



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 10:23 PM GMT

PDB ID : 4GWH
Title : Crystal structure of acyl-CoA thioesterase tesB from Yersinia pestis
Authors : Swarbrick, C.M.D.; Forwood, J.K.
Deposited on : 2012-09-03
Resolution : 2.00 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

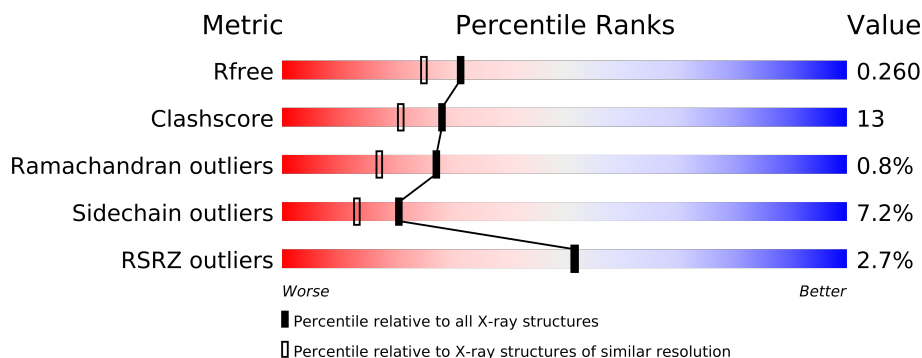
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	288	
1	B	288	
1	C	288	
1	D	288	
1	E	288	
1	F	288	
1	G	288	
1	H	288	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 17725 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Acyl-CoA thioesterase II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	263	Total	C	N	O	S	28	0	0
			2097	1334	367	389	7			
1	B	275	Total	C	N	O	S	11	0	0
			2199	1398	387	407	7			
1	C	269	Total	C	N	O	S	15	0	0
			2154	1371	377	399	7			
1	D	268	Total	C	N	O	S	30	0	0
			2139	1358	376	398	7			
1	E	270	Total	C	N	O	S	11	0	0
			2150	1367	377	399	7			
1	F	265	Total	C	N	O	S	24	0	0
			2112	1342	370	393	7			
1	G	268	Total	C	N	O	S	26	0	0
			2138	1359	375	397	7			
1	H	269	Total	C	N	O	S	13	0	0
			2147	1364	377	399	7			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	287	SER	-	EXPRESSION TAG	UNP Q0WCE2
A	288	ASN	-	EXPRESSION TAG	UNP Q0WCE2
B	287	SER	-	EXPRESSION TAG	UNP Q0WCE2
B	288	ASN	-	EXPRESSION TAG	UNP Q0WCE2
C	287	SER	-	EXPRESSION TAG	UNP Q0WCE2
C	288	ASN	-	EXPRESSION TAG	UNP Q0WCE2
D	287	SER	-	EXPRESSION TAG	UNP Q0WCE2
D	288	ASN	-	EXPRESSION TAG	UNP Q0WCE2
E	287	SER	-	EXPRESSION TAG	UNP Q0WCE2
E	288	ASN	-	EXPRESSION TAG	UNP Q0WCE2
F	287	SER	-	EXPRESSION TAG	UNP Q0WCE2
F	288	ASN	-	EXPRESSION TAG	UNP Q0WCE2
G	287	SER	-	EXPRESSION TAG	UNP Q0WCE2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	288	ASN	-	EXPRESSION TAG	UNP Q0WCE2
H	287	SER	-	EXPRESSION TAG	UNP Q0WCE2
H	288	ASN	-	EXPRESSION TAG	UNP Q0WCE2

- Molecule 2 is water.

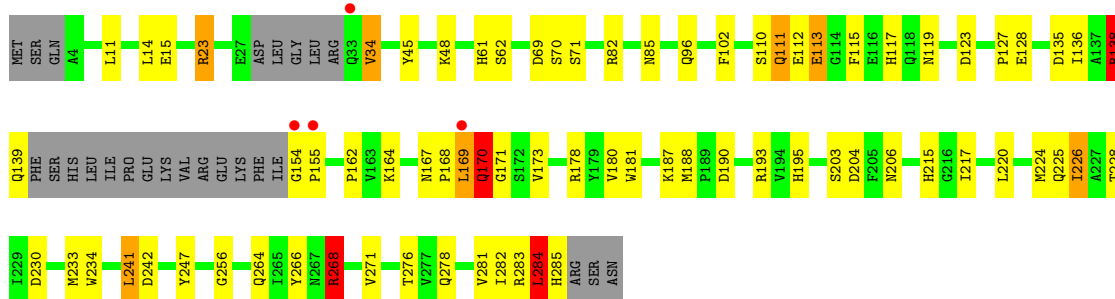
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	68	Total O 68 68	0	0
2	B	82	Total O 82 82	0	0
2	C	67	Total O 67 67	0	0
2	D	89	Total O 89 89	0	0
2	E	74	Total O 74 74	0	0
2	F	59	Total O 59 59	0	0
2	G	80	Total O 80 80	0	0
2	H	70	Total O 70 70	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

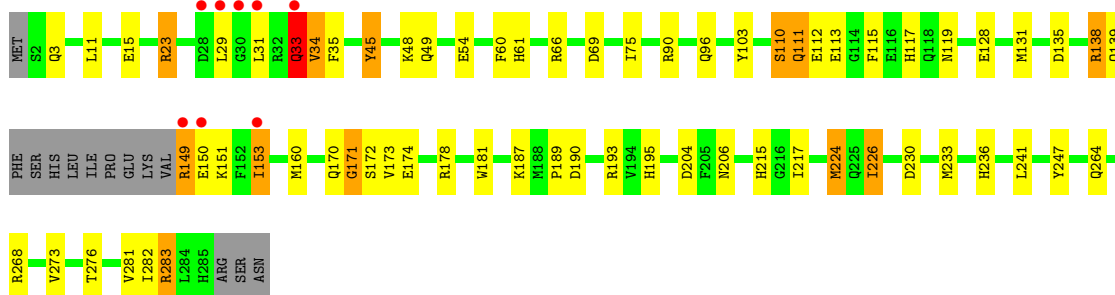
• Molecule 1: Acyl-CoA thioesterase II

Chain A: 



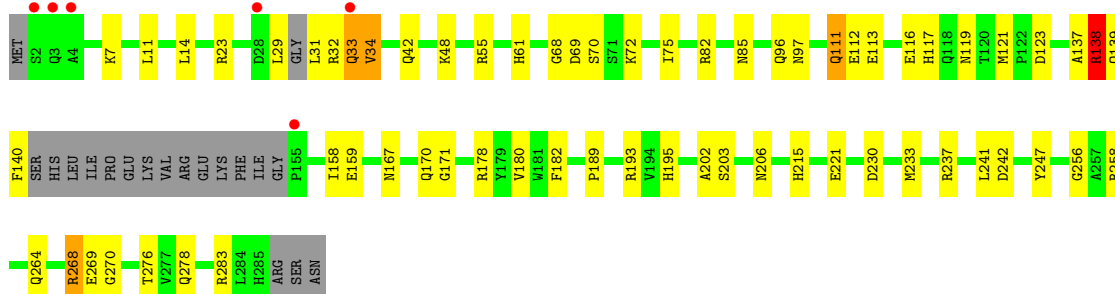
• Molecule 1: Acyl-CoA thioesterase II

Chain B: 



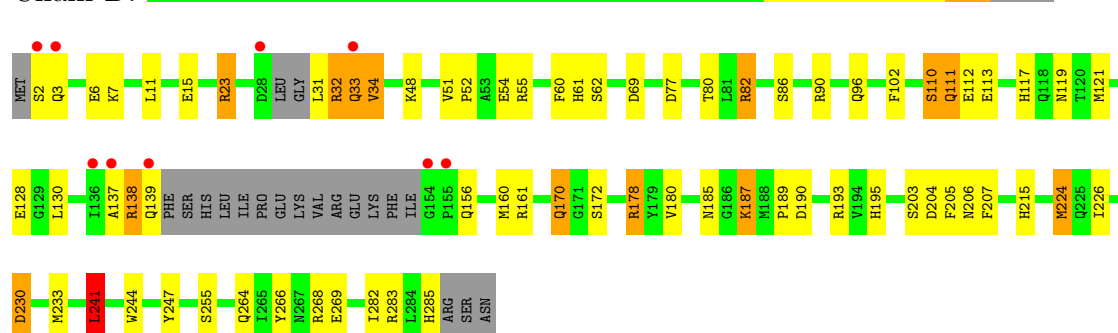
• Molecule 1: Acyl-CoA thioesterase II

Chain C: 



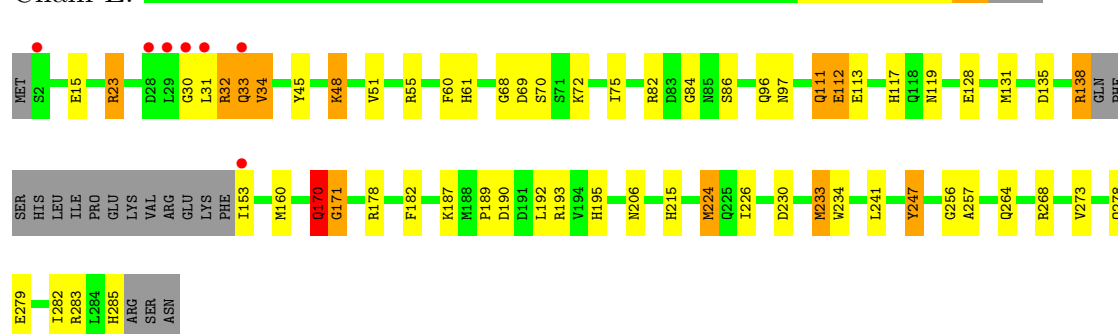
- Molecule 1: Acyl-CoA thioesterase II

Chain D:



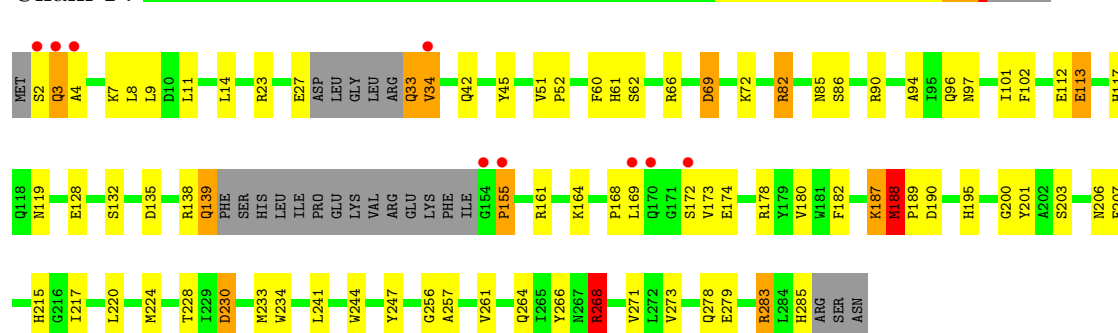
- Molecule 1: Acyl-CoA thioesterase II

Chain E:



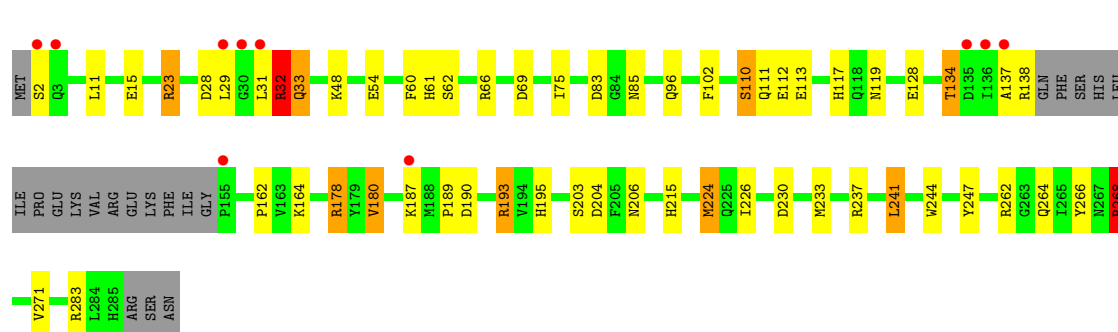
- Molecule 1: Acyl-CoA thioesterase II

Chain F:



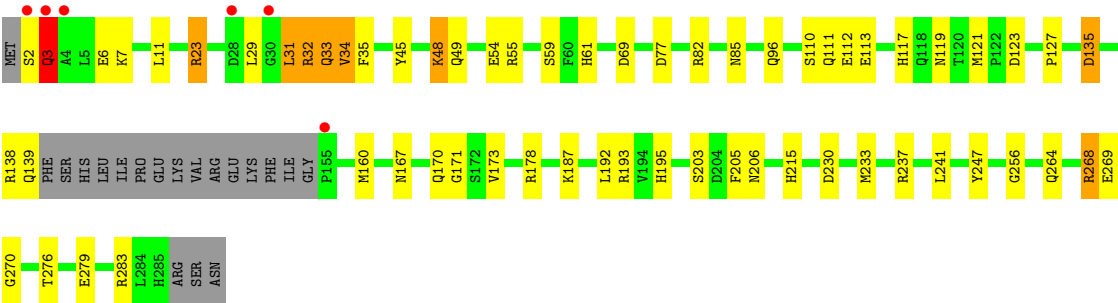
- Molecule 1: Acyl-CoA thioesterase II

Chain G:



● Molecule 1: Acyl-CoA thioesterase II

Chain H: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	73.55Å 170.82Å 101.98Å 90.00° 109.44° 90.00°	Depositor
Resolution (Å)	49.04 – 2.00 49.04 – 2.00	Depositor EDS
% Data completeness (in resolution range)	89.8 (49.04-2.00) 89.7 (49.04-2.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.16 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.210 , 0.261 0.210 , 0.260	Depositor DCC
R_{free} test set	7216 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	19.6	Xtriage
Anisotropy	0.061	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.43 , 35.0	EDS
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 143004 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	17725	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 85.51 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.2323e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

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	1.36	8/2152 (0.4%)	1.18	16/2915 (0.5%)
1	B	1.16	5/2256 (0.2%)	1.19	12/3054 (0.4%)
1	C	1.18	4/2210 (0.2%)	1.16	18/2992 (0.6%)
1	D	1.21	8/2194 (0.4%)	1.27	22/2971 (0.7%)
1	E	1.26	5/2206 (0.2%)	1.24	12/2989 (0.4%)
1	F	1.22	9/2167 (0.4%)	1.22	15/2935 (0.5%)
1	G	1.33	5/2194 (0.2%)	1.16	16/2972 (0.5%)
1	H	1.13	3/2203 (0.1%)	1.16	15/2984 (0.5%)
All	All	1.23	47/17582 (0.3%)	1.20	126/23812 (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	1
1	D	0	2
All	All	0	4

The worst 5 of 47 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	155	PRO	CA-CB	-24.92	1.03	1.53
1	G	31	LEU	CA-CB	-22.53	1.01	1.53
1	G	138	ARG	CA-CB	-20.76	1.08	1.53
1	F	33	GLN	CA-CB	-20.55	1.08	1.53
1	A	155	PRO	N-CD	-19.27	1.20	1.47

The worst 5 of 126 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	33	GLN	N-CA-CB	15.28	138.10	110.60
1	E	23	ARG	NE-CZ-NH2	-14.30	113.15	120.30
1	E	23	ARG	NE-CZ-NH1	14.24	127.42	120.30
1	B	23	ARG	NE-CZ-NH1	14.20	127.40	120.30
1	B	23	ARG	NE-CZ-NH2	-13.14	113.73	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	138	ARG	Peptide
1	B	33	GLN	Peptide
1	D	32	ARG	Mainchain,Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2097	0	2024	64	1
1	B	2199	0	2132	71	0
1	C	2154	0	2083	60	0
1	D	2139	0	2065	71	0
1	E	2150	0	2083	48	0
1	F	2112	0	2037	62	0
1	G	2138	0	2070	45	0
1	H	2147	0	2078	55	1
2	A	68	0	0	2	0
2	B	82	0	0	3	0
2	C	67	0	0	0	0
2	D	89	0	0	0	0
2	E	74	0	0	0	0
2	F	59	0	0	1	0
2	G	80	0	0	0	0
2	H	70	0	0	0	0
All	All	17725	0	16572	418	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 13.

The worst 5 of 418 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:149:ARG:CD	1:B:153:ILE:HD11	1.48	1.40
1:B:149:ARG:CD	1:B:153:ILE:CD1	2.17	1.21
1:B:149:ARG:HD2	1:B:153:ILE:CD1	1.70	1.20
1:E:112:GLU:OE2	1:F:283:ARG:NH2	1.80	1.14
1:H:138:ARG:O	1:H:139:GLN:HG3	1.49	1.12

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	Distance(Å)	Clash(Å)
1:A:85:ASN:OD1	1:H:85:ASN:OD1[2_456]	1.91	0.29

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	257/288 (89%)	247 (96%)	8 (3%)	2 (1%)	27	17
1	B	271/288 (94%)	261 (96%)	7 (3%)	3 (1%)	21	10
1	C	263/288 (91%)	254 (97%)	8 (3%)	1 (0%)	43	36
1	D	262/288 (91%)	250 (95%)	9 (3%)	3 (1%)	21	10
1	E	266/288 (92%)	254 (96%)	9 (3%)	3 (1%)	21	10
1	F	259/288 (90%)	247 (95%)	12 (5%)	0	100	100
1	G	264/288 (92%)	252 (96%)	10 (4%)	2 (1%)	27	17
1	H	265/288 (92%)	254 (96%)	9 (3%)	2 (1%)	27	17
All	All	2107/2304 (91%)	2019 (96%)	72 (3%)	16 (1%)	27	17

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	34	VAL
1	G	33	GLN
1	B	33	GLN
1	B	171	GLY
1	D	33	GLN

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	224/248 (90%)	206 (92%)	18 (8%)	17	10
1	B	235/248 (95%)	222 (94%)	13 (6%)	30	23
1	C	231/248 (93%)	219 (95%)	12 (5%)	32	25
1	D	229/248 (92%)	214 (93%)	15 (7%)	24	16
1	E	230/248 (93%)	214 (93%)	16 (7%)	21	14
1	F	226/248 (91%)	205 (91%)	21 (9%)	13	7
1	G	229/248 (92%)	212 (93%)	17 (7%)	20	12
1	H	230/248 (93%)	210 (91%)	20 (9%)	15	8
All	All	1834/1984 (92%)	1702 (93%)	132 (7%)	21	13

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	33	GLN
1	E	273	VAL
1	H	54	GLU
1	E	70	SER
1	E	138	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	195	HIS
1	E	96	GLN
1	H	111	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	206	ASN
1	D	264	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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(Å²)	Q<0.9	
1	A	263/288 (91%)	-0.20	4 (1%)	70	70	12, 23, 48, 62	6 (2%)
1	B	275/288 (95%)	-0.32	8 (2%)	49	49	9, 17, 39, 76	2 (0%)
1	C	269/288 (93%)	-0.22	6 (2%)	59	59	10, 23, 49, 62	3 (1%)
1	D	268/288 (93%)	-0.21	9 (3%)	43	43	10, 20, 49, 71	6 (2%)
1	E	270/288 (93%)	-0.36	7 (2%)	53	53	9, 18, 39, 62	2 (0%)
1	F	265/288 (92%)	-0.08	9 (3%)	43	43	12, 24, 50, 64	5 (1%)
1	G	268/288 (93%)	-0.18	10 (3%)	39	39	9, 20, 50, 75	5 (1%)
1	H	269/288 (93%)	-0.20	6 (2%)	59	59	11, 22, 49, 60	3 (1%)
All	All	2147/2304 (93%)	-0.22	59 (2%)	52	52	9, 21, 49, 76	32 (1%)

The worst 5 of 59 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	31	LEU	6.9
1	D	154	GLY	6.5
1	E	153	ILE	5.3
1	B	153	ILE	4.8
1	D	137	ALA	4.5

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 carbohydrates 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.