



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 13, 2020 – 09:46 PM BST

PDB ID : 5N0L
Title : The structure of the cofactor binding GAF domain of the nutrient sensor CodY from *Clostridium difficile*
Authors : Levdikov, V.M.; Blagova, E.V.; Wilkinson, A.J.; Sonenshein, A.L.
Deposited on : 2017-02-03
Resolution : 1.68 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.14.4.dev1
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) : 2.14.4.dev1

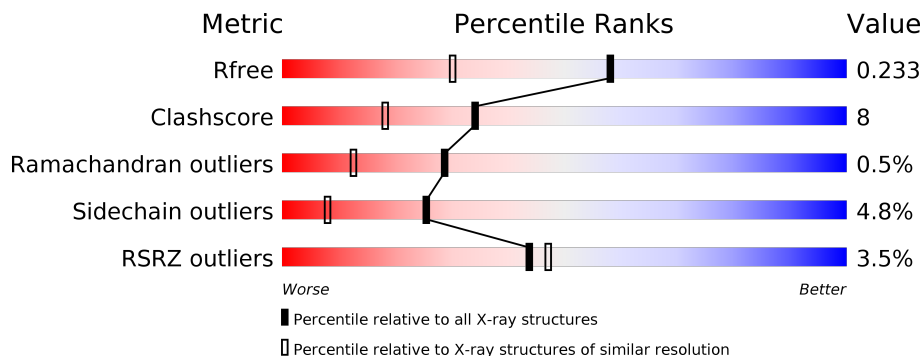
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.68 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	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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 $\leq 5\%$. 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	159	 4% 75% 18% 5% ..
1	B	159	 6% 74% 19% ..
1	C	159	 % 81% 14% ...
1	D	159	 3% 82% 13% ..
1	E	159	 6% 80% 13% ..
1	F	159	 2% 77% 17% ...

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8169 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 GTP-sensing transcriptional pleiotropic repressor CodY.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	156	1240	777	203	259	1	0	7	0
1	B	156	1247	782	204	260	1	0	8	0
1	C	156	1253	785	205	262	1	0	9	0
1	D	156	1247	782	204	260	1	0	8	0
1	E	156	1240	777	203	259	1	0	7	0
1	F	156	1240	777	203	259	1	0	7	0

There are 18 discrepancies between the modelled and reference sequences:

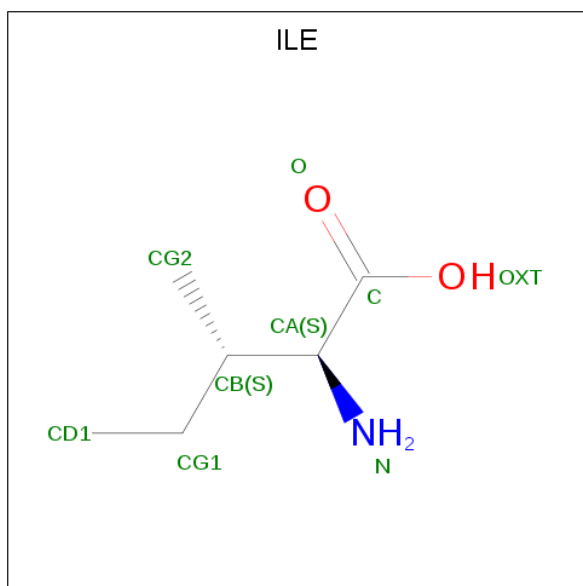
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP A0A031WD25
A	-1	PRO	-	expression tag	UNP A0A031WD25
A	0	ALA	-	expression tag	UNP A0A031WD25
B	-2	GLY	-	expression tag	UNP A0A031WD25
B	-1	PRO	-	expression tag	UNP A0A031WD25
B	0	ALA	-	expression tag	UNP A0A031WD25
C	-2	GLY	-	expression tag	UNP A0A031WD25
C	-1	PRO	-	expression tag	UNP A0A031WD25
C	0	ALA	-	expression tag	UNP A0A031WD25
D	-2	GLY	-	expression tag	UNP A0A031WD25
D	-1	PRO	-	expression tag	UNP A0A031WD25
D	0	ALA	-	expression tag	UNP A0A031WD25
E	-2	GLY	-	expression tag	UNP A0A031WD25
E	-1	PRO	-	expression tag	UNP A0A031WD25
E	0	ALA	-	expression tag	UNP A0A031WD25
F	-2	GLY	-	expression tag	UNP A0A031WD25
F	-1	PRO	-	expression tag	UNP A0A031WD25

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	ALA	-	expression tag	UNP A0A031WD25

- Molecule 2 is ISOLEUCINE (three-letter code: ILE) (formula: C₆H₁₃NO₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	9	6	1	2	0	0
2	B	1	9	6	1	2	0	0
2	C	1	9	6	1	2	0	0
2	D	1	9	6	1	2	0	0
2	E	1	9	6	1	2	0	0
2	F	1	9	6	1	2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	133	133	133	0	0
3	B	101	101	101	0	0
3	C	98	98	98	0	0

Continued on next page...

Continued from previous page...

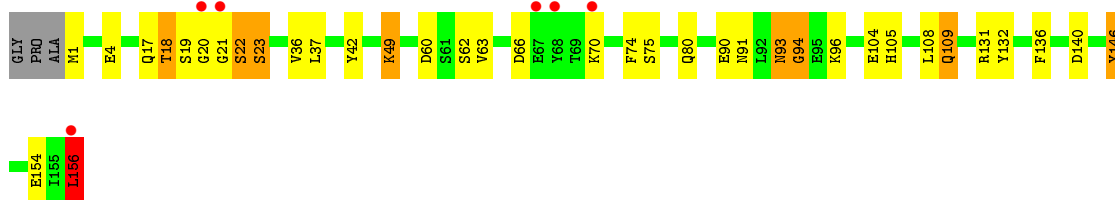
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	91	Total O 91 91	0	0
3	E	94	Total O 94 94	0	0
3	F	131	Total O 131 131	0	0

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 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.

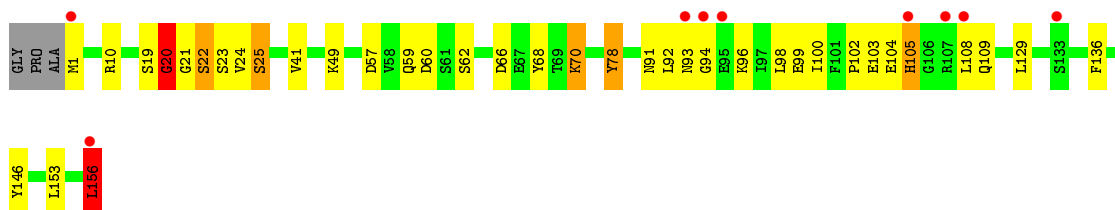
- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain A: 




- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain B: 




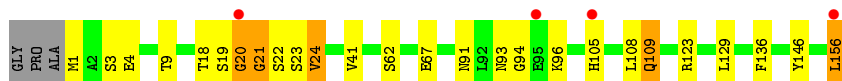
- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

Chain C: 

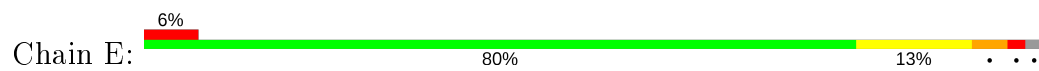


- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY

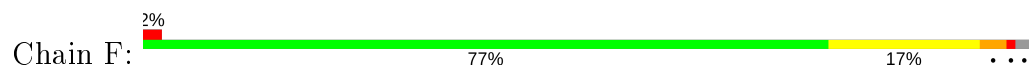
Chain D: 



- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY



- Molecule 1: GTP-sensing transcriptional pleiotropic repressor CodY



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	120.89Å 190.39Å 43.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	56.19 – 1.68 56.19 – 1.68	Depositor EDS
% Data completeness (in resolution range)	100.0 (56.19-1.68) 100.0 (56.19-1.68)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.52 (at 1.68Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.154 , 0.228 0.164 , 0.233	Depositor DCC
R_{free} test set	5726 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	20.2	Xtrriage
Anisotropy	1.014	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 48.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	8169	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% 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	1.69	22/1259 (1.7%)	1.40	11/1699 (0.6%)
1	B	1.20	4/1264 (0.3%)	1.23	8/1703 (0.5%)
1	C	1.17	2/1272 (0.2%)	1.18	11/1717 (0.6%)
1	D	1.17	4/1266 (0.3%)	1.19	4/1709 (0.2%)
1	E	1.28	5/1259 (0.4%)	1.18	6/1699 (0.4%)
1	F	1.55	13/1259 (1.0%)	1.37	15/1699 (0.9%)
All	All	1.36	50/7579 (0.7%)	1.26	55/10226 (0.5%)

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	1
1	B	0	4
1	C	0	3
1	D	0	5
1	E	0	4
1	F	0	2
All	All	0	19

The worst 5 of 50 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	106	GLY	N-CA	9.91	1.60	1.46
1	A	94	GLY	N-CA	-9.90	1.31	1.46
1	F	101	PHE	C-O	9.69	1.41	1.23
1	A	4	GLU	CD-OE1	9.48	1.36	1.25
1	D	4	GLU	CD-OE2	-9.47	1.15	1.25

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	131	ARG	NE-CZ-NH1	10.00	125.30	120.30
1	A	74	PHE	CB-CG-CD2	-9.63	114.06	120.80
1	D	156	LEU	CA-CB-CG	9.09	136.21	115.30
1	D	21[A]	GLY	C-N-CA	9.06	144.35	121.70
1	D	21[B]	GLY	C-N-CA	9.06	144.35	121.70

There are no chirality outliers.

5 of 19 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	93	ASN	Peptide
1	B	20[A]	GLY	Peptide
1	B	21[A]	GLY	Peptide
1	B	22[B]	SER	Peptide
1	B	93	ASN	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	1240	0	1249	18	0
1	B	1247	0	1259	24	7
1	C	1253	0	1264	21	7
1	D	1247	0	1259	23	0
1	E	1240	0	1250	31	0
1	F	1240	0	1250	17	0
2	A	9	0	10	1	0
2	B	9	0	10	1	0
2	C	9	0	10	1	0
2	D	9	0	10	1	0
2	E	9	0	10	1	0
2	F	9	0	10	1	0
3	A	133	0	0	4	0
3	B	101	0	0	7	1
3	C	98	0	0	7	1
3	D	91	0	0	1	0
3	E	94	0	0	8	0
3	F	131	0	0	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	8169	0	7591	127	8

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 127 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:E:19[B]:SER:O	1:E:22[B]:SER:CB	1.73	1.36
1:D:19[B]:SER:O	1:D:22[B]:SER:HB3	1.28	1.28
1:E:19[B]:SER:O	1:E:22[B]:SER:HB3	1.26	1.26
1:F:22[B]:SER:HB2	3:F:401:HOH:O	1.43	1.16
1:B:23[B]:SER:O	1:C:24[B]:VAL:O	1.63	1.14

The worst 5 of 8 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:68:TYR:OH	1:C:95:GLU:OE2[1_556]	1.57	0.63
1:B:103:GLU:OE2	1:C:105:HIS:ND1[2_556]	1.64	0.56
1:B:105:HIS:CB	1:C:95:GLU:OE1[2_556]	1.79	0.41
1:B:105:HIS:CA	1:C:95:GLU:OE1[2_556]	2.00	0.20
1:B:103:GLU:OE2	1:C:105:HIS:CE1[2_556]	2.07	0.13

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	161/159 (101%)	148 (92%)	11 (7%)	2 (1%)	13 2
1	B	162/159 (102%)	154 (95%)	8 (5%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	163/159 (102%)	155 (95%)	8 (5%)	0	100	100
1	D	162/159 (102%)	151 (93%)	9 (6%)	2 (1%)	13	2
1	E	161/159 (101%)	147 (91%)	8 (5%)	6 (4%)	3	0
1	F	161/159 (101%)	153 (95%)	8 (5%)	0	100	100
All	All	970/954 (102%)	908 (94%)	52 (5%)	10 (1%)	29	3

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	20[A]	GLY
1	A	20[B]	GLY
1	D	21[A]	GLY
1	D	21[B]	GLY
1	E	22[A]	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	143/139 (103%)	135 (94%)	8 (6%)	21	6
1	B	144/139 (104%)	137 (95%)	7 (5%)	25	7
1	C	145/139 (104%)	140 (97%)	5 (3%)	37	15
1	D	144/139 (104%)	138 (96%)	6 (4%)	30	10
1	E	143/139 (103%)	137 (96%)	6 (4%)	30	10
1	F	143/139 (103%)	133 (93%)	10 (7%)	15	3
All	All	862/834 (103%)	820 (95%)	42 (5%)	25	7

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	146	TYR
1	D	109	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	105	HIS
1	C	156	LEU
1	D	96	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	109	GLN
1	D	81	ASN
1	F	91	ASN
1	D	71	GLN
1	D	91	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](#)

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	ILE	E	201	-	5,8,8	0.70	0	4,10,10	1.84	1 (25%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	ILE	C	201	-	5,8,8	0.71	0	4,10,10	0.76	0
2	ILE	A	201	-	5,8,8	0.93	0	4,10,10	1.24	1 (25%)
2	ILE	F	201	-	5,8,8	1.20	1 (20%)	4,10,10	0.98	0
2	ILE	B	201	-	5,8,8	0.44	0	4,10,10	0.81	0
2	ILE	D	201	-	5,8,8	0.33	0	4,10,10	0.22	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ILE	E	201	-	-	0/6/10/10	-
2	ILE	C	201	-	-	0/6/10/10	-
2	ILE	A	201	-	-	1/6/10/10	-
2	ILE	F	201	-	-	0/6/10/10	-
2	ILE	B	201	-	-	0/6/10/10	-
2	ILE	D	201	-	-	0/6/10/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	201	ILE	CA-N	-2.40	1.43	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	201	ILE	CG2-CB-CG1	-3.65	102.53	111.78
2	A	201	ILE	CG2-CB-CG1	-2.28	106.00	111.78

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	201	ILE	CG2-CB-CG1-CD1

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	201	ILE	1	0
2	C	201	ILE	1	0
2	A	201	ILE	1	0
2	F	201	ILE	1	0
2	B	201	ILE	1	0
2	D	201	ILE	1	0

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	2
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	24[A]:VAL	C	25:SER	N	2.16
1	B	24[B]:VAL	C	25:SER	N	1.73
1	A	23[A]:SER	C	24:VAL	N	1.13

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	156/159 (98%)	-0.21	6 (3%) 40 43	20, 27, 60, 87	0
1	B	156/159 (98%)	0.06	9 (5%) 23 24	22, 36, 81, 124	0
1	C	156/159 (98%)	-0.30	2 (1%) 77 80	22, 34, 68, 116	0
1	D	156/159 (98%)	-0.26	4 (2%) 56 58	22, 35, 70, 89	0
1	E	156/159 (98%)	-0.08	9 (5%) 23 24	22, 34, 73, 101	0
1	F	156/159 (98%)	-0.33	3 (1%) 66 70	19, 27, 63, 102	0
All	All	936/954 (98%)	-0.18	33 (3%) 44 47	19, 32, 70, 124	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	105	HIS	8.8
1	B	94	GLY	6.9
1	B	108	LEU	5.3
1	A	156	LEU	4.6
1	B	156	LEU	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 monosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	ILE	B	201	9/9	0.94	0.08	37,41,44,44	0
2	ILE	E	201	9/9	0.95	0.08	34,37,44,45	0
2	ILE	A	201	9/9	0.97	0.07	28,31,33,34	0
2	ILE	D	201	9/9	0.97	0.07	31,36,40,41	0
2	ILE	C	201	9/9	0.98	0.07	29,34,37,38	0
2	ILE	F	201	9/9	0.98	0.07	24,28,31,36	0

6.5 Other polymers [i](#)

There are no such residues in this entry.