wwPDB X-ray Structure Validation Summary Report (i)

Aug 19, 2023 – 10:32 PM EDT

PDB ID	:	2F8T
Title	:	Crystal structure of Aa-Ago with externally-bound siRNA
Authors	:	Yuan, Y.R.; Chen, H.Y.; Patel, D.J.
Deposited on	:	2005-12-03
Resolution	:	3.10 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

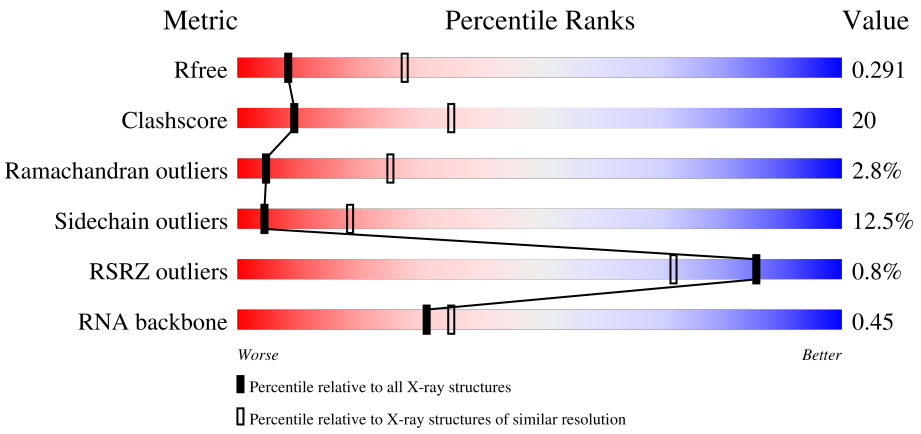
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)
RNA backbone	3102	1116 (3.40-2.80)

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 of 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 <=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	C	26	15% 46% 19% 19%
1	D	26	15% 38% 35% • 8%
2	A	706	58% 36% 6%
2	B	706	% 52% 39% 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12841 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 RNA chain called 26-MER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	26	551	246	94	185	26	0	0	0
1	D	24	511	228	90	169	24	0	0	0

- Molecule 2 is a protein called Argonaute protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	704	5864	3818	993	1042	11	0	0	0
2	B	704	5864	3818	993	1042	11	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	C	2	2	2	0	0
3	D	2	2	2	0	0
3	A	24	24	24	0	0
3	B	23	23	23	0	0

3 Residue-property plots (i)

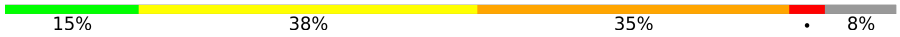
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

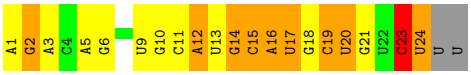
- Molecule 1: 26-MER

Chain C:

- Molecule 1: 26-MER

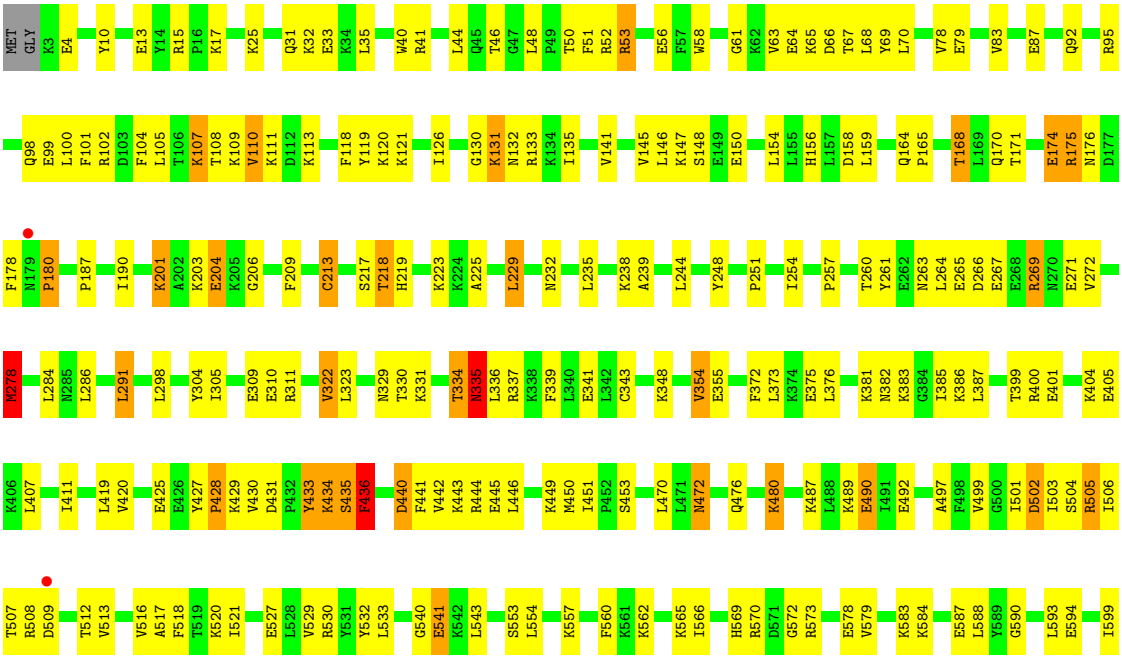
Chain D:

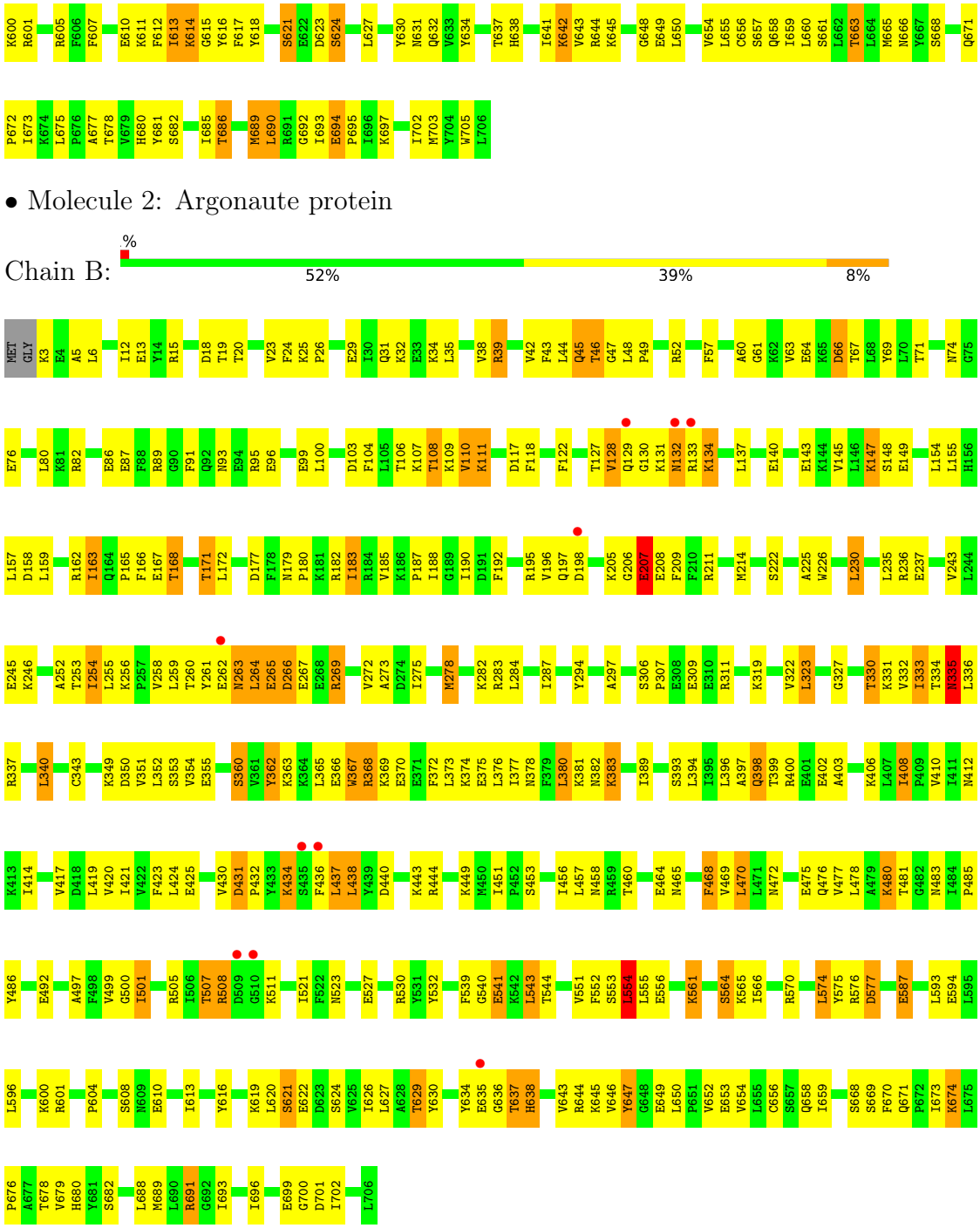




- Molecule 2: Argonaute protein

Chain A:




K600 R601 R605 F606 F607 E610 K611 F612 I613 K614 G615 Y616 F617 Y618 S621 E622 D623 S624 L627 Y630 N631 Q632 V633 Y634 T637 H638 I641 K642 V643 R644 K645 G648 E649 L650 V654 L655 C656 S657 Q658 I659 L660 S661 L662 T663 L664 M665 N666 Y667 S668 Q671
P672 I673 K674 L675 P676 A677 T678 V679 H680 Y681 S682 I685 T686 M689 L690 R691 G692 I693 E694 P695 I696 K697 I702 M703 Y704 W705 L706
• Molecule 2: Argonaute protein
Chain B:
%
52%
39%
8%
MET GLY K3 E4 A5 L6 I12 E13 Y14 R15 D18 T19 T20 V23 F24 K25 P26 E29 I30 Q31 K32 E33 K34 L35 V38 R39 V42 F43 L44 Q45 T46 G47 L48 P49 R52 F57 A60 G61 K62 V63 E64 K65 D66 T67 L68 Y69 L70 T71 N74 G75
E76 L80 R81 R82 E86 E87 F88 R89 G90 F91 Q92 N93 E94 R95 E96 E99 L100 D103 F104 L105 T106 K107 T108 K109 V110 K111 D117 F118 F122 T127 V128 Q129 G130 K131 N132 R133 K134 L137 E140 E143 K144 V145 L146 K147 S148 E149 L154 L155 H156
L157 D158 L159 R162 I163 Q164 P165 F166 E167 T168 T171 L172 D177 F178 N179 P180 K181 R182 I183 R184 V185 K186 P187 I188 G189 I190 D191 F192 R195 V196 Q197 D198 K205 G206 E207 E208 F209 F210 R211 M214 S222 A225 W226 L230 L235 R236 E237 V243 L244
E245 K246 A252 T253 I254 L255 K256 P257 V258 L259 T260 Y261 E262 N263 L264 E265 D266 E267 E268 R269 V272 A273 D274 I275 M278 K282 R283 L284 I287 Y294 A297 S306 P307 E308 E309 E310 R311 K319 V322 L323 G327 T330 K331 V332 I333 T334 N335 L336
R337 L340 C343 K349 D350 V351 L352 S353 V354 E355 S360 V361 Y362 K363 K364 L365 E366 W367 R368 K369 E370 E371 F372 L373 K374 E375 L376 I377 N378 F379 L380 K381 N382 K383 I389 S393 L394 I395 L396 A397 Q398 T399 R400 E401 E402 A403 K406 L407 I408 P409 V410 I411 N412
K413 I414 V417 D418 L419 V420 I421 V422 F423 L424 E425 V430 D431 P432 Y433 K434 S435 F436 L437 L438 Y439 D440 K443 R444 K449 M450 I451 P452 S453 I456 L457 N458 R459 T460 E464 N465 F468 V469 L470 L471 N472 E475 Q476 V477 L478 A479 K480 T481 G482 N483 I484 P485
Y486 E492 A497 F498 V499 G500 I501 R505 I506 T507 R508 D509 G510 K511 I521 F522 N523 E527 R530 Y531 Y532 F539 G540 E541 K542 L543 T544 V551 F552 S553 I554 L555 E556 K561 S564 K565 I566 R570 L574 Y575 R576 D577 E587 L593 E594 L595
L596 K600 R601 P604 S608 N609 E610 I613 Y616 K619 L620 S621 E622 D623 S624 V625 I626 L627 A628 T629 Y630 Y634 E635 G636 T637 H638 V643 R644 K645 V646 Y647 G648 E649 L650 P651 V652 E653 V654 L655 C656 S657 Q658 I659 S668 S669 F670 Q671 P672 I673 K674 L675
P676 A677 T678 V679 H680 Y681 S682 L688 M689 L690 R691 G692 I693 I696 E699 G700 D701 I702 L706

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	81.20Å 118.08Å 98.50Å 90.00° 99.57° 90.00°	Depositor
Resolution (Å)	40.00 – 3.10 32.44 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (40.00-3.10) 100.0 (32.44-3.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.71 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.204 , 0.297 0.201 , 0.291	Depositor DCC
R_{free} test set	1672 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	58.5	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 61.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12841	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% 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, 0.333 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	C	2.45	24/614 (3.9%)	2.20	31/952 (3.3%)
1	D	1.33	5/570 (0.9%)	1.75	14/884 (1.6%)
2	A	0.76	5/5979 (0.1%)	0.82	2/8027 (0.0%)
2	B	0.72	0/5979	0.81	2/8027 (0.0%)
All	All	0.93	34/13142 (0.3%)	1.00	49/17890 (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	C	0	1
2	A	0	1
2	B	0	4
All	All	0	6

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	2	G	N9-C8	21.71	1.53	1.37
1	C	2	G	C5-C4	16.53	1.50	1.38
1	C	1	A	N7-C5	16.30	1.49	1.39
1	C	1	A	N9-C4	14.03	1.46	1.37
1	C	2	G	N3-C4	13.16	1.44	1.35

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	2	G	C8-N9-C4	-17.28	99.49	106.40
1	C	2	G	C2-N3-C4	-13.63	105.08	111.90
1	C	2	G	N3-C2-N2	-11.04	112.17	119.90
1	C	2	G	N3-C4-N9	-10.15	119.91	126.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	2	G	N7-C8-N9	10.09	118.15	113.10

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	335	ASN	Peptide
2	B	507	THR	Peptide
2	B	576	ARG	Peptide
2	B	621	SER	Peptide
1	C	1	A	Sidechain

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	C	551	0	279	25	0
1	D	511	0	259	31	0
2	A	5864	0	6128	210	0
2	B	5864	0	6128	251	0
3	A	24	0	0	3	0
3	B	23	0	0	1	0
3	C	2	0	0	0	0
3	D	2	0	0	1	0
All	All	12841	0	12794	504	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 20.

The worst 5 of 504 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)
2:A:642:LYS:CD	2:A:642:LYS:CG	1.76	1.61
2:B:532:TYR:CZ	2:B:554:LEU:HD21	1.70	1.27
2:B:532:TYR:CE1	2:B:554:LEU:HD21	1.78	1.18

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:66:ASP:HB3	3:A:708:HOH:O	1.51	1.11
2:B:122:PHE:CZ	2:B:275:ILE:HG21	1.87	1.09

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	
2	A	702/706 (99%)	605 (86%)	77 (11%)	20 (3%)	5	25
2	B	702/706 (99%)	605 (86%)	78 (11%)	19 (3%)	5	25
All	All	1404/1412 (99%)	1210 (86%)	155 (11%)	39 (3%)	5	25

5 of 39 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	434	LYS
2	A	435	SER
2	B	130	GLY
2	B	132	ASN
2	B	265	GLU

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
2	A	646/647 (100%)	580 (90%)	66 (10%)	7 27
2	B	646/647 (100%)	550 (85%)	96 (15%)	3 13
All	All	1292/1294 (100%)	1130 (88%)	162 (12%)	4 18

5 of 162 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	355	GLU
2	B	564	SER
2	B	368	ARG
2	B	464	GLU
2	B	637	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	523	ASN
2	B	335	ASN
2	A	602	ASN
2	B	263	ASN
2	A	569	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	C	25/26 (96%)	6 (24%)	0
1	D	23/26 (88%)	6 (26%)	1 (4%)
All	All	48/52 (92%)	12 (25%)	1 (2%)

5 of 12 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	C	3	A
1	C	6	G
1	C	9	U
1	C	10	G
1	C	11	C

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	D	23	C

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)

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 (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	C	26/26 (100%)	-0.10	0	100	100	31, 58, 79, 82	0
1	D	24/26 (92%)	0.23	0	100	100	26, 59, 89, 91	0
2	A	704/706 (99%)	-0.48	2 (0%)	94	88	27, 54, 82, 96	0
2	B	704/706 (99%)	-0.31	10 (1%)	75	56	30, 55, 78, 93	0
All	All	1458/1464 (99%)	-0.38	12 (0%)	86	72	26, 54, 80, 96	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	132	ASN	3.0
2	A	179	ASN	2.9
2	B	509	ASP	2.8
2	B	435	SER	2.5
2	B	129	GLN	2.5

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.