wwPDB X-ray Structure Validation Summary Report (i)

Jun 22, 2024 – 12:02 PM EDT

PDB ID	:	5I6H
Title	:	Crystal structure of CD-CT domains of Chaetomium thermophilum acetyl-CoA carboxylase
Authors	:	Hunkeler, M.; Stuttfeld, E.; Hagmann, A.; Imseng, S.; Maier, T.
Deposited on	:	2016-02-16
Resolution	:	7.20 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

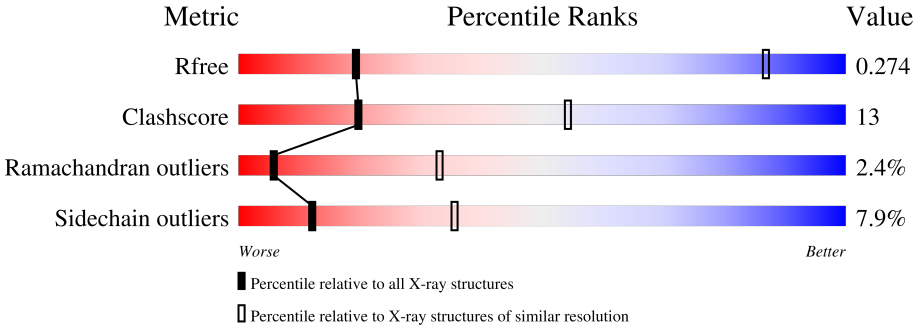
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 7.20 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	1004 (10.00-3.90)
Clashscore	141614	1069 (10.00-3.90)
Ramachandran outliers	138981	1002 (10.00-3.90)
Sidechain outliers	138945	1002 (10.00-3.86)

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 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%

Mol	Chain	Length	Quality of chain
1	A	1487	63% 28% • 6%
1	B	1487	64% 27% • 5%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 22543 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.

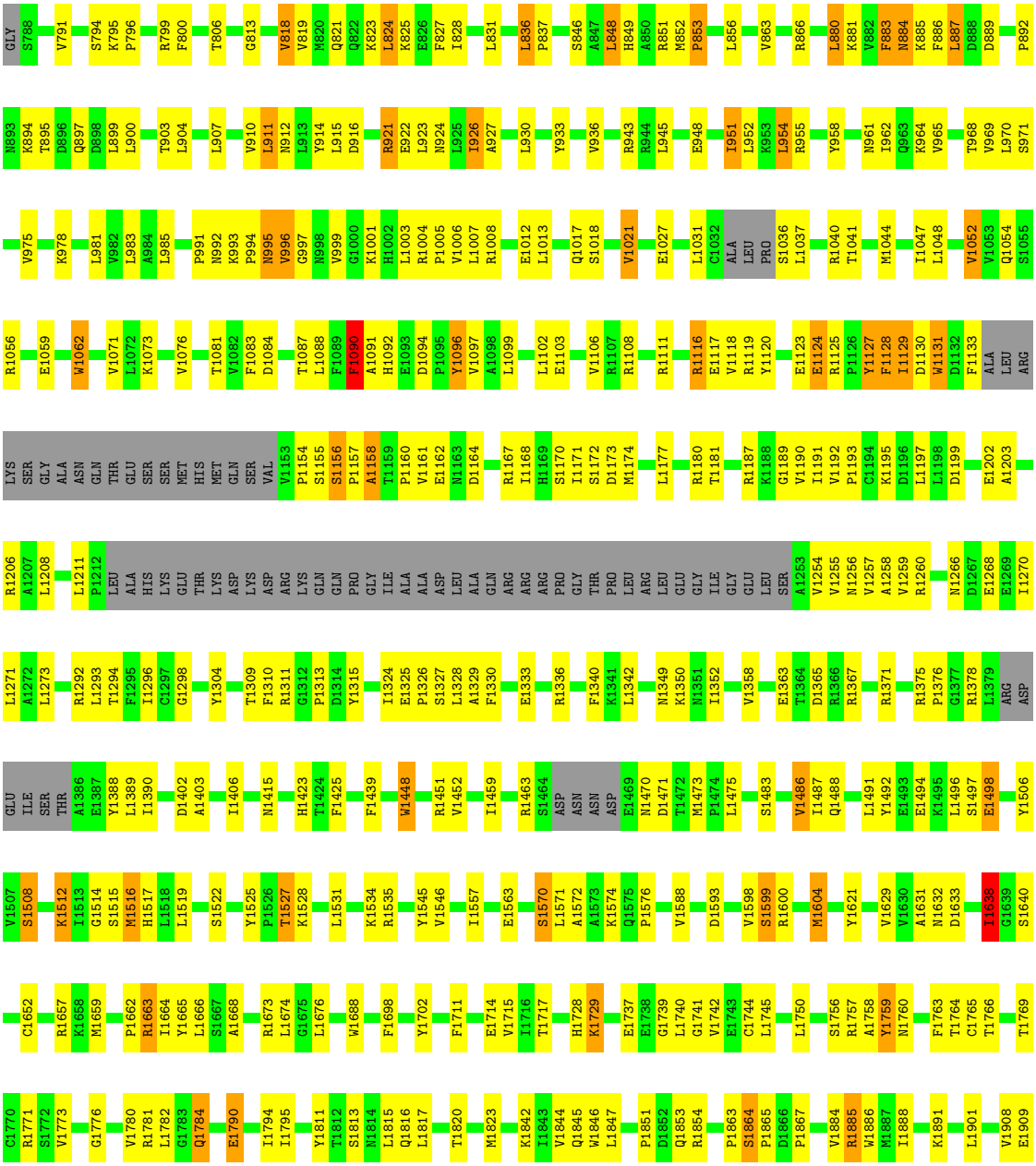
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1404	11262	7144	1983	2096	39	0	0	0
1	B	1406	11281	7156	1988	2098	39	0	0	0

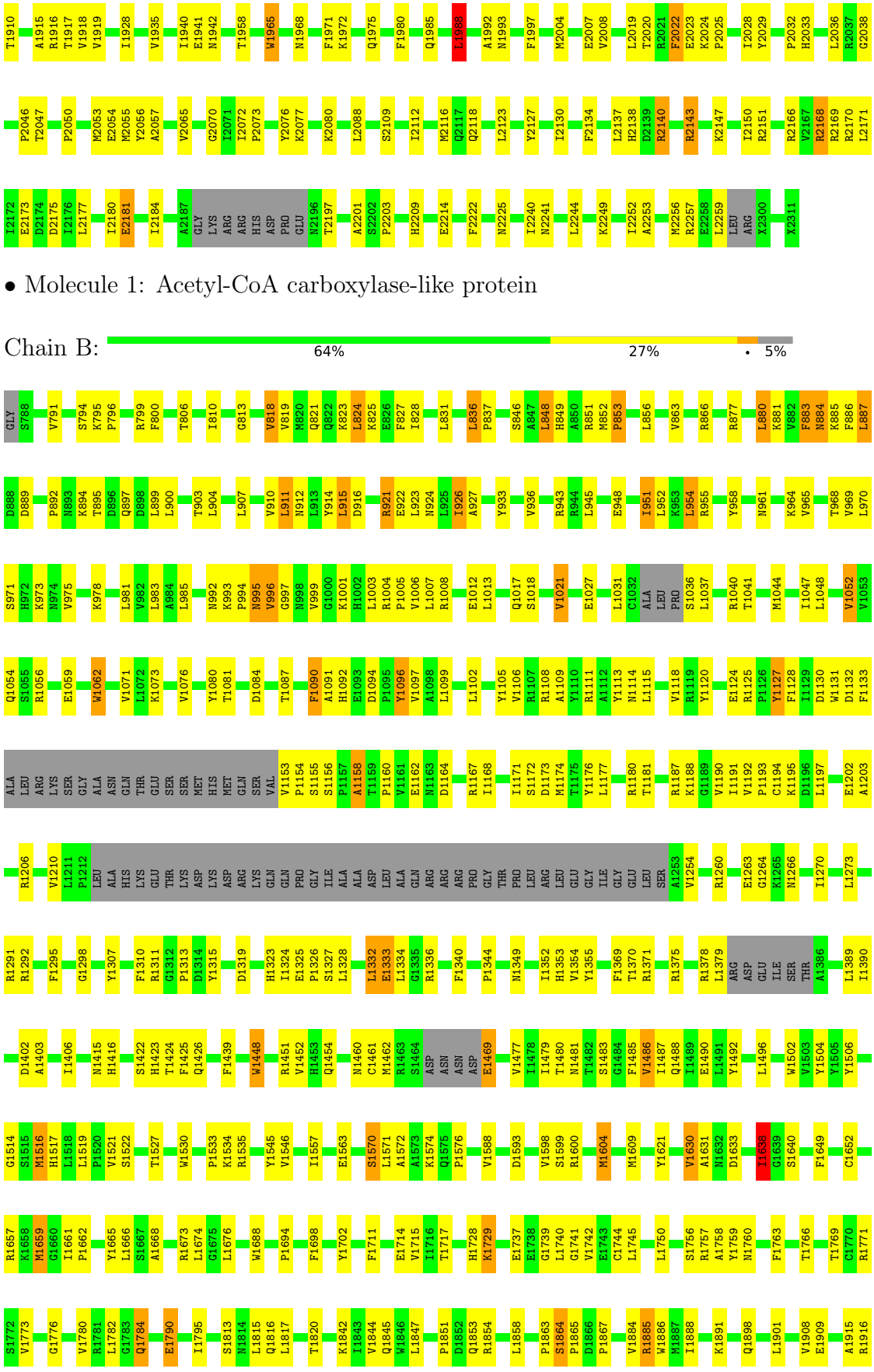
3 Residue-property plots (i)

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. 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

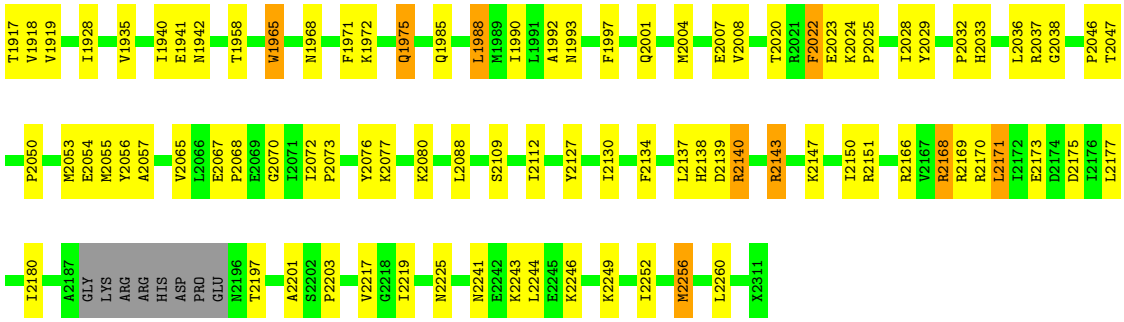
Chain A: 63% 28% • 6%





• Molecule 1: Acetyl-CoA carboxylase-like protein

Chain B: 64% 27% • 5%


T1917
V1918
V1919
I1928
V1935
I1940
E1941
N1942
T1958
W1965
N1968
F1971
K1972
Q1975
Q1985
L1988
M1989
I1990
L1991
A1992
N1993
F1997
Q2001
M2004
E2007
V2008
T2020
R2021
F2022
E2023
K2024
P2025
I2028
Y2029
P2032
H2033
L2036
R2037
G2038
P2046
T2047
P2050
M2053
E2054
M2055
Y2056
A2057
V2065
L2066
E2067
P2068
E2069
G2070
I2071
I2072
P2073
Y2076
K2077
K2080
L2088
S2109
I2112
Y2127
I2130
F2134
L2137
H2138
D2139
R2140
R2143
K2147
I2150
R2151
R2166
V2167
R2168
R2169
R2170
L2171
I2172
E2173
D2174
D2175
I2176
L2177
I2180
A2187
GLY
LYS
ARG
ARG
HIS
ASP
PRO
GLU
N2196
T2197
A2201
S2202
P2203
V2217
G2218
I2219
N2225
N2241
E2242
K2243
L2244
E2245
K2246
K2249
I2252
M2256
L2260
X2311

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	295.02Å 295.02Å 189.52Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.85 – 7.20 49.85 – 7.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.85-7.20) 100.0 (49.85-7.20)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.72 (at 7.37Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.232 , 0.248 0.263 , 0.274	Depositor DCC
R_{free} test set	696 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	662.8	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 410.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.021 for -h,-k,l	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	22543	wwPDB-VP
Average B, all atoms (Å^2)	271.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.82% 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 (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	0.49	0/11439	0.72	3/15486 (0.0%)
1	B	0.50	0/11458	0.71	0/15511
All	All	0.49	0/22897	0.72	3/30997 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1988	LEU	CB-CG-CD2	5.41	120.20	111.00
1	A	1988	LEU	CB-CG-CD1	5.17	119.78	111.00
1	A	1090	PHE	C-N-CA	5.09	134.41	121.70

There are no chirality outliers.

There are no planarity outliers.

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	11262	0	11181	289	0
1	B	11281	0	11205	295	0
All	All	22543	0	22386	571	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 13.

The worst 5 of 571 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:1389:LEU:HD11	1:B:1425:PHE:HB3	1.33	1.08
1:A:994:PRO:HA	1:A:1037:LEU:HB3	1.35	1.07
1:B:994:PRO:HA	1:B:1037:LEU:HB3	1.40	1.01
1:A:1090:PHE:HB2	1:A:1193:PRO:HB3	1.39	0.99
1:B:1176:TYR:HE2	1:B:1291:ARG:HD3	1.23	0.98

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	1378/1487 (93%)	1221 (89%)	121 (9%)	36 (3%)	5 31
1	B	1380/1487 (93%)	1221 (88%)	130 (9%)	29 (2%)	7 36
All	All	2758/2974 (93%)	2442 (88%)	251 (9%)	65 (2%)	6 33

5 of 65 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	819	VAL
1	A	943	ARG
1	A	1092	HIS
1	A	1117	GLU
1	A	1128	PHE

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	1213/1282 (95%)	1116 (92%)	97 (8%)	12	35
1	B	1215/1282 (95%)	1119 (92%)	96 (8%)	12	35
All	All	2428/2564 (95%)	2235 (92%)	193 (8%)	12	35

5 of 193 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	992	ASN
1	B	1483	SER
1	B	1012	GLU
1	B	1127	TYR
1	B	1574	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	995	ASN
1	B	1043	GLN
1	B	1460	ASN

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 monosaccharides 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	2261:ARG	C	2300:UNK	N	9.73

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

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

6.4 Ligands (i)

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

6.5 Other polymers (i)

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