wwPDB X-ray Structure Validation Summary Report (i)

Aug 16, 2023 – 02:29 PM EDT

PDB ID	:	2ACI
Title	:	Structure of D166A arginine deiminase
Authors	:	Galkin, A.; Lu, X.; Dunaway-Mariano, D.; Herzberg, O.
Deposited on	:	2005-07-18
Resolution	:	2.50 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 (i)) were used in the production of this report:

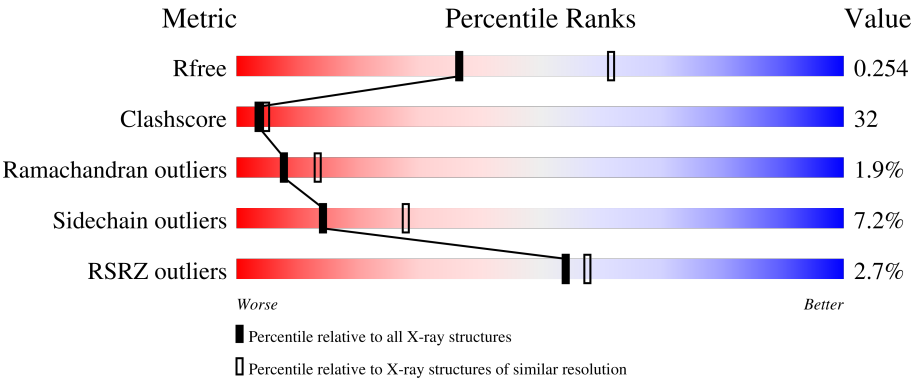
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	A	418	2% 42% 49% 5%
1	B	418	3% 61% 32%
1	C	418	2% 45% 47%
1	D	418	3% 46% 46%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12953 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 Arginine deiminase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	402	Total	C	N	O	S	0	0	0
			3139	1988	546	588	17			
1	B	409	Total	C	N	O	S	0	0	0
			3193	2020	556	600	17			
1	C	403	Total	C	N	O	S	0	0	0
			3147	1992	547	591	17			
1	D	406	Total	C	N	O	S	0	0	0
			3174	2009	553	595	17			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	166	ALA	ASP	engineered mutation	UNP P13981
B	166	ALA	ASP	engineered mutation	UNP P13981
C	166	ALA	ASP	engineered mutation	UNP P13981
D	166	ALA	ASP	engineered mutation	UNP P13981

- Molecule 2 is water.

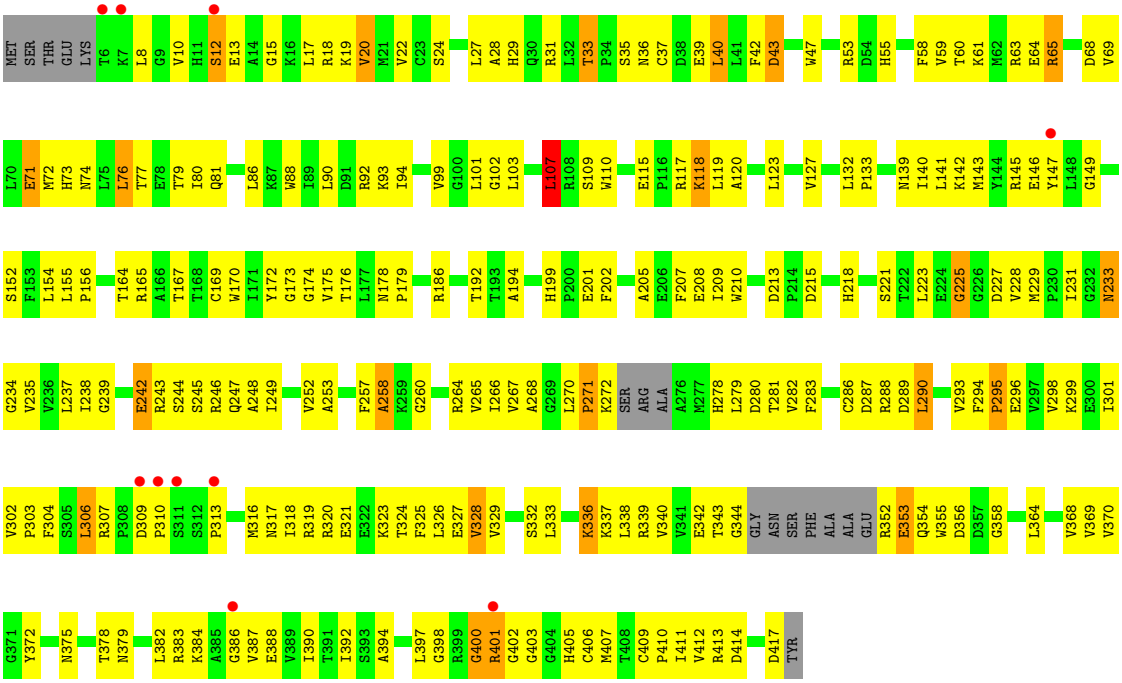
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	60	Total	O	0	0
			60	60		
2	B	90	Total	O	0	0
			90	90		
2	C	86	Total	O	0	0
			86	86		
2	D	64	Total	O	0	0
			64	64		

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.

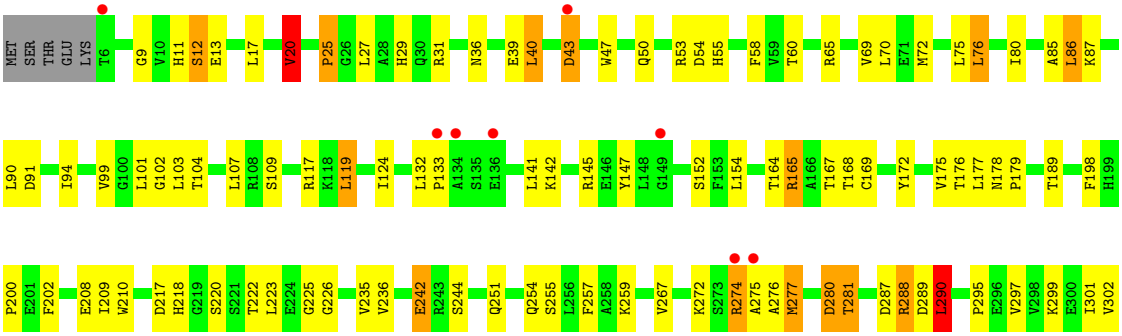
• Molecule 1: Arginine deiminase

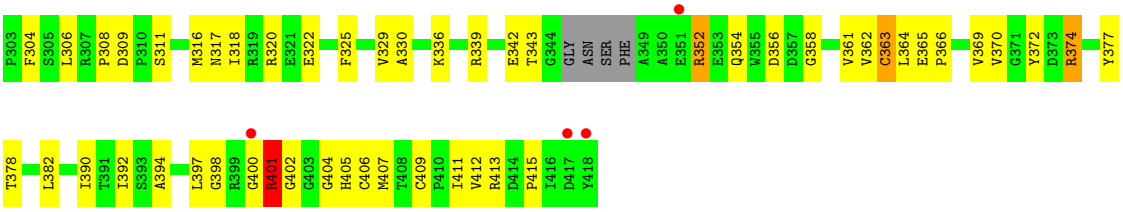
Chain A:

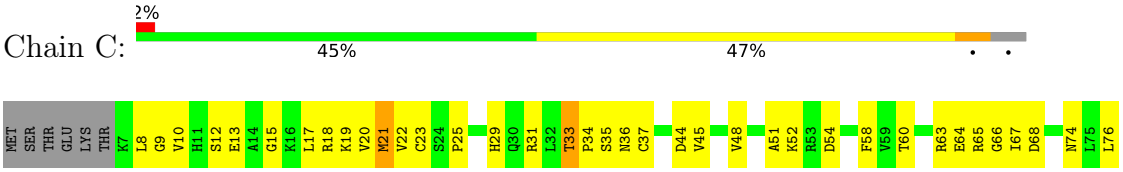


• Molecule 1: Arginine deiminase

Chain B:

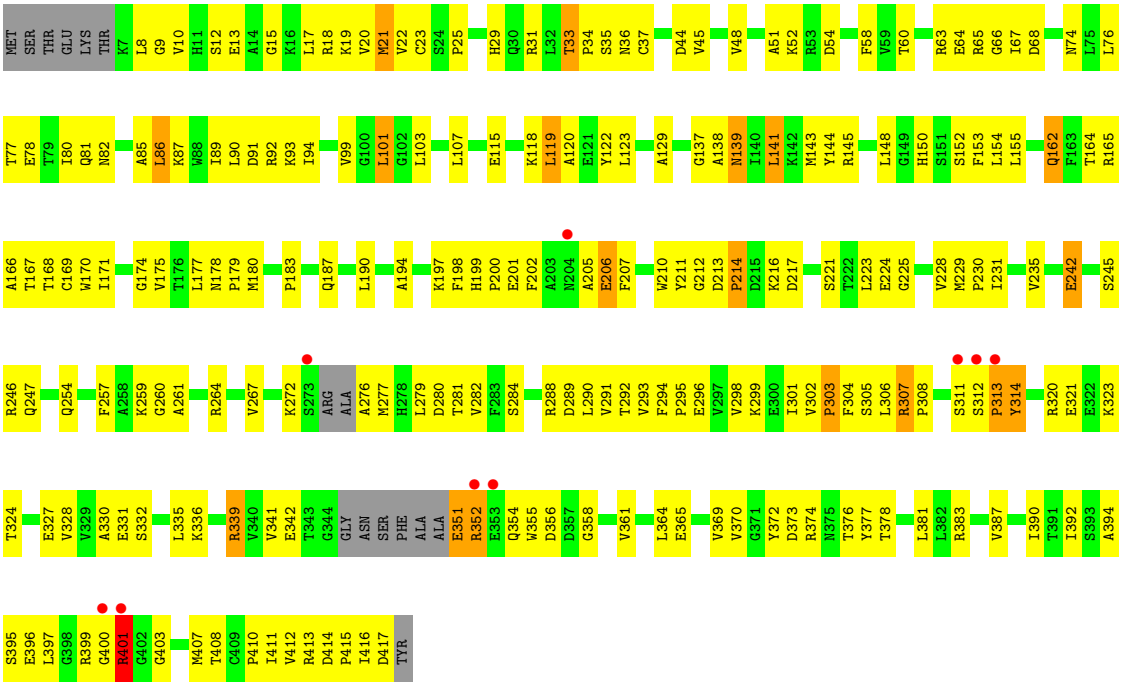






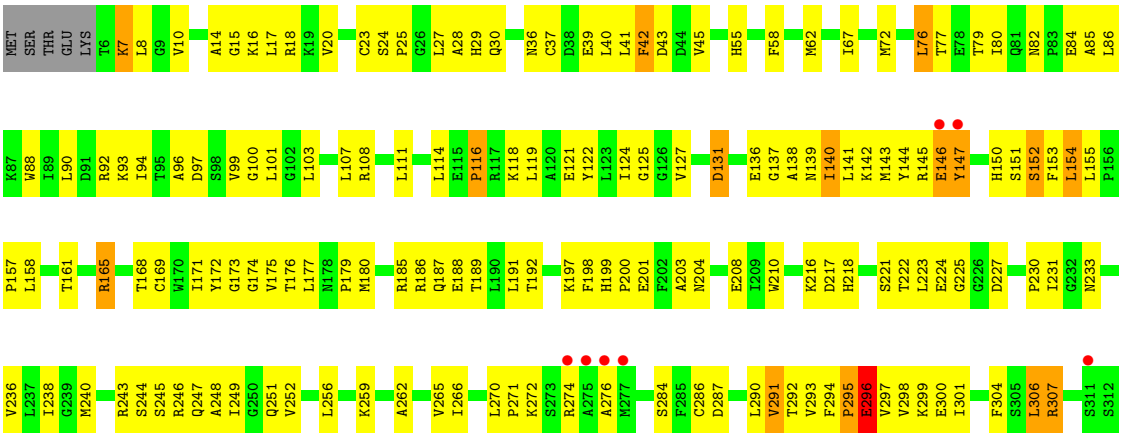
• Molecule 1: Arginine deiminase

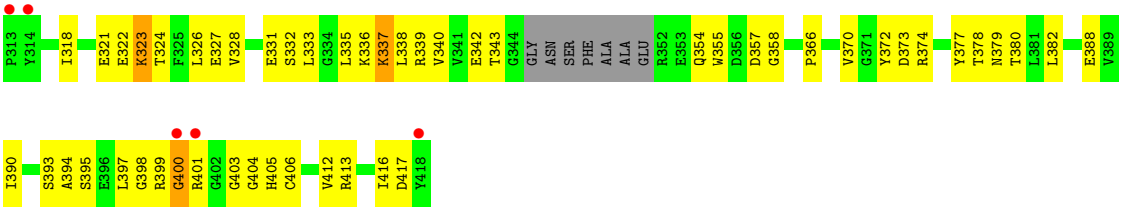
Chain C:



• Molecule 1: Arginine deiminase

Chain D:




P313
Y314
I318
E321
E322
K323
T324
F325
L326
E327
V328
E331
S332
L333
G334
L335
K336
K337
L338
R339
V340
V341
E342
T343
G344
GLY
ASN
SER
PHE
ALA
ALA
GLU
R352
E353
Q354
W355
D356
D357
G358
P366
V370
G371
Y372
D373
R374
Y377
T378
N379
T380
L381
L382
E388
V389

I390
S393
A394
S395
E396
L397
G398
R399
G400
R401
G402
G403
G404
H405
C406
V412
R413
I416
D417
Y418

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	91.20Å 123.90Å 150.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 19.96 – 2.90	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.50) 93.9 (19.96-2.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.16 (at 2.88Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.198 , 0.272 0.182 , 0.254	Depositor DCC
R_{free} test set	1806 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	46.0	Xtriage
Anisotropy	0.128	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 63.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12953	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.68% 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

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	0.70	1/3205 (0.0%)	0.93	3/4346 (0.1%)
1	B	0.77	0/3261	0.96	3/4422 (0.1%)
1	C	0.74	3/3213 (0.1%)	0.92	1/4356 (0.0%)
1	D	0.73	0/3242	0.92	2/4396 (0.0%)
All	All	0.74	4/12921 (0.0%)	0.93	9/17520 (0.1%)

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

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	303	PRO	N-CD	10.53	1.62	1.47
1	C	214	PRO	N-CD	-5.74	1.39	1.47
1	A	37	CYS	CB-SG	-5.34	1.73	1.81
1	C	242	GLU	CG-CD	5.18	1.59	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	186	ARG	NE-CZ-NH2	-5.55	117.53	120.30
1	D	86	LEU	CA-CB-CG	5.33	127.55	115.30
1	C	339	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	B	306	LEU	CA-CB-CG	5.28	127.45	115.30
1	B	20	VAL	CB-CA-C	-5.19	101.54	111.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	144	TYR	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	A	3139	0	3128	256	0
1	B	3193	0	3177	153	0
1	C	3147	0	3132	215	0
1	D	3174	0	3161	208	0
2	A	60	0	0	23	0
2	B	90	0	0	22	0
2	C	86	0	0	15	0
2	D	64	0	0	14	0
All	All	12953	0	12598	798	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 32.

The worst 5 of 798 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:D:43:ASP:HB2	1:D:401:ARG:HH12	1.05	1.19
1:C:277:MET:HE2	1:C:281:THR:HG21	1.29	1.13
1:C:33:THR:HG22	1:C:35:SER:H	1.17	1.10
1:A:17:LEU:HD11	1:A:20:VAL:HG13	1.34	1.08
1:D:343:THR:HG21	1:D:358:GLY:N	1.69	1.07

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	396/418 (95%)	342 (86%)	45 (11%)	9 (2%)	6 10
1	B	405/418 (97%)	366 (90%)	36 (9%)	3 (1%)	22 39
1	C	397/418 (95%)	346 (87%)	45 (11%)	6 (2%)	10 18
1	D	402/418 (96%)	336 (84%)	53 (13%)	13 (3%)	4 5
All	All	1600/1672 (96%)	1390 (87%)	179 (11%)	31 (2%)	8 13

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	43	ASP
1	D	152	SER
1	D	276	ALA
1	D	323	LYS
1	A	172	TYR

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	342/354 (97%)	314 (92%)	28 (8%)	11 22
1	B	346/354 (98%)	315 (91%)	31 (9%)	9 19
1	C	343/354 (97%)	321 (94%)	22 (6%)	17 33
1	D	345/354 (98%)	327 (95%)	18 (5%)	23 44
All	All	1376/1416 (97%)	1277 (93%)	99 (7%)	14 28

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	401	ARG
1	C	280	ASP
1	C	20	VAL
1	C	119	LEU
1	C	314	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	81	GLN
1	B	317	ASN
1	C	139	ASN
1	D	81	GLN
1	D	150	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	402/418 (96%)	-0.06	10 (2%)	57	61	18, 41, 64, 78	0
1	B	409/418 (97%)	-0.20	12 (2%)	51	55	16, 30, 60, 78	0
1	C	403/418 (96%)	-0.21	9 (2%)	62	65	17, 36, 62, 78	0
1	D	406/418 (97%)	-0.07	12 (2%)	50	53	17, 39, 66, 79	0
All	All	1620/1672 (96%)	-0.13	43 (2%)	54	58	16, 37, 64, 79	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	275	ALA	5.8
1	D	418	TYR	5.6
1	A	6	THR	5.1
1	D	274	ARG	4.3
1	C	352	ARG	4.2

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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.