wwPDB X-ray Structure Validation Summary Report (i)

May 17, 2020 – 04:27 am BST

PDB ID	:	5I6I
Title	:	Crystal structure of a dBCCP-variant of Chaetomium thermophilum acetyl-CoA carboxylase
Authors	:	Hunkeler, M.; Stuttfeld, E.; Hagmann, A.; Imseng, S.; Maier, T.
Deposited on	:	2016-02-16
Resolution	:	8.40 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

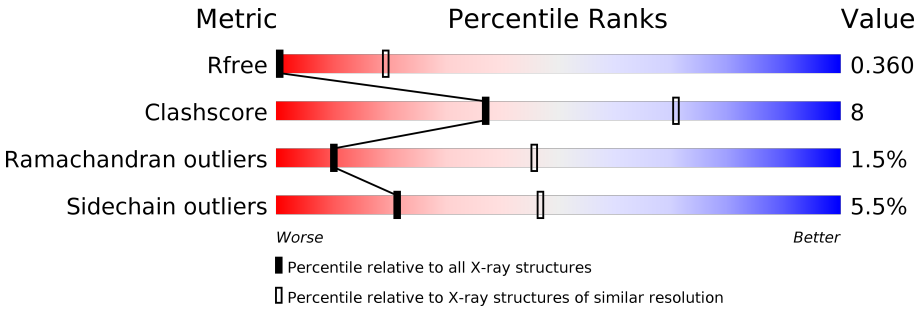
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 8.40 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1005 (11.50-3.90)
Clashscore	141614	1070 (11.50-3.90)
Ramachandran outliers	138981	1003 (11.50-3.90)
Sidechain outliers	138945	1003 (11.50-3.86)

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 on 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%

Mol	Chain	Length	Quality of chain
1	A	2211	48% 13% • 37%
1	B	2211	49% 13% • 37%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 22445 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 Acetyl-CoA carboxylase-like protein,Acetyl-CoA carboxylase-like protein,Acetyl-CoA carboxylase-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1399	Total	C	N	O	S	0	0	0
			11224	7122	1978	2086	38			
1	B	1394	Total	C	N	O	S	0	0	0
			11221	7120	1976	2086	39			

There are 8 discrepancies between the modelled and reference sequences:

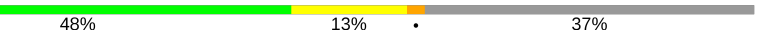
Chain	Residue	Modelled	Actual	Comment	Reference
A	63	GLY	-	expression tag	UNP G0S3L5
A	763	GLY	-	linker	UNP G0S3L5
A	764	SER	-	linker	UNP G0S3L5
A	765	GLY	-	linker	UNP G0S3L5
B	63	GLY	-	expression tag	UNP G0S3L5
B	763	GLY	-	linker	UNP G0S3L5
B	764	SER	-	linker	UNP G0S3L5
B	765	GLY	-	linker	UNP G0S3L5

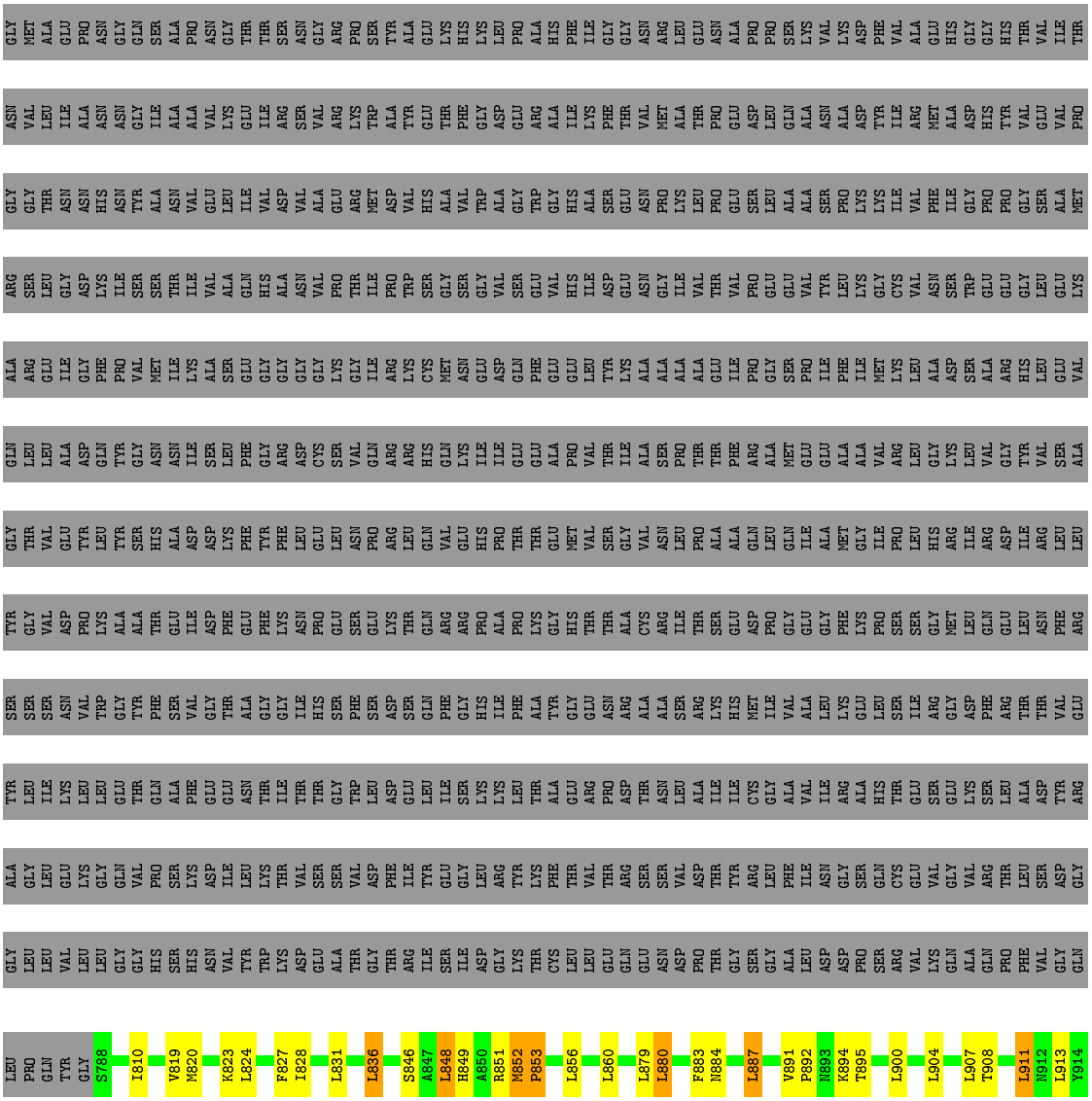
3 Residue-property plots (i)

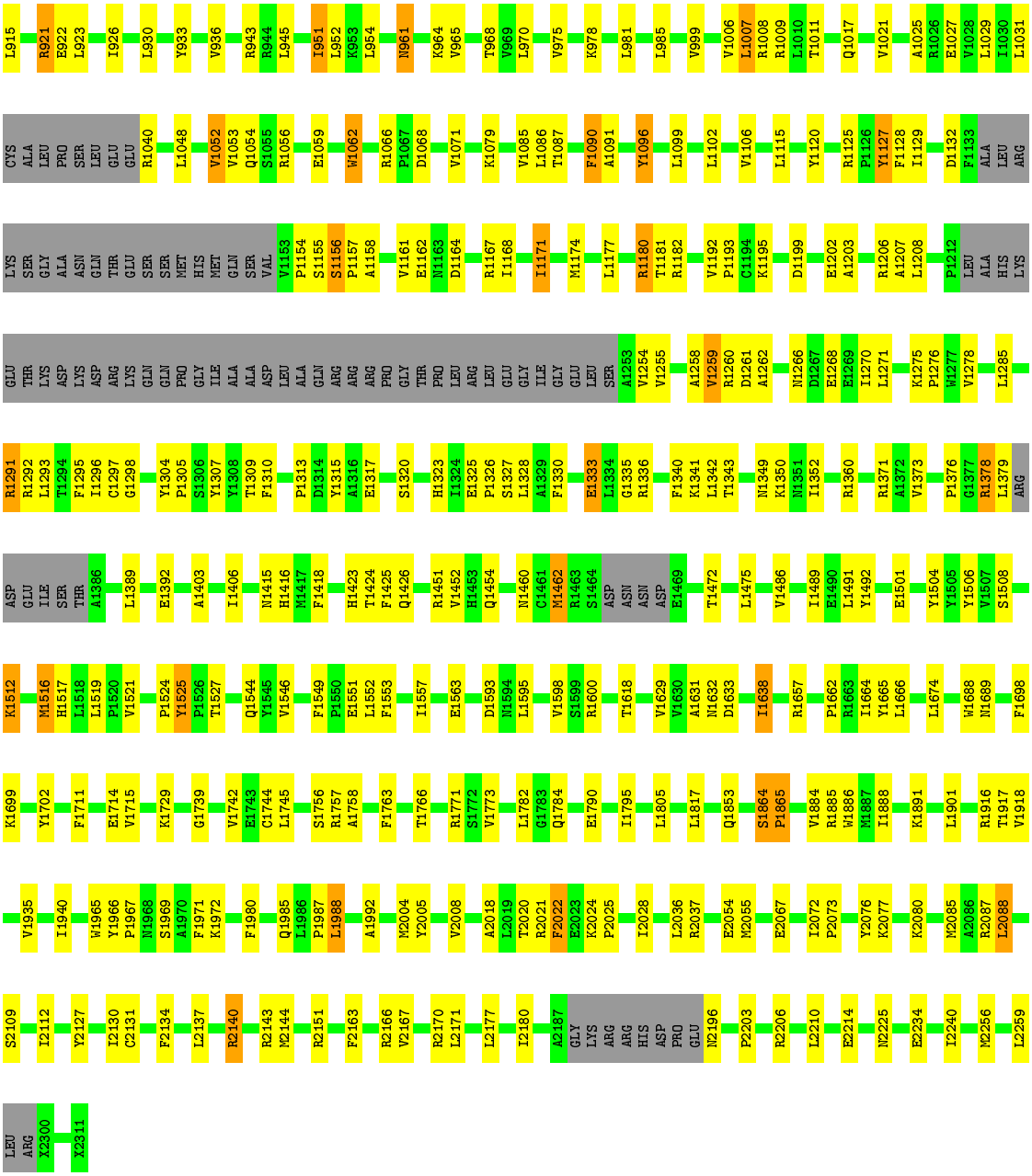
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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. 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: Acetyl-CoA carboxylase-like protein,Acetyl-CoA carboxylase-like protein,Acetyl-CoA carboxylase-like protein

Chain A:

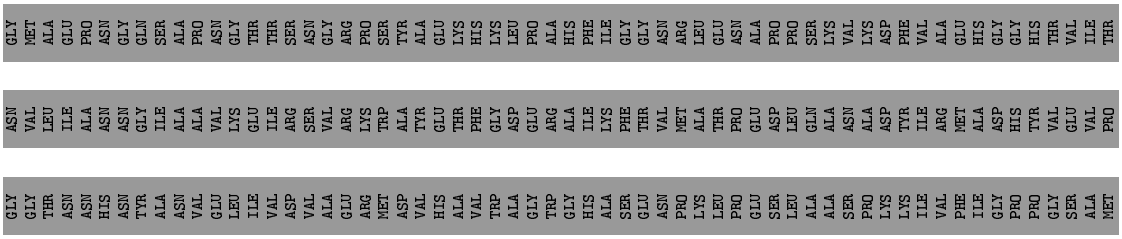


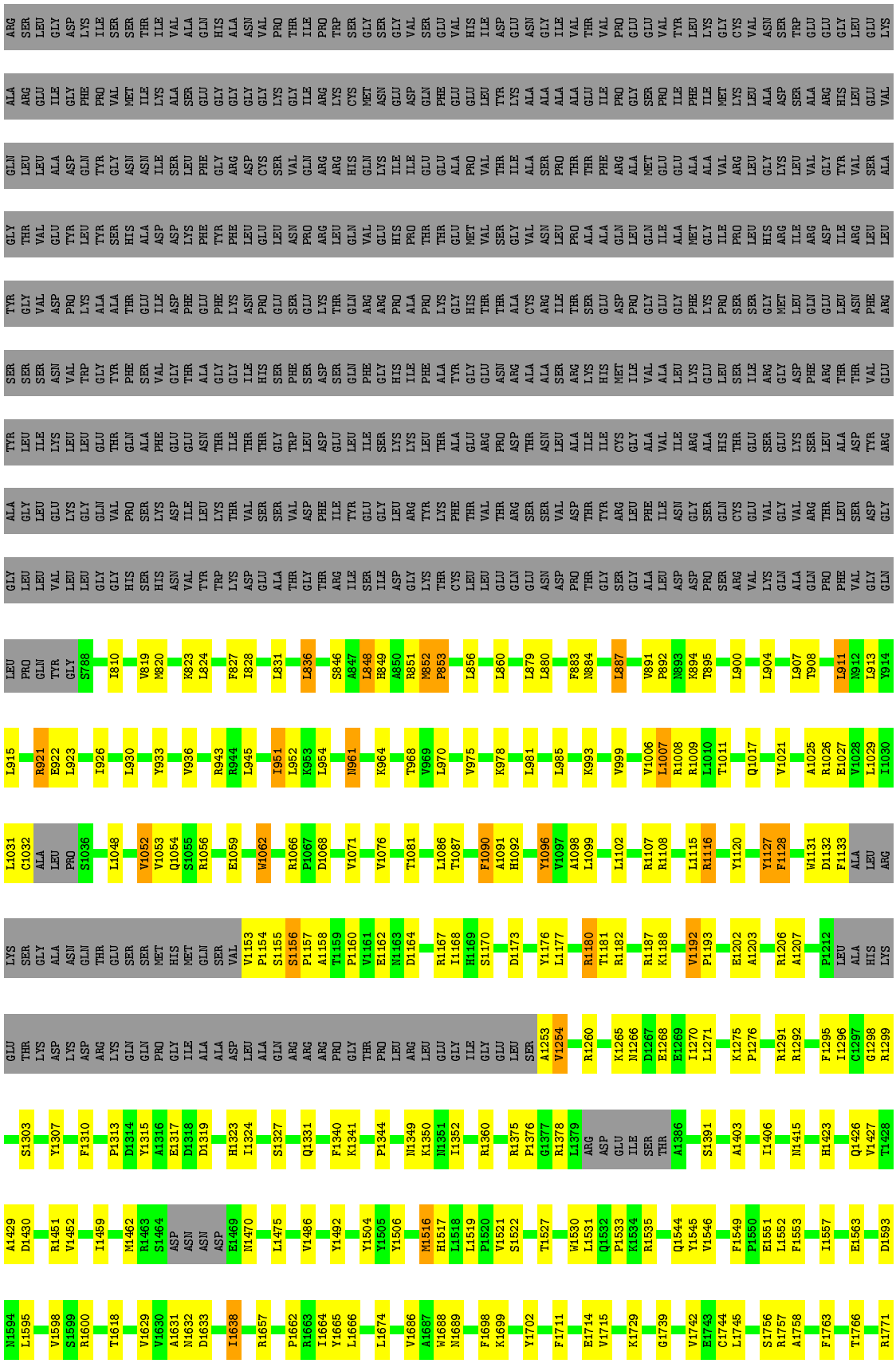


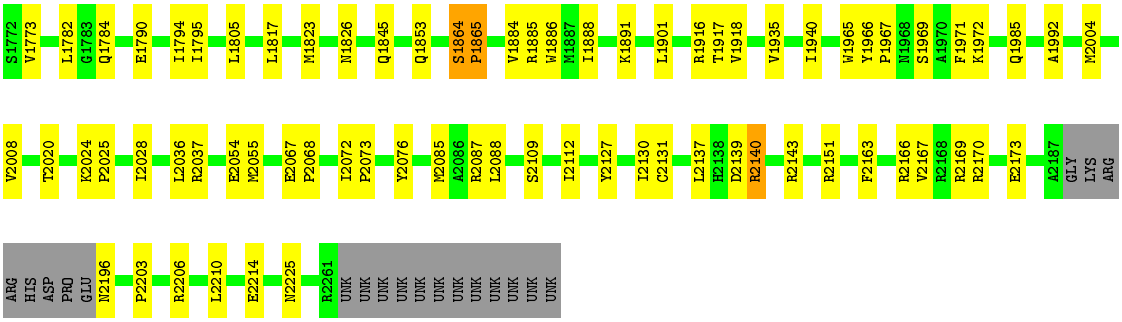


• Molecule 1: Acetyl-CoA carboxylase-like protein,Acetyl-CoA carboxylase-like protein,Acetyl-CoA carboxylase-like protein

Chain B:






S1772
V1773
L1782
G1783
Q1784
E1790
I1794
I1795
L1805
L1817
M1823
N1826
Q1845
Q1853
S1864
P1865
V1884
R1885
W1886
M1887
I1888
K1891
L1901
R1916
T1917
V1918
V1935
I1940
W1965
Y1966
P1967
N1968
S1969
A1970
F1971
K1972
Q1985
A1992
M2004
V2008
T2020
K2024
P2025
I2028
L2036
R2037
E2054
M2055
E2067
P2068
I2072
P2073
Y2076
M2085
A2086
R2087
L2088
S2109
I2112
Y2127
I2130
C2131
L2137
H2138
D2139
R2140
R2143
R2151
F2163
R2166
V2167
R2168
R2169
R2170
E2173
A2187
GLY
LYS
ARG
ARG
HIS
ASP
PRO
GLU
N2196
P2203
R2206
L2210
E2214
N2225
R2261
UNK
UNK
UNK
UNK
UNK
UNK
UNK
UNK
UNK
UNK
UNK
UNK

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, α, β, γ	462.20Å 462.20Å 204.64Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.95 – 8.40 49.95 – 8.40	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.95-8.40) 99.8 (49.95-8.40)	Depositor EDS
R_{merge}	0.29	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 8.33Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.297 , 0.324 0.340 , 0.360	Depositor DCC
R_{free} test set	549 reflections (4.53%)	wwPDB-VP
Wilson B-factor (Å^2)	572.4	Xtriage
Anisotropy	0.424	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 899.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.36$, $< L^2 > = 0.18$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.76	EDS
Total number of atoms	22445	wwPDB-VP
Average B, all atoms (Å^2)	250.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.05% 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	0.41	0/11401	0.64	0/15435
1	B	0.41	0/11458	0.63	0/15511
All	All	0.41	0/22859	0.63	0/30946

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	11224	0	11148	178	0
1	B	11221	0	11191	174	0
All	All	22445	0	22339	347	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 8.

The worst 5 of 347 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:1192:VAL:HG21	1:A:1203:ALA:HB1	1.48	0.94
1:B:1108:ARG:HH21	1:B:1375:ARG:HH22	1.21	0.85

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1864:SER:HB2	1:B:1865:PRO:HD3	1.65	0.79
1:B:2054:GLU:HG3	1:B:2203:PRO:HG2	1.65	0.79
1:A:1864:SER:HB2	1:A:1865:PRO:HD3	1.65	0.78

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	1373/2211 (62%)	1229 (90%)	123 (9%)	21 (2%)	10	46
1	B	1380/2211 (62%)	1249 (90%)	110 (8%)	21 (2%)	10	46
All	All	2753/4422 (62%)	2478 (90%)	233 (8%)	42 (2%)	10	46

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	943	ARG
1	A	1156	SER
1	A	1516	MET
1	A	1784	GLN
1	A	1864	SER

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	1208/1879 (64%)	1138 (94%)	70 (6%)	20	45
1	B	1215/1879 (65%)	1151 (95%)	64 (5%)	22	47
All	All	2423/3758 (64%)	2289 (94%)	134 (6%)	21	47

5 of 134 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2022	PHE
1	B	883	PHE
1	B	1965	TRP
1	A	2088	LEU
1	A	2234	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1415	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 carbohydrates 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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.