wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 07:43 AM GMT

PDB ID	:	3BWO
Title	:	L-tryptophan aminotransferase
Authors	:	Ferrer, J.-L.; Noel, J.P.; Pojer, F.; Bowman, M.; Chory, J.; Tao, Y.
Deposited on	:	2008-01-10
Resolution	:	2.40 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
http://wwpdb.org/validation/2016/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:

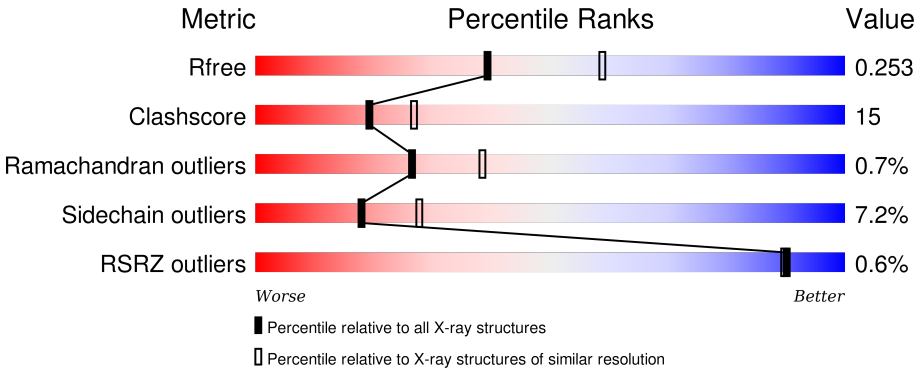
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.40 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.



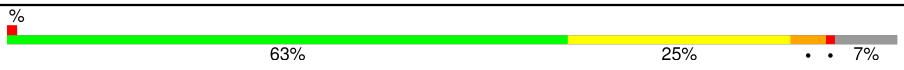
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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. 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% 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	391	67% 21% • • 7%
1	B	391	% 63% 26% • • 7%
1	C	391	% 62% 24% 5% • 8%
1	D	391	% 62% 26% • • 8%
1	E	391	% 65% 23% • 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	391	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18479 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 L-tryptophan aminotransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	363	Total	C	N	O	P	S	0	0	0
			2935	1863	499	553	1	19			
1	B	364	Total	C	N	O	P	S	0	0	0
			2943	1871	499	553	1	19			
1	C	360	Total	C	N	O	P	S	0	0	0
			2910	1849	493	548	1	19			
1	D	360	Total	C	N	O	P	S	0	0	0
			2909	1847	493	549	1	19			
1	E	361	Total	C	N	O	P	S	0	0	0
			2920	1856	495	549	1	19			
1	F	363	Total	C	N	O	P	S	0	0	0
			2935	1867	497	551	1	19			

- Molecule 2 is water.

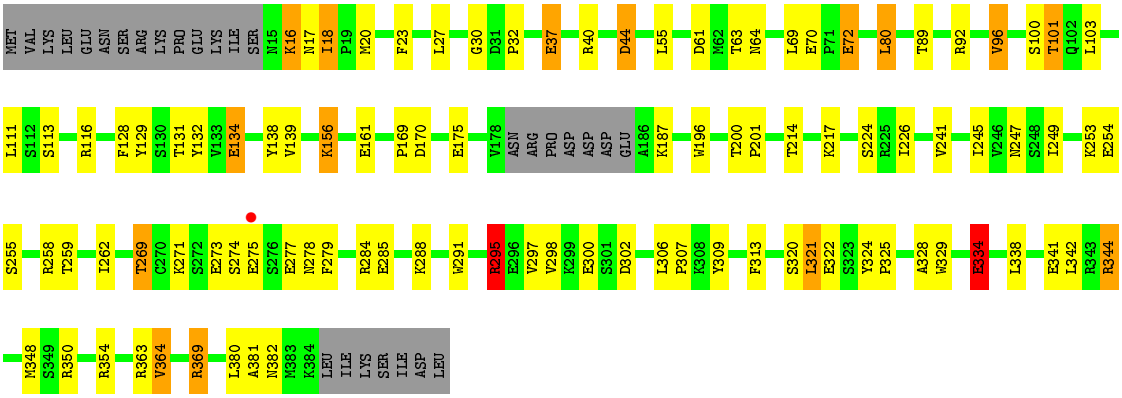
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	161	Total	O	0	0
			161	161		
2	B	184	Total	O	0	0
			184	184		
2	C	132	Total	O	0	0
			132	132		
2	D	159	Total	O	0	0
			159	159		
2	E	135	Total	O	0	0
			135	135		
2	F	156	Total	O	0	0
			156	156		

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 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.

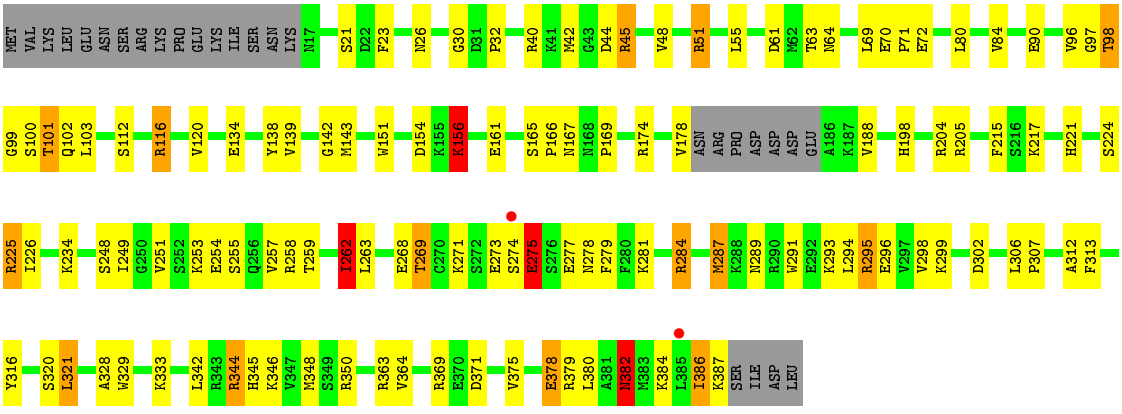
- Molecule 1: L-tryptophan aminotransferase

Chain A:



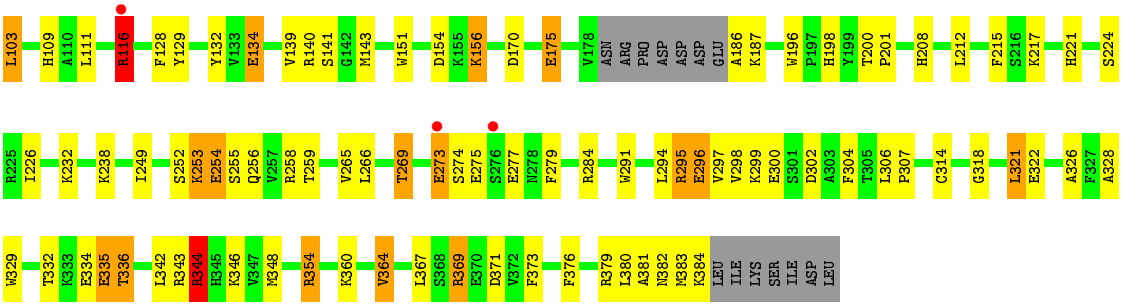
- Molecule 1: L-tryptophan aminotransferase

Chain B:



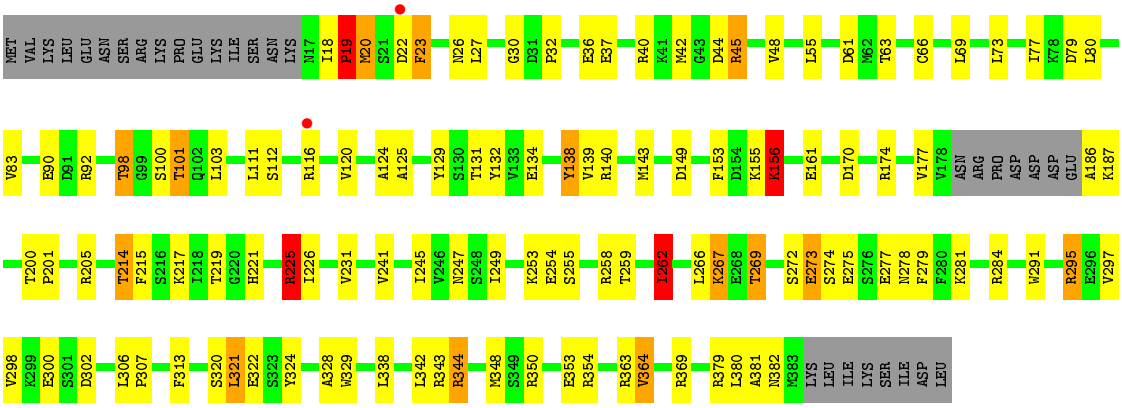
- Molecule 1: L-tryptophan aminotransferase

Chain C:



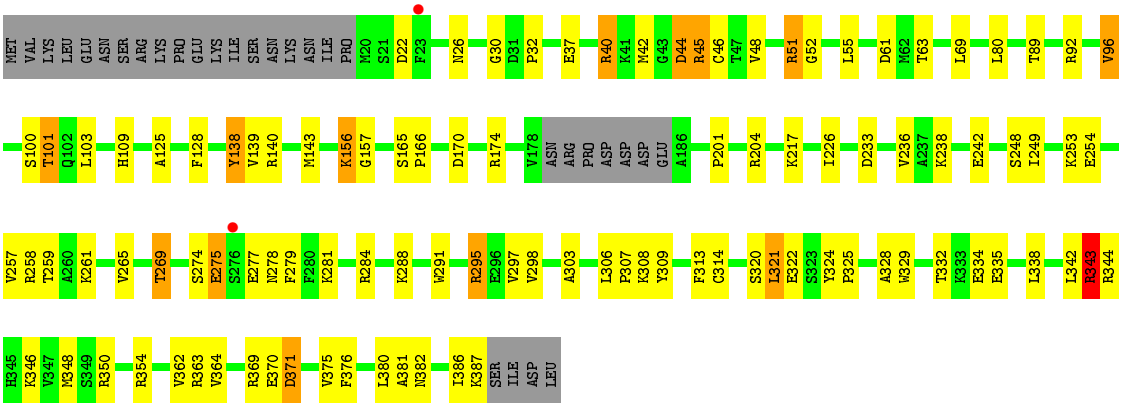
• Molecule 1: L-tryptophan aminotransferase

Chain D:



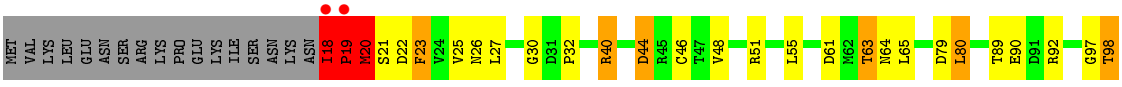
• Molecule 1: L-tryptophan aminotransferase

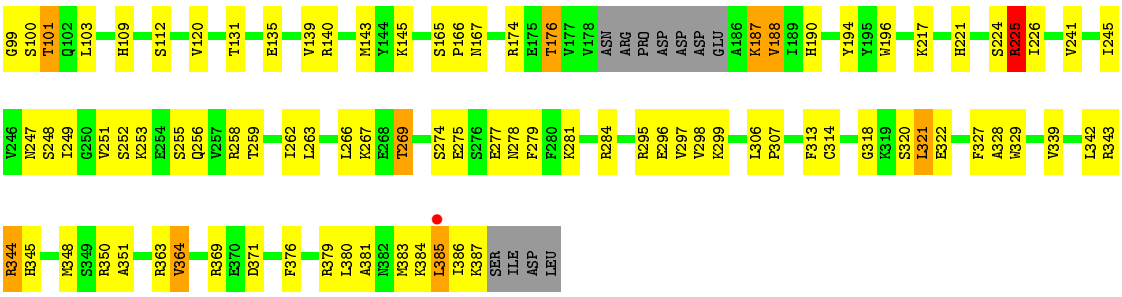
Chain E:



• Molecule 1: L-tryptophan aminotransferase

Chain F:




G99 S100 T101 Q102 L103 H109 S112 V120 T131 E135 V139 R140 M143 Y144 K145 S165 P166 N167 R174 E175 T176 V177 V178 ASN ARG PRO ASP ASP ASP GLU A186 K187 V188 I189 H190 Y194 Y195 W196 K217 H221 S224 R225 I226 V241 I245
V246 M247 S248 I249 G250 V251 S252 K253 E254 S255 Q256 V257 R258 T259 I262 L263 L266 K267 E268 T269 S274 E275 S276 E277 N278 F279 F280 K281 R284 R295 E296 V297 V298 K299 L306 P307 F313 C314 G318 K319 S320 L321 E322 F327 A328 W329 V339 L342 R343
R344 H345 M348 S349 R350 A351 R363 V364 R369 E370 D371 F376 R379 L380 A381 N382 M383 K384 L385 I386 K387 SER ILE ASP LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	93.33Å 98.66Å 139.36Å 90.00° 104.50° 90.00°	Depositor
Resolution (Å)	29.66 – 2.40 29.66 – 2.40	Depositor EDS
% Data completeness (in resolution range)	97.9 (29.66-2.40) 97.9 (29.66-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	3.15 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.238 , 0.253 0.238 , 0.253	Depositor DCC
R_{free} test set	4682 reflections (5.25%)	DCC
Wilson B-factor (Å^2)	29.7	Xtriage
Anisotropy	0.253	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 25.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Outliers	3 of 93815 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	18479	wwPDB-VP
Average B, all atoms (Å^2)	33.0	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 27.82 % 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 2.0544e-03.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: LLP

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	0.81	12/2979 (0.4%)	0.87	13/4025 (0.3%)
1	B	0.80	8/2987 (0.3%)	0.85	8/4036 (0.2%)
1	C	0.78	7/2954 (0.2%)	0.92	12/3992 (0.3%)
1	D	0.78	7/2953 (0.2%)	0.85	11/3992 (0.3%)
1	E	0.76	7/2963 (0.2%)	0.90	10/4002 (0.2%)
1	F	0.88	11/2979 (0.4%)	0.90	15/4025 (0.4%)
All	All	0.80	52/17815 (0.3%)	0.88	69/24072 (0.3%)

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	D	0	1
1	F	0	2
All	All	0	3

The worst 5 of 52 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	344	ARG	NE-CZ	11.99	1.48	1.33
1	F	140	ARG	CZ-NH2	-11.61	1.18	1.33
1	D	273	GLU	CD-OE1	10.22	1.36	1.25
1	F	40	ARG	CD-NE	8.77	1.61	1.46
1	F	19	PRO	N-CD	8.63	1.59	1.47

The worst 5 of 69 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	354	ARG	NE-CZ-NH2	21.37	130.98	120.30
1	E	343	ARG	NE-CZ-NH2	20.19	130.40	120.30
1	E	51	ARG	NE-CZ-NH2	-17.06	111.77	120.30
1	F	140	ARG	NE-CZ-NH2	15.61	128.10	120.30
1	F	344	ARG	NE-CZ-NH1	14.49	127.55	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	19	PRO	Peptide
1	F	18	ILE	Peptide
1	F	19	PRO	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	2935	0	2864	73	0
1	B	2943	0	2880	92	0
1	C	2910	0	2840	89	0
1	D	2909	0	2832	95	0
1	E	2920	0	2857	91	0
1	F	2935	0	2874	119	0
2	A	161	0	0	4	0
2	B	184	0	0	19	0
2	C	132	0	0	7	0
2	D	159	0	0	10	0
2	E	135	0	0	7	0
2	F	156	0	0	14	0
All	All	18479	0	17147	524	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 15.

The worst 5 of 524 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:156:LYS:NZ	1:E:156:LYS:HB2	1.57	1.12
1:B:344:ARG:HH11	1:B:344:ARG:HB2	1.09	1.12
1:F:386:ILE:HG22	1:F:387:LYS:H	1.13	1.07
1:F:385:LEU:HB3	1:F:386:ILE:HD12	1.40	1.02
1:F:20:MET:HG2	1:F:21:SER:N	1.72	1.02

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	358/391 (92%)	337 (94%)	20 (6%)	1 (0%)	46	63
1	B	359/391 (92%)	334 (93%)	20 (6%)	5 (1%)	14	19
1	C	355/391 (91%)	330 (93%)	23 (6%)	2 (1%)	30	43
1	D	355/391 (91%)	332 (94%)	21 (6%)	2 (1%)	30	43
1	E	356/391 (91%)	333 (94%)	21 (6%)	2 (1%)	30	43
1	F	358/391 (92%)	332 (93%)	23 (6%)	3 (1%)	24	35
All	All	2141/2346 (91%)	1998 (93%)	128 (6%)	15 (1%)	26	38

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	386	ILE
1	D	20	MET
1	E	386	ILE
1	F	19	PRO
1	B	248	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	317/345 (92%)	291 (92%)	26 (8%)	14	21
1	B	318/345 (92%)	293 (92%)	25 (8%)	15	23
1	C	314/345 (91%)	285 (91%)	29 (9%)	11	16
1	D	314/345 (91%)	290 (92%)	24 (8%)	16	25
1	E	315/345 (91%)	299 (95%)	16 (5%)	29	46
1	F	317/345 (92%)	300 (95%)	17 (5%)	27	43
All	All	1895/2070 (92%)	1758 (93%)	137 (7%)	18	28

5 of 137 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	134	GLU
1	C	344	ARG
1	F	103	LEU
1	C	175	GLU
1	C	295	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	382	ASN
1	F	289	ASN
1	E	289	ASN
1	B	382	ASN
1	D	382	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)

6 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	LLP	A	217	1	23,24,25	2.40	7 (30%)	28,32,34	1.82	8 (28%)
1	LLP	B	217	1	23,24,25	2.21	4 (17%)	28,32,34	1.49	5 (17%)
1	LLP	C	217	1	23,24,25	2.29	7 (30%)	28,32,34	1.42	6 (21%)
1	LLP	D	217	1	23,24,25	2.29	5 (21%)	28,32,34	1.66	7 (25%)
1	LLP	E	217	1	23,24,25	2.42	8 (34%)	28,32,34	1.47	6 (21%)
1	LLP	F	217	1	23,24,25	2.21	6 (26%)	28,32,34	1.61	6 (21%)

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
1	LLP	A	217	1	-	0/15/17/19	0/1/1/1
1	LLP	B	217	1	-	0/15/17/19	0/1/1/1
1	LLP	C	217	1	-	0/15/17/19	0/1/1/1
1	LLP	D	217	1	-	0/15/17/19	0/1/1/1
1	LLP	E	217	1	-	0/15/17/19	0/1/1/1
1	LLP	F	217	1	-	0/15/17/19	0/1/1/1

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	217	LLP	P-OP2	-2.59	1.45	1.54
1	F	217	LLP	P-OP2	-2.56	1.45	1.54
1	A	217	LLP	P-OP2	-2.42	1.46	1.54
1	C	217	LLP	P-OP2	-2.06	1.47	1.54
1	E	217	LLP	C2-N1	2.02	1.38	1.34

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	217	LLP	C5-C6-N1	-3.68	117.47	123.86
1	F	217	LLP	C4-C4'-NZ	-3.54	105.37	125.06
1	D	217	LLP	C5-C6-N1	-3.24	118.23	123.86
1	E	217	LLP	C4-C4'-NZ	-2.92	108.81	125.06
1	D	217	LLP	C5'-C5-C4	-2.89	116.61	121.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	217	LLP	2	0
1	B	217	LLP	2	0
1	C	217	LLP	2	0
1	D	217	LLP	2	0
1	E	217	LLP	2	0
1	F	217	LLP	3	0

5.5 Carbohydrates (i)

There are no carbohydrates 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

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, 95^{th} 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	362/391 (92%)	-0.58	1 (0%)	94	94	8, 30, 60, 74	0
1	B	363/391 (92%)	-0.51	2 (0%)	90	90	8, 29, 61, 74	0
1	C	359/391 (91%)	-0.49	3 (0%)	87	87	14, 32, 61, 77	0
1	D	359/391 (91%)	-0.56	2 (0%)	90	90	11, 31, 61, 77	0
1	E	360/391 (92%)	-0.49	2 (0%)	90	90	14, 32, 58, 71	0
1	F	362/391 (92%)	-0.56	3 (0%)	87	87	12, 30, 59, 77	0
All	All	2165/2346 (92%)	-0.53	13 (0%)	90	90	8, 31, 61, 77	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	274	SER	3.3
1	F	19	PRO	3.1
1	B	385	LEU	3.1
1	E	276	SER	2.8
1	E	23	PHE	2.8

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} 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	LLDF	B-factors(Å^2)	Q<0.9
1	LLP	C	217	24/25	0.96	0.12	-	23,25,29,32	0
1	LLP	F	217	24/25	0.96	0.13	-	15,23,27,32	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
1	LLP	E	217	24/25	0.97	0.12	-	18,23,26,27	0
1	LLP	D	217	24/25	0.96	0.12	-	19,23,32,37	0
1	LLP	B	217	24/25	0.97	0.12	-	18,21,25,29	0
1	LLP	A	217	24/25	0.96	0.13	-	18,23,25,29	0

6.3 Carbohydrates (i)

There are no carbohydrates 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.