wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:05 PM GMT

PDB ID	:	1NDI
Title	:	Carnitine Acetyltransferase in complex with CoA
Authors	:	Jogl, G.; Tong, L.
Deposited on	:	2002-12-09
Resolution	:	2.30 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:

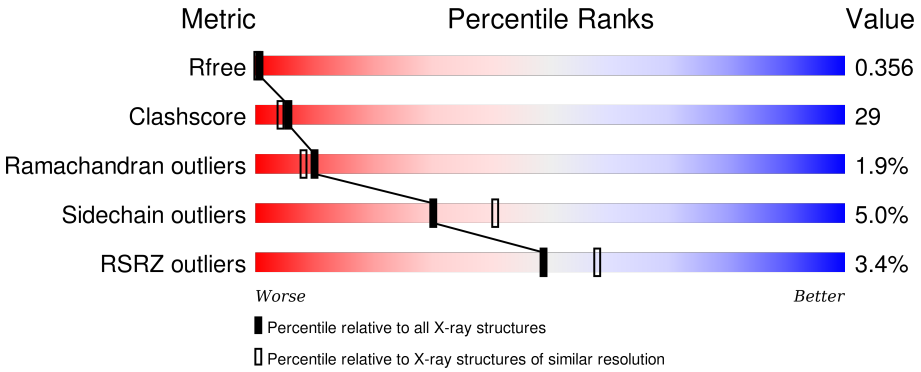
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	596	4% 49% 46%
1	B	596	2% 51% 46%

2 Entry composition

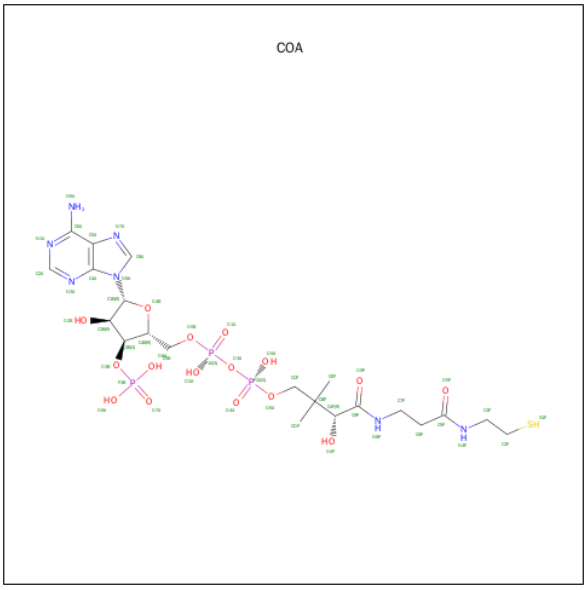
There are 3 unique types of molecules in this entry. The entry contains 10303 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Carnitine Acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	596	Total	C	N	O	S	0	0	0
			4757	3034	820	875	28			
1	B	596	Total	C	N	O	S	0	0	0
			4757	3034	820	875	28			

- Molecule 2 is COENZYME A (three-letter code: COA) (formula: $C_{21}H_{36}N_7O_{16}P_3S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		
2	B	1	Total	C	N	O	P	S	0	0
			48	21	7	16	3	1		

- Molecule 3 is water.

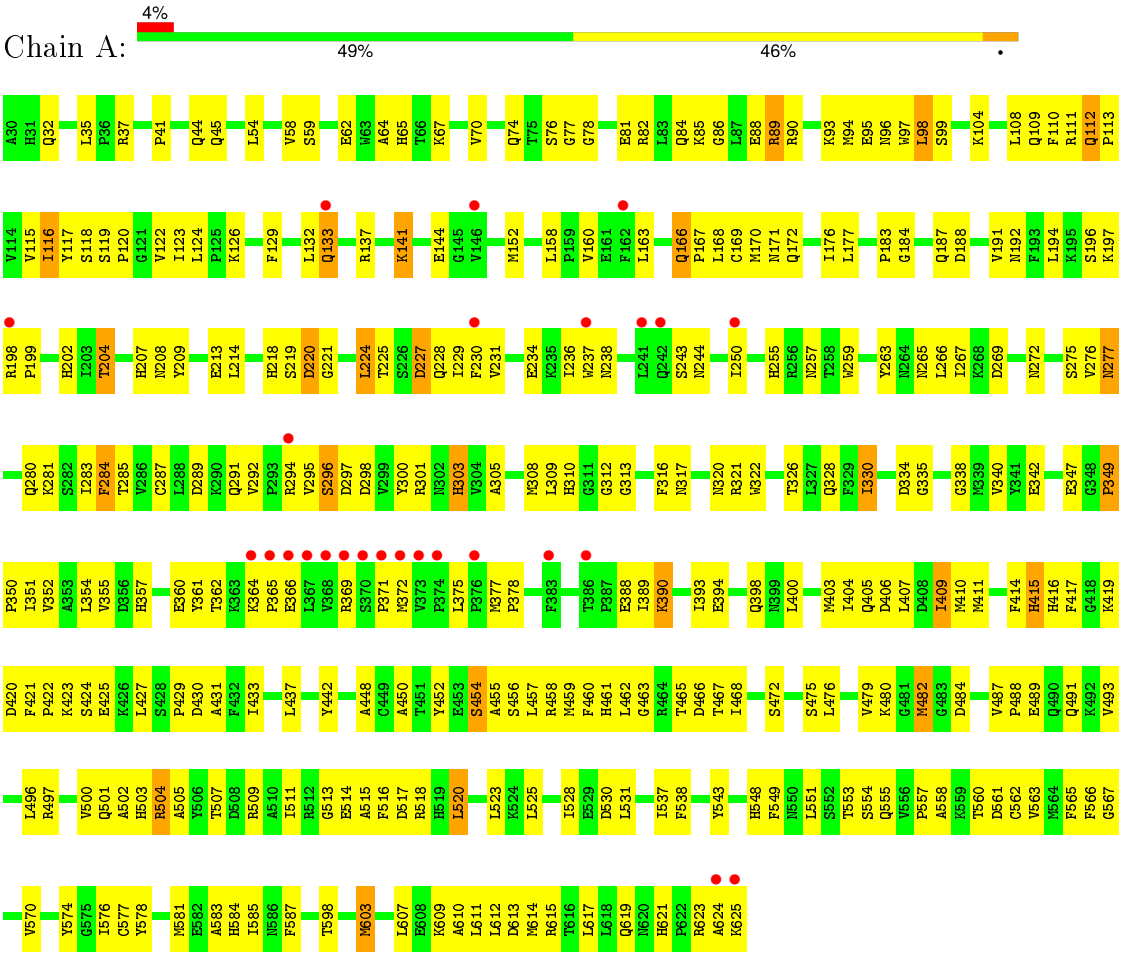
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	342	Total O 342 342	0	0
3	B	351	Total O 351 351	0	0

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.

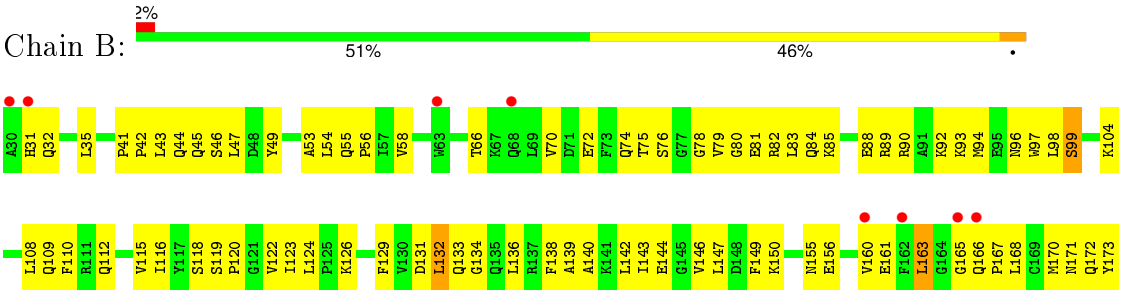
- Molecule 1: Carnitine Acetyltransferase

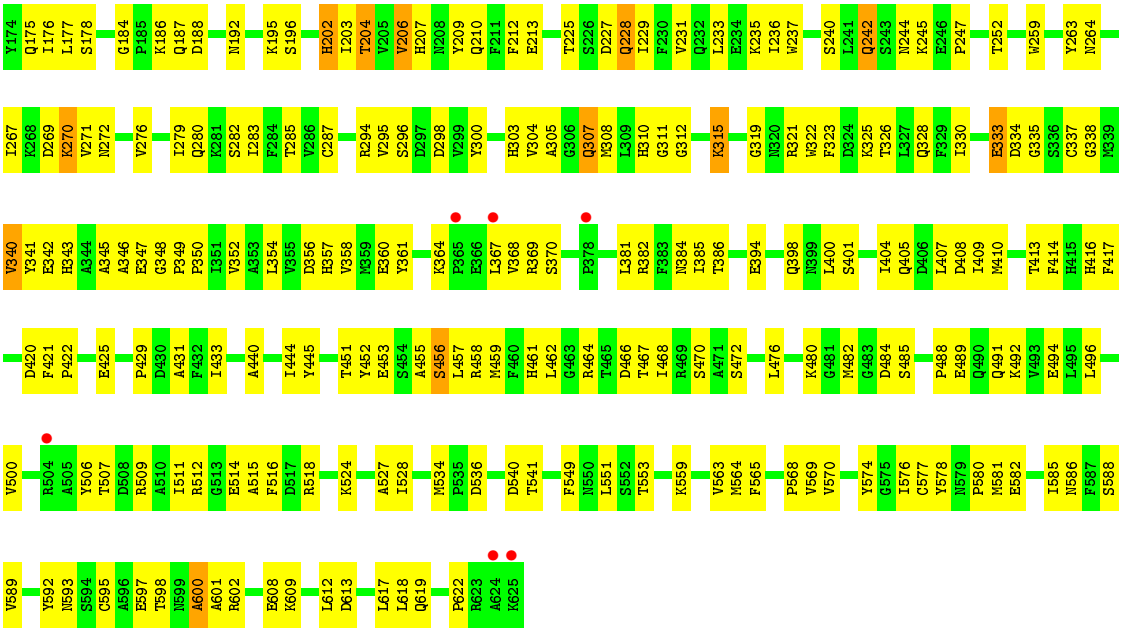
Chain A:



- Molecule 1: Carnitine Acetyltransferase

Chain B:




Y174 Q175 I176 L177 S178 G184 P185 K186 Q187 D188 N192 K195 S196 H202 I203 T204 Y205 V206 H207 N208 Y209 Q210 F211 F212 E213 T225 S226 D227 Q228 I229 F230 V231 Q232 L233 E234 K235 I236 W237 S240 L241 Q242 S243 N244 K245 E246 P247 T252 W259 Y263 N264
I267 K268 D269 K270 V271 N272 V276 I279 Q280 K281 S282 I283 F284 T285 V286 C287 R294 V295 S296 D297 D298 V299 Y300 H303 V304 A305 G306 Q307 M308 L309 H310 G311 G312 K315 G319 N320 R321 W322 F323 D324 K325 T326 L327 Q328 F329 I330 E333 D334 G335 S336 C337 G338 M339
V340 Y341 E342 H343 A344 A345 A346 E347 G348 P349 P350 I351 V352 A353 L354 V355 D356 H357 V358 M359 E360 Y361 K364 P365 E366 L367 V368 R369 S370 P378 L381 R382 F383 N384 I385 T386 E394 Q398 N399 L400 S401 I404 Q405 D406 L407 D408 I409 M410 T413 F414 H415 H416 F417
D420 F421 P422 E425 P429 D430 A431 F432 I433 A440 I444 Y445 T451 Y452 E453 S454 A455 S456 L457 R458 M459 F460 H461 L462 G463 R464 T465 D466 T467 I468 R469 S470 A471 S472 L476 K480 G481 M482 G483 D484 S485 P488 E489 Q490 Q491 K492 V493 E494 L495 L496
V500 R504 A505 Y506 T507 D508 R509 A510 I511 R512 G513 E514 A515 F516 D517 R518 K524 A527 I528 M534 P535 D536 D540 T541 F549 N550 L551 S552 T553 K559 V563 M564 F565 P568 V569 V570 Y574 G575 I576 C577 Y578 N579 P580 M581 E582 I585 N586 F587 S588
V589 Y592 N593 S594 C595 A596 E597 T598 N599 A600 A601 R602 E608 K609 L612 D613 L617 L618 Q619 P622 R623 A624 K625

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	162.26Å 92.05Å 122.90Å 90.00° 128.98° 90.00°	Depositor
Resolution (Å)	30.00 – 2.30 29.75 – 2.28	Depositor EDS
% Data completeness (in resolution range)	85.3 (30.00-2.30) 76.0 (29.75-2.28)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 2.29Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.270 , 0.363 0.271 , 0.356	Depositor DCC
R_{free} test set	4090 reflections (7.66%)	DCC
Wilson B-factor (Å^2)	34.7	Xtriage
Anisotropy	0.525	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 26.4	EDS
Estimated twinning fraction	0.055 for h,-k,-h-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Outliers	0 of 56477 reflections	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	10303	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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, 0.375 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: COA

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	0.44	0/4872	0.63	0/6600
1	B	0.47	0/4872	0.64	0/6600
All	All	0.46	0/9744	0.64	0/13200

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	4757	0	4733	279	0
1	B	4757	0	4733	281	0
2	A	48	0	32	5	0
2	B	48	0	32	7	0
3	A	342	0	0	31	0
3	B	351	0	0	52	0
All	All	10303	0	9530	563	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 29.

The worst 5 of 563 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:410:MET:HG2	1:B:601:ALA:HA	1.36	1.05
1:B:585:ILE:HG22	3:B:6421:HOH:O	1.60	1.02
1:A:90:ARG:HD3	1:A:97:TRP:HB2	1.44	0.98
1:A:32:GLN:HE22	1:A:170:MET:H	1.15	0.93
1:A:310:HIS:HD2	1:A:312:GLY:H	1.15	0.90

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	594/596 (100%)	501 (84%)	81 (14%)	12 (2%)	9 7
1	B	594/596 (100%)	532 (90%)	52 (9%)	10 (2%)	11 10
All	All	1188/1192 (100%)	1033 (87%)	133 (11%)	22 (2%)	10 8

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	99	SER
1	B	515	ALA
1	B	485	SER
1	B	622	PRO
1	A	172	GLN

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	524/524 (100%)	494 (94%)	30 (6%)	25 34
1	B	524/524 (100%)	502 (96%)	22 (4%)	36 49
All	All	1048/1048 (100%)	996 (95%)	52 (5%)	30 41

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	459	MET
1	A	625	LYS
1	B	470	SER
1	A	482	MET
1	A	517	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	501	GLN
1	B	84	GLN
1	B	526	GLN
1	A	586	ASN
1	B	109	GLN

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 (i)

2 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	COA	A	5170	-	40,50,50	1.25	2 (5%)	50,75,75	1.91	5 (10%)
2	COA	B	6170	-	40,50,50	1.15	2 (5%)	50,75,75	1.87	4 (8%)

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	COA	A	5170	-	-	0/44/64/64	0/3/3/3
2	COA	B	6170	-	-	0/44/64/64	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	5170	COA	P3B-O7A	3.32	1.62	1.51
2	B	6170	COA	P3B-O7A	3.41	1.62	1.51
2	B	6170	COA	O4B-C1B	4.06	1.46	1.41
2	A	5170	COA	O4B-C1B	4.54	1.46	1.41

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	6170	COA	N3A-C2A-N1A	-10.90	120.55	128.89
2	A	5170	COA	N3A-C2A-N1A	-10.88	120.56	128.89
2	A	5170	COA	P2A-O3A-P1A	-3.73	122.25	132.73
2	B	6170	COA	P2A-O3A-P1A	-3.48	122.94	132.73
2	A	5170	COA	C5B-C4B-C3B	-2.23	106.28	114.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	5170	COA	5	0
2	B	6170	COA	7	0

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, 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	596/596 (100%)	0.29	26 (4%)	38	47	13, 47, 89, 134	0
1	B	596/596 (100%)	0.13	14 (2%)	64	72	14, 42, 72, 99	0
All	All	1192/1192 (100%)	0.21	40 (3%)	49	58	13, 44, 85, 134	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	624	ALA	9.0
1	A	373	VAL	7.3
1	A	367	LEU	6.9
1	A	371	PRO	5.7
1	A	370	SER	5.1

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

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
2	COA	A	5170	48/48	0.91	0.13	0.13	39,54,60,62	0
2	COA	B	6170	48/48	0.92	0.13	-0.39	36,53,64,65	0

6.5 Other polymers (i)

There are no such residues in this entry.