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: Preprotein translocase secA subunit
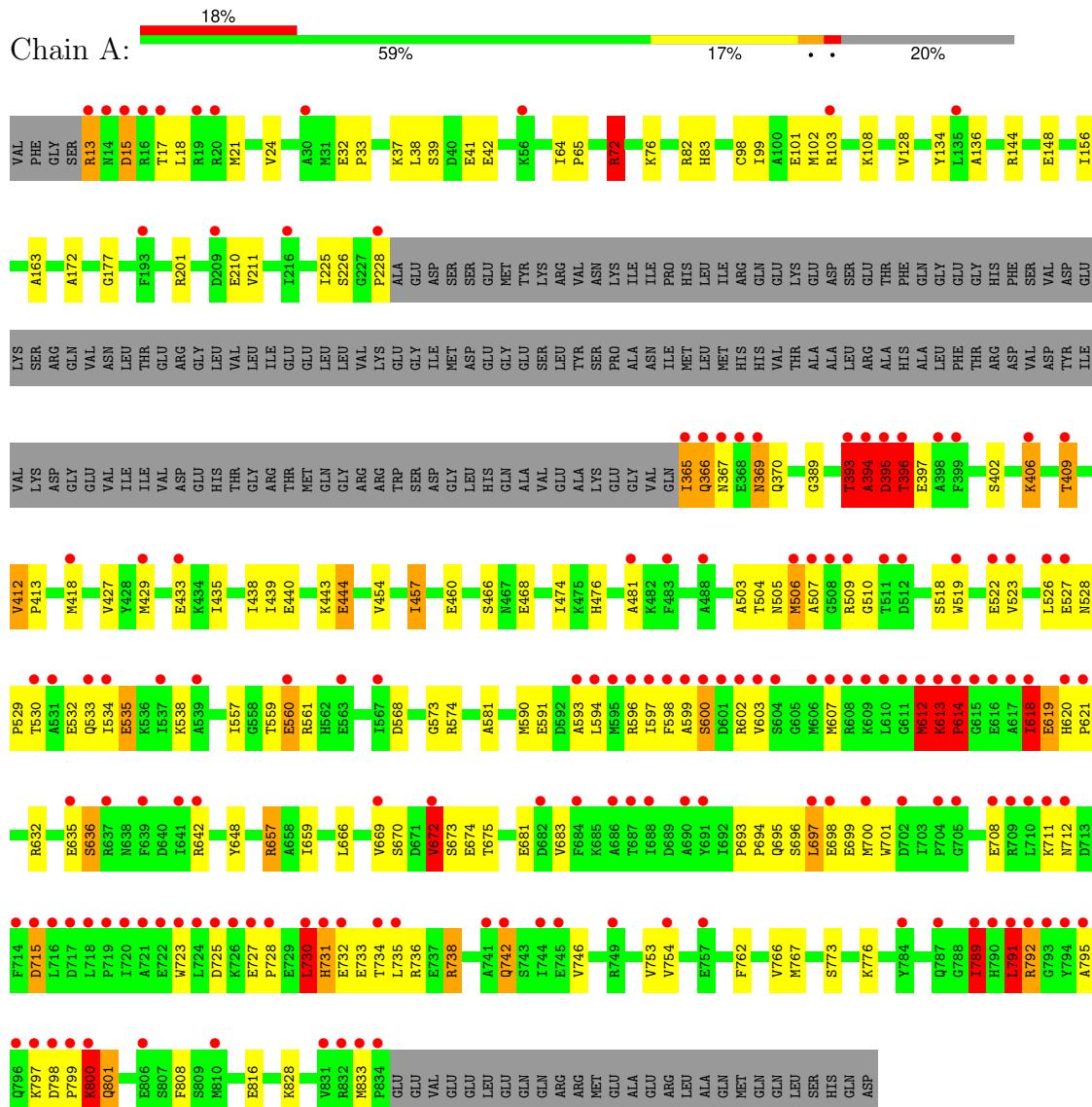
• Molecule 1: Preprotein translocase secA subunit
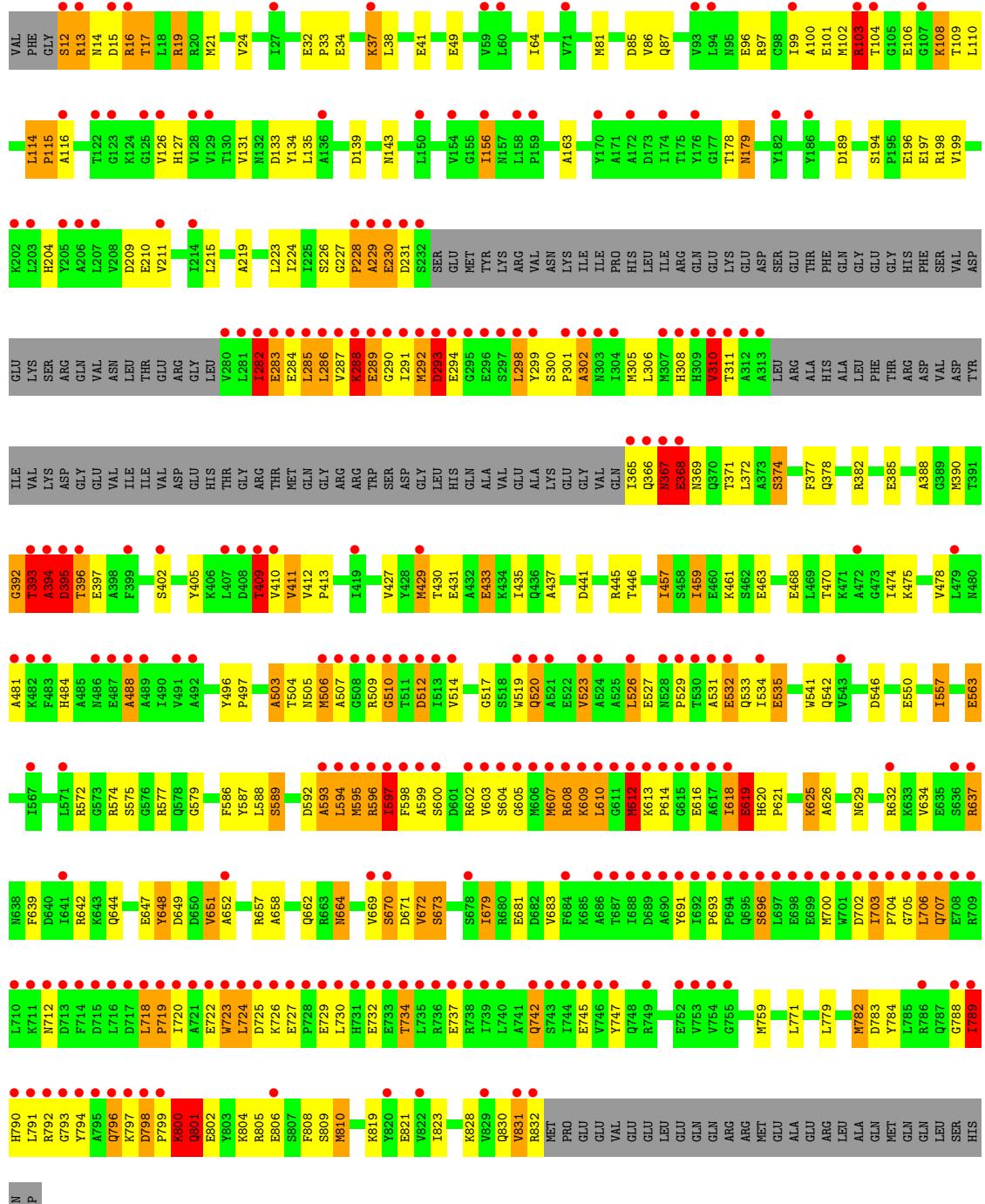
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	75.03Å 90.17Å 163.05Å 90.00° 100.48° 90.00°	Depositor
Resolution (Å)	19.98 – 2.00 19.98 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.1 (19.98-2.00) 97.0 (19.98-2.00)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.213 , 0.261 0.260 , 0.297	Depositor DCC
R_{free} test set	7010 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	35.8	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 53.8	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.015 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	11637	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% 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	1.61	44/5552 (0.8%)	1.34	29/7491 (0.4%)
1	B	1.84	95/5833 (1.6%)	1.48	63/7870 (0.8%)
All	All	1.73	139/11385 (1.2%)	1.41	92/15361 (0.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	B	0	9
All	All	0	13

The worst 5 of 139 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	697	LEU	C-N	14.88	1.56	1.33
1	B	393	THR	N-CA	13.38	1.63	1.46
1	B	395	ASP	N-CA	13.01	1.62	1.46
1	B	64	ILE	CA-CB	12.69	1.61	1.54
1	A	698	GLU	C-O	11.84	1.39	1.24

The worst 5 of 92 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	613	LYS	CA-C-N	10.33	132.76	119.84
1	A	613	LYS	C-N-CA	10.33	132.76	119.84
1	A	800	LYS	N-CA-C	-9.18	91.25	110.80
1	B	608	ARG	N-CA-C	-8.89	101.53	111.14
1	B	435	ILE	N-CA-C	-8.75	102.33	110.82

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	13	ARG	Peptide
1	A	393	THR	Peptide
1	A	395	ASP	Peptide
1	A	697	LEU	Mainchain
1	B	229	ALA	Peptide

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	5462	0	5460	95	0
1	B	5741	0	5731	167	1
2	A	177	0	0	12	0
2	B	257	0	0	29	0
All	All	11637	0	11191	262	1

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 12.

The worst 5 of 262 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:96:GLU:HG3	2:B:1033:HOH:O	1.46	1.13
1:A:799:PRO:HA	2:A:1026:HOH:O	1.50	1.08
1:B:800:LYS:O	1:B:801:GLN:HB2	1.61	1.01
1:A:799:PRO:C	1:A:800:LYS:O	1.95	0.97
1:B:103:ARG:HG3	1:B:103:ARG:HH11	1.26	0.97

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:600:SER:OG	1:B:830:GLN:O[1_655]	1.97	0.23

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	682/853 (80%)	635 (93%)	31 (4%)	16 (2%)	5	2
1	B	717/853 (84%)	650 (91%)	39 (5%)	28 (4%)	2	1
All	All	1399/1706 (82%)	1285 (92%)	70 (5%)	44 (3%)	3	1

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	394	ALA
1	A	396	THR
1	A	507	ALA
1	A	599	ALA
1	A	614	PRO

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	582/728 (80%)	525 (90%)	57 (10%)	6	4
1	B	613/728 (84%)	541 (88%)	72 (12%)	4	2
All	All	1195/1456 (82%)	1066 (89%)	129 (11%)	5	3

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	725	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	742	GLN
1	A	738	ARG
1	A	735	LEU
1	B	783	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	528	ASN
1	B	545	HIS
1	B	761	HIS
1	B	542	GLN
1	B	620	HIS

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	686/853 (80%)	1.41	153 (22%)	3	2	40, 53, 84, 102	0
1	B	723/853 (84%)	1.86	247 (34%)	1	1	38, 54, 87, 102	0
All	All	1409/1706 (82%)	1.64	400 (28%)	1	1	38, 53, 86, 102	0

The worst 5 of 400 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	728	PRO	10.5
1	A	597	ILE	10.3
1	B	789	ILE	8.7
1	B	598	PHE	8.4
1	B	291	ILE	8.1

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK