



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 25, 2026 – 08:39 PM EDT

PDB ID : 3NL2 / pdb_00003nl2
Title : The Crystal Structure of *Candida glabrata* THI6, a Bifunctional Enzyme involved in Thiamin Biosynthesis of Eukaryotes
Authors : Paul, D.; Chatterjee, A.; Begley, T.P.; Ealick, S.E.
Deposited on : 2010-06-21
Resolution : 3.08 Å (reported)

This is a wwPDB X-ray Structure Validation Summary 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

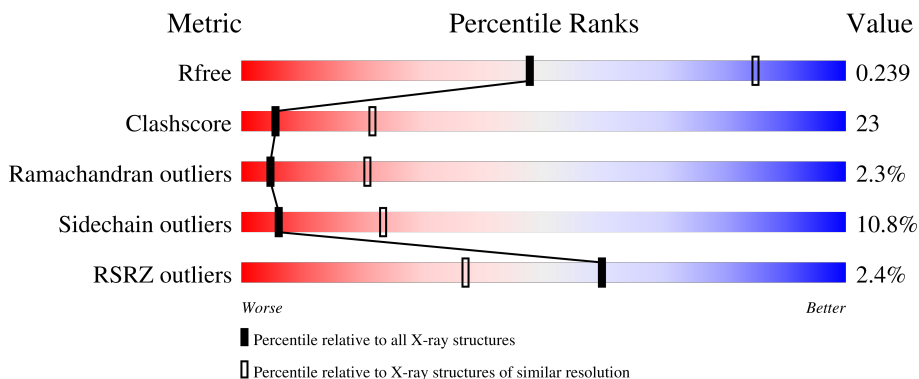
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.08 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.



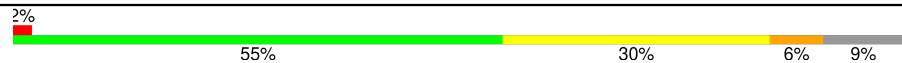
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	2010 (3.10-3.06)
Clashscore	190562	2102 (3.10-3.06)
Ramachandran outliers	187476	1982 (3.10-3.06)
Sidechain outliers	187428	1981 (3.10-3.06)
RSRZ outliers	180081	2010 (3.10-3.06)

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	540	
1	B	540	
1	C	540	
1	D	540	
1	E	540	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	540	 <p>A horizontal bar chart showing the quality distribution of chain F. The bar is divided into five segments: a small red segment (2%), a large green segment (55%), a yellow segment (30%), a small orange segment (6%), and a small grey segment (9%).</p>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 22460 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.

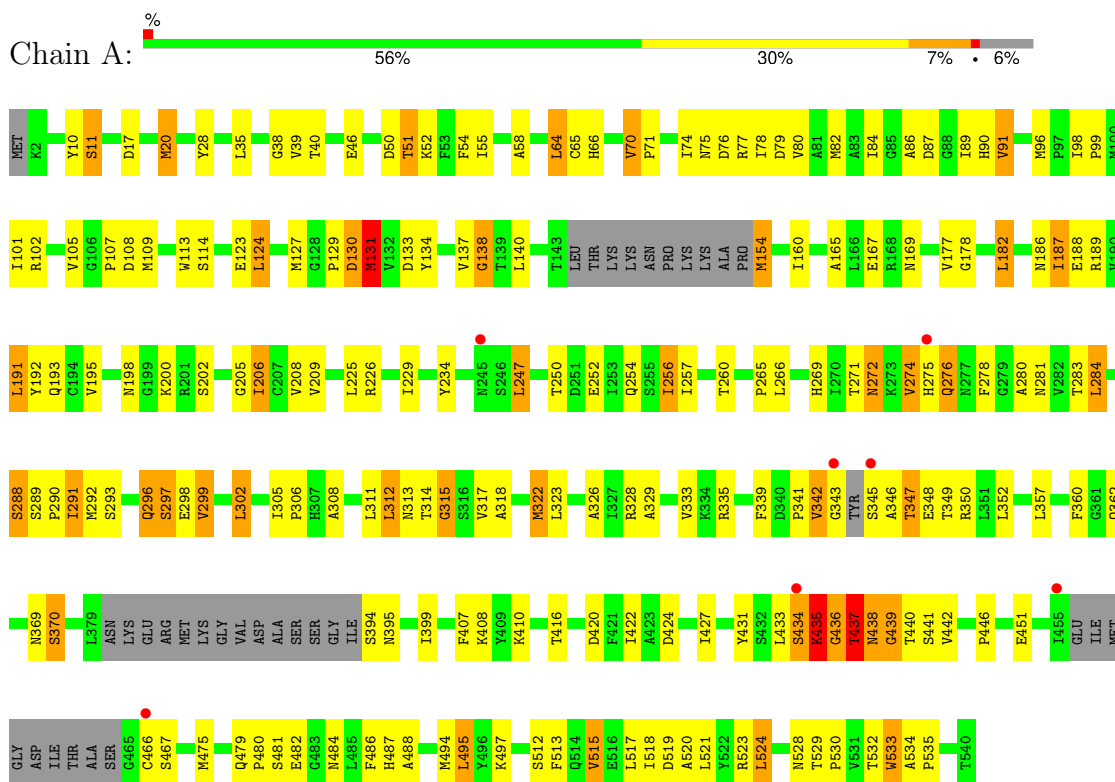
- Molecule 1 is a protein called Thiamine biosynthetic bifunctional enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	505	3764	2376	634	733	21	0	0	0
1	B	505	3766	2382	635	728	21	0	0	0
1	C	498	3727	2356	626	724	21	0	0	0
1	D	501	3739	2360	630	727	22	0	0	0
1	E	503	3767	2383	634	728	22	0	0	0
1	F	494	3697	2342	623	711	21	0	0	0

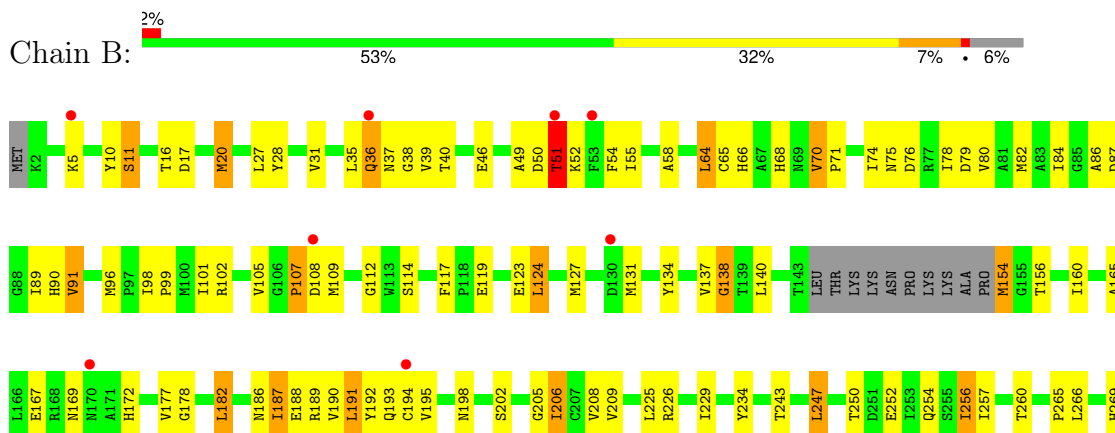
3 Residue-property plots

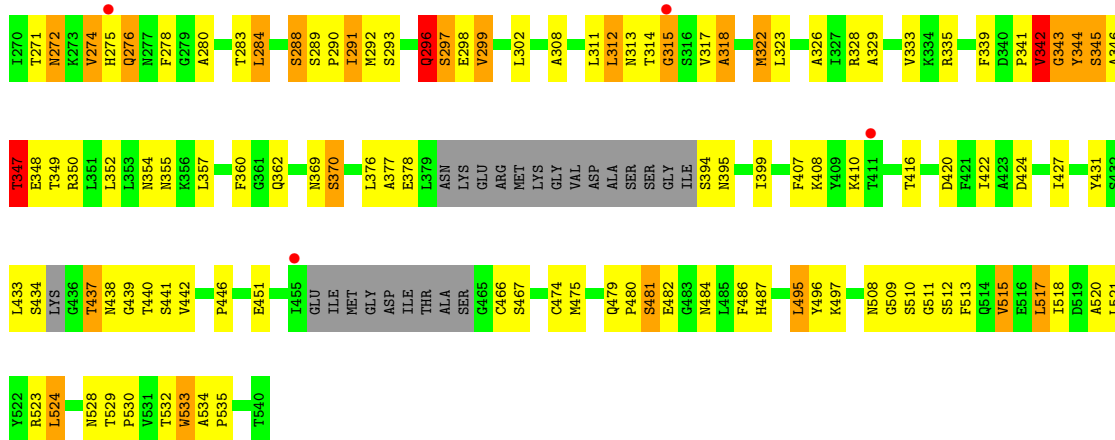
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Thiamine biosynthetic bifunctional enzyme

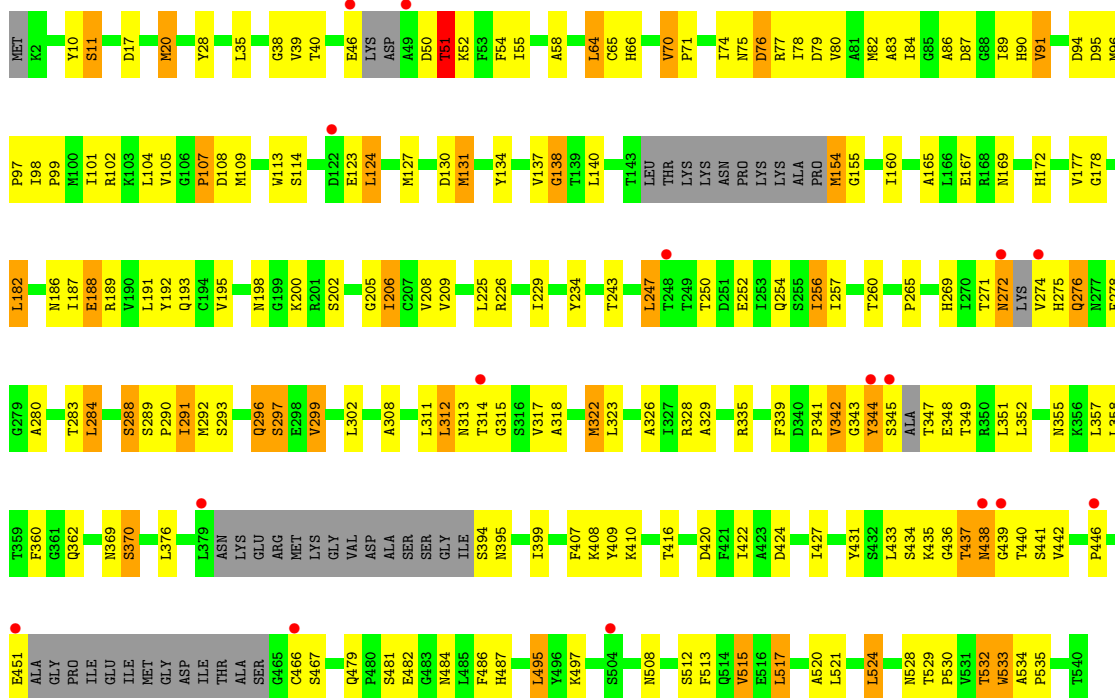


- Molecule 1: Thiamine biosynthetic bifunctional enzyme

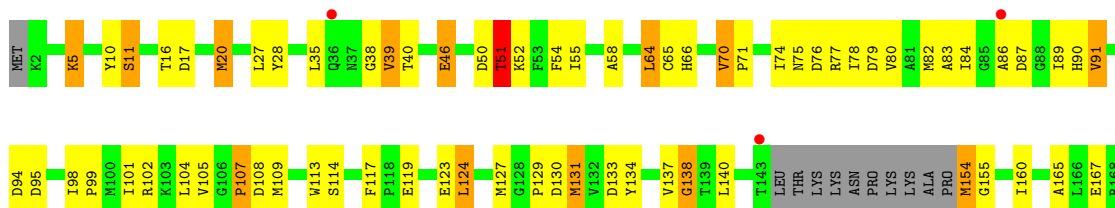


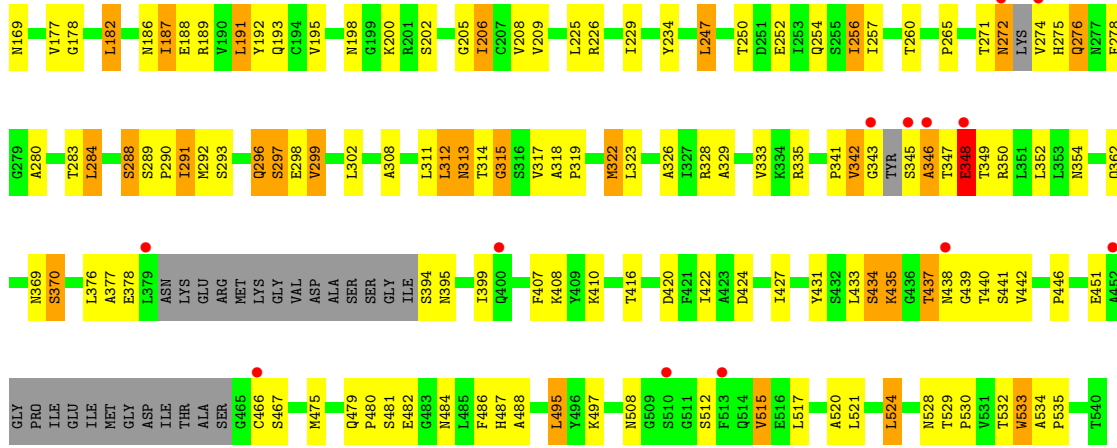


• Molecule 1: Thiamine biosynthetic bifunctional enzyme

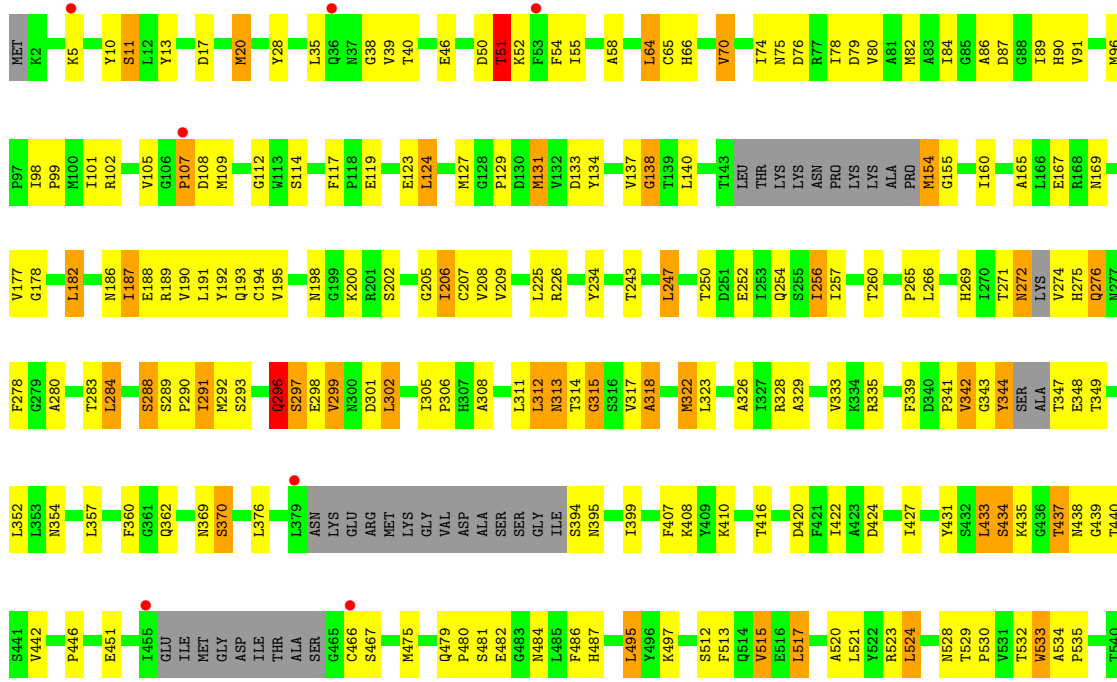


• Molecule 1: Thiamine biosynthetic bifunctional enzyme

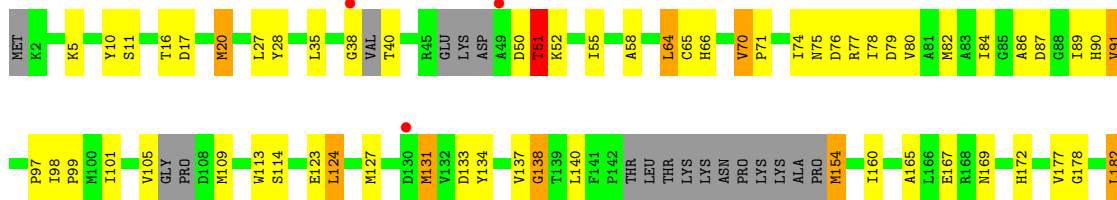




● Molecule 1: Thiamine biosynthetic bifunctional enzyme



● Molecule 1: Thiamine biosynthetic bifunctional enzyme



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	102.47Å 153.01Å 146.75Å 90.00° 102.29° 90.00°	Depositor
Resolution (Å)	37.00 – 3.08 37.00 – 3.08	Depositor EDS
% Data completeness (in resolution range)	85.8 (37.00-3.08) 93.2 (37.00-3.08)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 3.07Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.214 , 0.244 0.213 , 0.239	Depositor DCC
R_{free} test set	3826 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	66.9	Xtrriage
Anisotropy	0.599	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 46.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	22460	wwPDB-VP
Average B, all atoms (Å ²)	78.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.37% 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	0.53	0/3823	0.98	15/5189 (0.3%)
1	B	0.53	0/3826	0.94	10/5192 (0.2%)
1	C	0.54	0/3784	0.90	3/5131 (0.1%)
1	D	0.54	0/3796	0.92	11/5149 (0.2%)
1	E	0.54	0/3826	0.93	9/5188 (0.2%)
1	F	0.53	0/3752	0.90	5/5083 (0.1%)
All	All	0.54	0/22807	0.93	53/30932 (0.2%)

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	1
1	D	0	1
1	E	0	2
All	All	0	8

There are no bond length outliers.

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	434	SER	CB-CA-C	-11.65	92.96	111.48
1	A	437	THR	N-CA-C	10.43	133.02	110.80
1	E	434	SER	N-CA-C	-9.72	102.67	112.97
1	D	345	SER	N-CA-CB	-9.37	94.57	110.50
1	A	434	SER	N-CA-C	7.91	122.28	112.47

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	129	PRO	Peptide
1	A	435	LYS	Peptide
1	A	436	GLY	Peptide
1	A	438	ASN	Peptide
1	B	347	THR	Peptide

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	3764	0	3745	178	0
1	B	3766	0	3751	201	0
1	C	3727	0	3706	193	0
1	D	3739	0	3715	204	0
1	E	3767	0	3765	172	0
1	F	3697	0	3690	163	0
All	All	22460	0	22372	1052	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 23.

The worst 5 of 1052 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:F:339:PHE:CE2	1:F:341:PRO:HG3	1.66	1.27
1:A:339:PHE:CE2	1:A:341:PRO:HG3	1.76	1.19
1:C:265:PRO:HD2	1:C:288:SER:HB3	1.39	1.04
1:B:345:SER:HA	1:B:347:THR:N	1.73	1.03
1:D:274:VAL:HG13	1:E:292:MET:HG3	1.42	1.01

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	495/540 (92%)	445 (90%)	41 (8%)	9 (2%)	6	26
1	B	495/540 (92%)	439 (89%)	41 (8%)	15 (3%)	3	17
1	C	484/540 (90%)	434 (90%)	38 (8%)	12 (2%)	4	19
1	D	489/540 (91%)	440 (90%)	36 (7%)	13 (3%)	4	18
1	E	491/540 (91%)	444 (90%)	38 (8%)	9 (2%)	6	26
1	F	476/540 (88%)	434 (91%)	32 (7%)	10 (2%)	5	22
All	All	2930/3240 (90%)	2636 (90%)	226 (8%)	68 (2%)	5	21

5 of 68 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	131	MET
1	A	347	THR
1	A	370	SER
1	B	370	SER
1	C	344	TYR

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	408/449 (91%)	363 (89%)	45 (11%)	6	23
1	B	406/449 (90%)	362 (89%)	44 (11%)	6	23
1	C	404/449 (90%)	361 (89%)	43 (11%)	6	24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	404/449 (90%)	360 (89%)	44 (11%)	6	23
1	E	409/449 (91%)	364 (89%)	45 (11%)	6	23
1	F	399/449 (89%)	357 (90%)	42 (10%)	6	24
All	All	2430/2694 (90%)	2167 (89%)	263 (11%)	6	23

5 of 263 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	191	LEU
1	F	284	LEU
1	F	515	VAL
1	C	188	GLU
1	C	124	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 82 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	90	HIS
1	F	75	ASN
1	E	193	GLN
1	E	355	ASN
1	F	276	GLN

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 oligosaccharides 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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	505/540 (93%)	0.06	7 (1%) 73 51	51, 76, 110, 138	0
1	B	505/540 (93%)	0.23	12 (2%) 59 37	39, 77, 112, 137	0
1	C	498/540 (92%)	0.12	16 (3%) 50 28	52, 75, 108, 136	0
1	D	501/540 (92%)	0.11	16 (3%) 50 28	20, 75, 108, 139	0
1	E	503/540 (93%)	0.11	7 (1%) 73 51	39, 76, 111, 138	0
1	F	494/540 (91%)	0.11	13 (2%) 57 34	37, 75, 106, 134	0
All	All	3006/3240 (92%)	0.12	71 (2%) 59 37	20, 76, 110, 139	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	343	GLY	5.5
1	A	343	GLY	5.5
1	F	38	GLY	5.2
1	D	345	SER	4.4
1	F	438	ASN	4.4

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.