



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 5, 2017 – 05:54 AM EST

PDB ID : 4ECK
Title : Crystal Structure of the Toxoplasma gondii TS-DHFR
Authors : Sharma, H.; Anderson, K.S.
Deposited on : unknown
Resolution : 3.52 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : rb-20030345

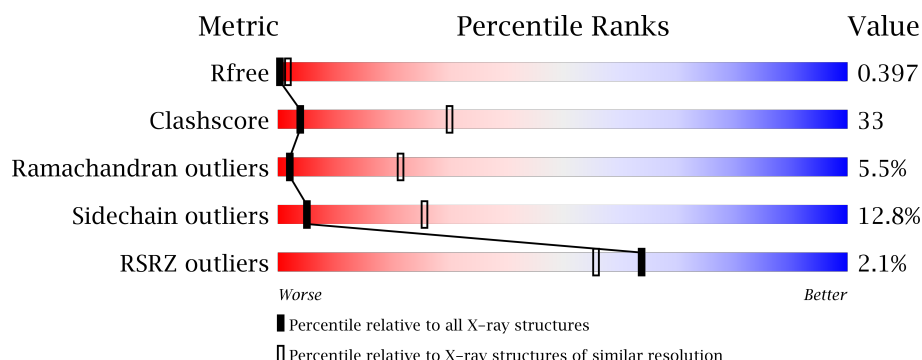
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.52 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}	100719	1239 (3.64-3.40)
Clashscore	112137	1007 (3.62-3.42)
Ramachandran outliers	110173	1328 (3.64-3.40)
Sidechain outliers	110143	1329 (3.64-3.40)
RSRZ outliers	101464	1270 (3.64-3.40)

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 $\leq 5\%$. 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	610	<div> <div>0.1%</div> <div>37%</div> <div>36%</div> <div>8%</div> <div>18%</div> </div>
1	B	610	<div> <div>2%</div> <div>36%</div> <div>37%</div> <div>9%</div> <div>18%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	UMP	A	701	-	-	X	-
2	UMP	B	701	-	-	X	-
3	CB3	A	702	-	-	X	X
3	CB3	B	702	-	-	X	X
4	FOL	A	703	-	-	X	X
4	FOL	B	703	-	-	X	X

2 Entry composition [i](#)

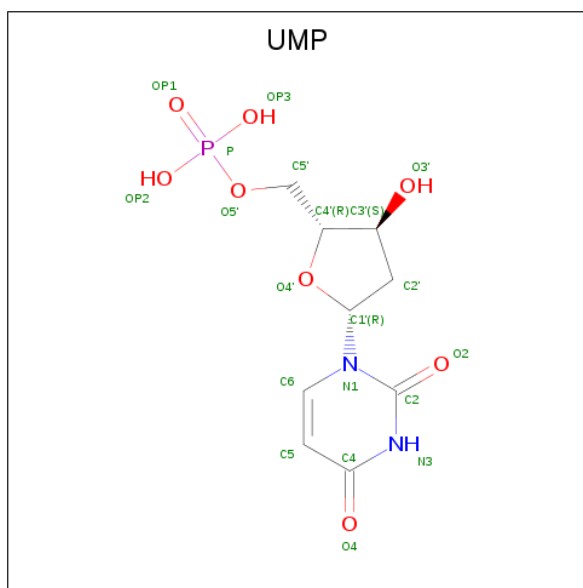
There are 5 unique types of molecules in this entry. The entry contains 8248 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Bifunctional dihydrofolate reductase-thymidylate synthase.

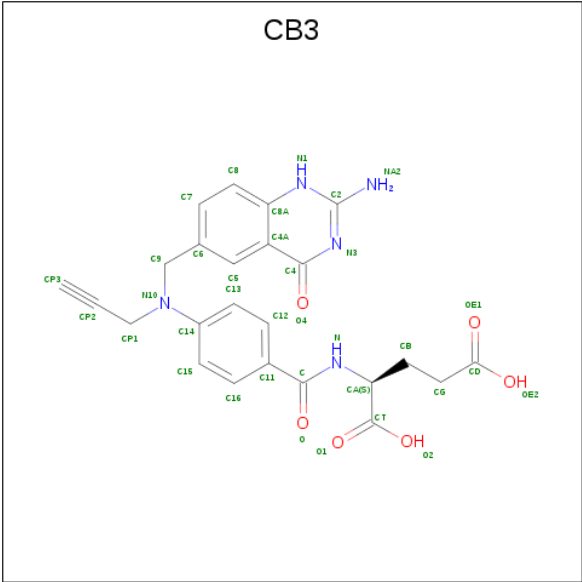
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	498	Total	C	N	O	S	0	0	0
			3989	2550	697	717	25			
1	B	498	Total	C	N	O	S	0	0	0
			3989	2550	697	717	25			

- Molecule 2 is 2'-DEOXYURIDINE 5'-MONOPHOSPHATE (three-letter code: UMP) (formula: $C_9H_{13}N_2O_8P$).



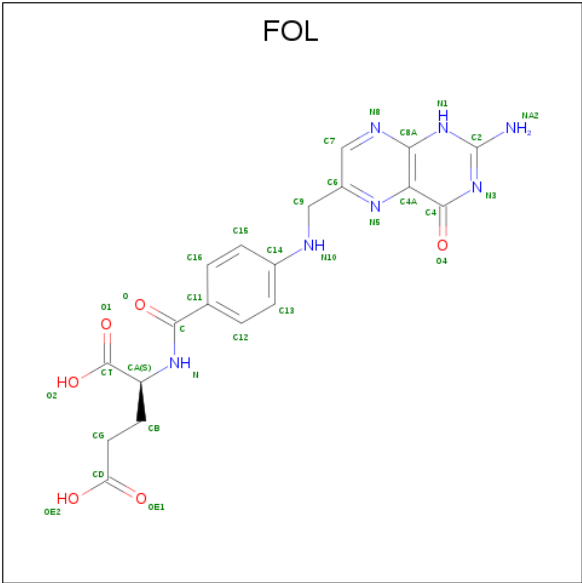
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
2	B	1	Total	C	N	O	P	0	0
			20	9	2	8	1		

- Molecule 3 is 10-PROPARGYL-5,8-DIDEAZAFOLIC ACID (three-letter code: CB3) (formula: $C_{24}H_{23}N_5O_6$).



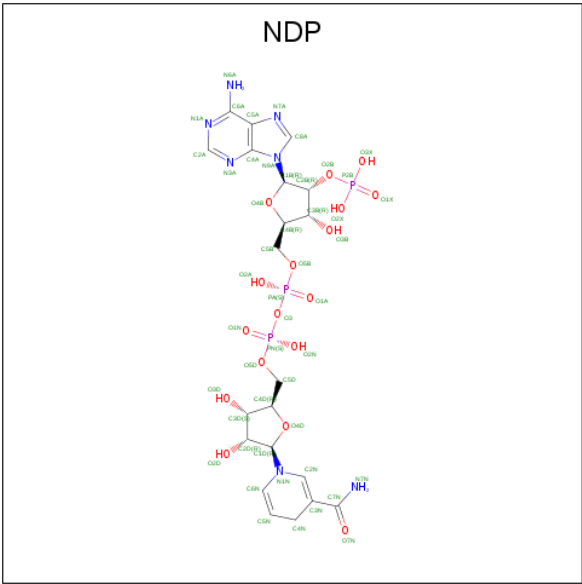
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			35	24	5	6		
3	B	1	Total	C	N	O	0	0
			35	24	5	6		

- Molecule 4 is FOLIC ACID (three-letter code: FOL) (formula: C₁₉H₁₉N₇O₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			32	19	7	6		
4	B	1	Total	C	N	O	0	0
			32	19	7	6		

- Molecule 5 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: C₂₁H₃₀N₇O₁₇P₃).

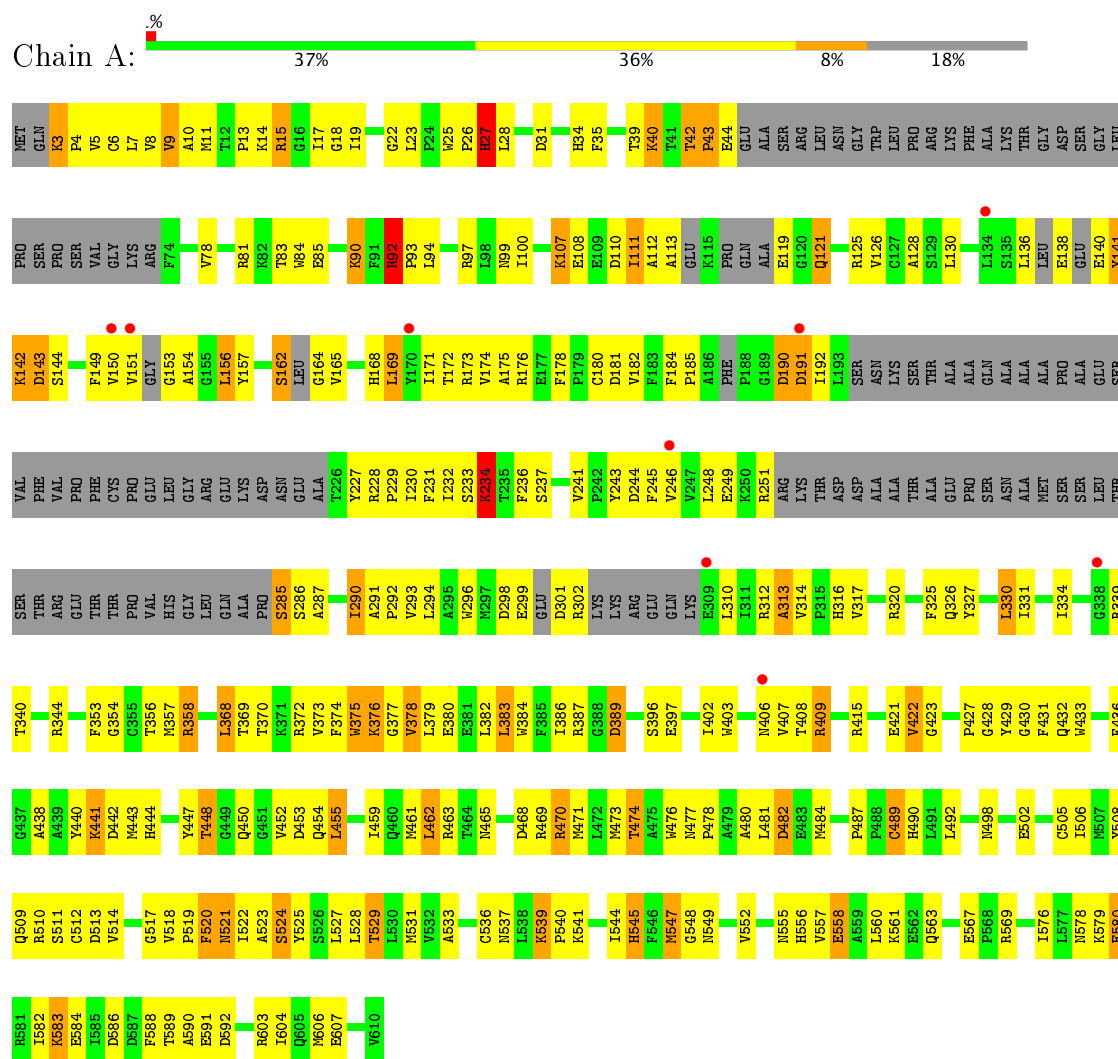


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
5	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

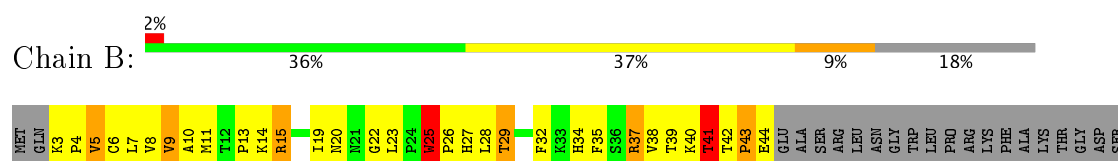
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Bifunctional dihydrofolate reductase-thymidylate synthase



- Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase



R569	N498	R409	I331	SER	GLN	GLY
P570	D499	E421	I334	ASN	ALA	LEU
I573	Q500	V422	I335	ALA	ALA	PRO
N578	C505	G423	R339	MET	ALA	SER
K579	I506	D424	R344	SER	PRO	PRO
I582	M507	P427	T345	LEU	ALA	SER
I585	Q509	N433	I350	THR	GLU	VAL
D586	R510	F436	F353	THR	PHE	GLY
D587	S511	Y440	G354	ARG	VAL	LYS
F588	C512	K441	C355	THR	PRO	ARG
T589	G515	D442	T356	GLU	CYS	F74
A590	G517	N443	R358	THR	PRO	N75
E591	V518	H444	L361	GLN	GLU	A76
D592	P519	T445	P366	ALA	GLY	V77
V596	F520	D446	L367	LYS	GLU	V78
V599	M521	Y447	L368	ASN	ASP	M79
P600	I522	D453	T369	PRO	GLU	G80
H601	A523	Q454	R371	ALA	ARG	G81
G602	S524	L455	R372	ALA	GLU	K82
R603	Y525	K456	T370	T226	LEU	R87
I604	S526	P457	V373	R228	L94	P88
Q605	L528	V458	F374	V229	V95	R89
M606	T529	L462	K375	I230	D96	K90
A609	M531	L465	G376	I232	R97	F91
V610	A533	V532	G377	I233	L98	R92
	H534	H535	V378	S233	M99	F93
	V535	D468	L379	R234	L100	P93
	C536	R469	E380	GLU	V101	V93
	K539	R470	E381	D301	F236	V95
P640	P640	M471	L382	LYS	I171	D96
F543	F543	L472	L383	LYS	T172	R97
I544	I544	M473	W384	ARG	R173	D96
H545	H545	T474	F385	GLU	V174	L98
F546	F546	A475	I386	GLN	D181	M99
M547	M547	M476	R387	LYS	C180	L100
		L481	G388	E309	D244	V101
		M484	D389	A313	F245	V102
		V552	T390	V314	V246	K107
		H556	N391	V317	V247	E108
			A392	H318	A186	E109
			N393	F319	P188	D110
			H394	R320	GLN	I111
			L395	G321	E250	A112
			S396	H322	R251	A113
			E397	E323	ARG	GLU
			I402	F324	D190	K115
			M403	D404	D191	PRO
			R405	K406	D192	GLN
			V496	L330	THR	ALA
					ASP	ALA
					LYS	E119
					THR	G120
					ALA	Q121
					ALA	Q122
					ALA	L193
					ALA	SER
					ALA	R123
					ALA	C127
					ALA	A128
					ALA	S129
					ALA	L130
					ALA	P131

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	185.94Å 143.01Å 59.84Å 90.00° 90.02° 90.00°	Depositor
Resolution (Å)	46.18 – 3.52 46.18 – 3.52	Depositor EDS
% Data completeness (in resolution range)	95.9 (46.18-3.52) 95.9 (46.18-3.52)	Depositor EDS
R_{merge}	0.30	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.81 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.332 , 0.395 0.328 , 0.397	Depositor DCC
R_{free} test set	982 reflections (5.53%)	DCC
Wilson B-factor (Å ²)	70.5	Xtriage
Anisotropy	0.145	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 30.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.387 for -h,-k,l	Xtriage
F_o, F_c correlation	0.77	EDS
Total number of atoms	8248	wwPDB-VP
Average B, all atoms (Å ²)	60.0	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 23.87 % 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 4.2797e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: FOL, CB3, UMP, NDP

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	0.58	1/4078 (0.0%)	0.73	0/5510
1	B	0.59	3/4078 (0.1%)	0.73	0/5510
All	All	0.59	4/8156 (0.0%)	0.73	0/11020

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	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	403	TRP	CD2-CE2	5.34	1.47	1.41
1	B	25	TRP	CD2-CE2	5.26	1.47	1.41
1	B	433	TRP	CD2-CE2	5.02	1.47	1.41
1	A	403	TRP	CD2-CE2	5.01	1.47	1.41

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	42	THR	Peptide
1	A	92	ARG	Peptide

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	3989	0	3939	267	0
1	B	3989	0	3940	295	0
2	A	20	0	11	25	0
2	B	20	0	11	15	0
3	A	35	0	21	15	0
3	B	35	0	21	18	0
4	A	32	0	17	10	0
4	B	32	0	17	26	0
5	A	48	0	26	8	0
5	B	48	0	26	16	0
All	All	8248	0	8029	543	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 33.

The worst 5 of 543 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:402:ILE:HG22	3:B:702:CB3:C16	1.09	1.56
1:B:402:ILE:CG2	3:B:702:CB3:C16	1.91	1.46
1:B:510:ARG:HH12	2:B:701:UMP:P	1.40	1.41
1:A:489:CYS:SG	2:A:701:UMP:C6	2.28	1.25
1:A:344:ARG:NH2	2:A:701:UMP:OP3	1.72	1.21

There are no symmetry-related clashes.

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	474/610 (78%)	379 (80%)	68 (14%)	27 (6%)	2	21
1	B	474/610 (78%)	385 (81%)	64 (14%)	25 (5%)	2	22
All	All	948/1220 (78%)	764 (81%)	132 (14%)	52 (6%)	2	21

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	43	PRO
1	A	92	ARG
1	A	128	ALA
1	A	165	VAL
1	A	313	ALA

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	433/525 (82%)	378 (87%)	55 (13%)	5	26
1	B	433/525 (82%)	377 (87%)	56 (13%)	5	25
All	All	866/1050 (82%)	755 (87%)	111 (13%)	5	26

5 of 111 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	547	MET
1	B	82	LYS
1	B	543	PHE
1	A	558	GLU
1	B	9	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	27	HIS
1	B	99	ASN
1	B	490	HIS
1	A	545	HIS
1	B	444	HIS

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 ⓘ

8 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	UMP	A	701	-	17,21,21	0.76	1 (5%)	23,31,31	2.10	3 (13%)
3	CB3	A	702	-	30,37,37	1.37	3 (10%)	39,51,51	2.60	8 (20%)
4	FOL	A	703	-	27,34,34	1.04	2 (7%)	35,47,47	1.99	8 (22%)
5	NDP	A	704	-	43,52,52	2.02	8 (18%)	49,80,80	2.54	5 (10%)
2	UMP	B	701	-	17,21,21	0.78	1 (5%)	23,31,31	2.15	4 (17%)
3	CB3	B	702	-	30,37,37	1.36	3 (10%)	39,51,51	2.50	10 (25%)
4	FOL	B	703	-	27,34,34	1.11	3 (11%)	35,47,47	2.22	9 (25%)
5	NDP	B	704	-	43,52,52	2.02	8 (18%)	49,80,80	2.47	6 (12%)

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	UMP	A	701	-	-	0/6/22/22	0/2/2/2
3	CB3	A	702	-	-	0/21/28/28	0/3/3/3
4	FOL	A	703	-	-	0/16/22/22	0/3/3/3
5	NDP	A	704	-	-	0/30/77/77	0/5/5/5
2	UMP	B	701	-	-	0/6/22/22	0/2/2/2
3	CB3	B	702	-	-	0/21/28/28	0/3/3/3
4	FOL	B	703	-	-	0/16/22/22	0/3/3/3
5	NDP	B	704	-	-	0/30/77/77	0/5/5/5

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	704	NDP	C3B-C2B	-4.10	1.43	1.53
5	A	704	NDP	C3B-C2B	-3.96	1.44	1.53
5	A	704	NDP	C2D-C3D	-3.38	1.44	1.53
5	B	704	NDP	C2D-C3D	-3.26	1.44	1.53
2	B	701	UMP	C2-N3	-2.31	1.33	1.38

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	702	CB3	C4A-C4-N3	-10.38	117.13	124.45
3	A	702	CB3	C4A-C4-N3	-10.29	117.19	124.45
3	A	702	CB3	C4A-C8A-N1	-9.24	118.40	123.67
5	A	704	NDP	N3A-C2A-N1A	-8.62	121.35	128.86
5	B	704	NDP	N3A-C2A-N1A	-8.34	121.60	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 132 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	UMP	25	0
3	A	702	CB3	15	0
4	A	703	FOL	10	0
5	A	704	NDP	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	701	UMP	15	0
3	B	702	CB3	18	0
4	B	703	FOL	26	0
5	B	704	NDP	16	0

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 [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	498/610 (81%)	0.08	9 (1%) 69 60	39, 57, 83, 115	2 (0%)
1	B	498/610 (81%)	0.05	12 (2%) 59 50	39, 58, 79, 103	2 (0%)
All	All	996/1220 (81%)	0.07	21 (2%) 64 56	39, 57, 81, 115	4 (0%)

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	150	VAL	3.7
1	A	151	VAL	3.0
1	A	338	GLY	2.9
1	A	309	GLU	2.9
1	A	406	ASN	2.9

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 carbohydrates 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. 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, 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	LLDF	B-factors(\AA^2)	Q<0.9
3	CB3	B	702	35/35	0.74	0.53	1.59	72,92,110,116	0
3	CB3	A	702	35/35	0.74	0.48	1.52	80,90,105,105	0
4	FOL	B	703	32/32	0.78	0.49	1.28	72,82,105,110	0
4	FOL	A	703	32/32	0.77	0.47	1.23	72,86,94,98	0
5	NDP	B	704	48/48	0.86	0.32	0.44	75,84,99,102	0
2	UMP	B	701	20/20	0.93	0.28	0.23	68,79,103,105	0
5	NDP	A	704	48/48	0.88	0.31	0.19	75,83,100,104	0
2	UMP	A	701	20/20	0.94	0.25	-0.22	71,78,95,97	0

6.5 Other polymers [i](#)

There are no such residues in this entry.