wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 12:54 PM GMT

PDB ID	:	3S7Y
Title	:	Crystal structure of mmNAGS in Space Group P3121 at 4.3 A resolution
Authors	:	Shi, D.; Li, Y.; Cabrera-Luque, J.; Jin, Z.; Yu, X.; Allewell, N.M.; Tuchman, M.
Deposited on	:	2011-05-27
Resolution	:	4.31 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:

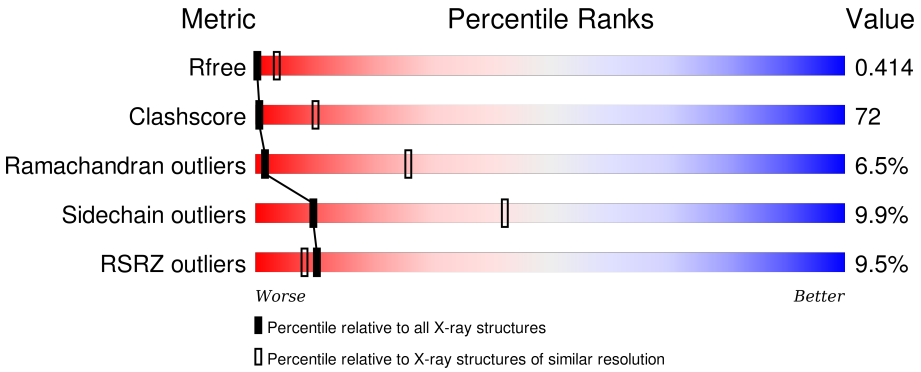
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)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

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 4.31 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	1059 (5.00-3.60)
Clashscore	102246	1166 (5.00-3.60)
Ramachandran outliers	100387	1106 (5.00-3.60)
Sidechain outliers	100360	1089 (5.00-3.60)
RSRZ outliers	91569	1062 (5.00-3.60)

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	461	7% 28% 54% 11% • 5%
1	X	461	11% 31% 54% 10% • 6%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 6683 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 N-acetylglutamate kinase / N-acetylglutamate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	436	Total	C	N	O	S	0	1	0
			3349	2107	605	631	6			
1	X	435	Total	C	N	O	S	0	0	0
			3334	2097	601	630	6			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q0ASS9
A	-18	GLY	-	EXPRESSION TAG	UNP Q0ASS9
A	-17	SER	-	EXPRESSION TAG	UNP Q0ASS9
A	-16	SER	-	EXPRESSION TAG	UNP Q0ASS9
A	-15	HIS	-	EXPRESSION TAG	UNP Q0ASS9
A	-14	HIS	-	EXPRESSION TAG	UNP Q0ASS9
A	-13	HIS	-	EXPRESSION TAG	UNP Q0ASS9
A	-12	HIS	-	EXPRESSION TAG	UNP Q0ASS9
A	-11	HIS	-	EXPRESSION TAG	UNP Q0ASS9
A	-10	HIS	-	EXPRESSION TAG	UNP Q0ASS9
A	-9	SER	-	EXPRESSION TAG	UNP Q0ASS9
A	-8	SER	-	EXPRESSION TAG	UNP Q0ASS9
A	-7	GLY	-	EXPRESSION TAG	UNP Q0ASS9
A	-6	LEU	-	EXPRESSION TAG	UNP Q0ASS9
A	-5	VAL	-	EXPRESSION TAG	UNP Q0ASS9
A	-4	PRO	-	EXPRESSION TAG	UNP Q0ASS9
A	-3	ARG	-	EXPRESSION TAG	UNP Q0ASS9
A	-2	GLY	-	EXPRESSION TAG	UNP Q0ASS9
A	-1	SER	-	EXPRESSION TAG	UNP Q0ASS9
A	0	HIS	-	EXPRESSION TAG	UNP Q0ASS9
X	-19	MET	-	EXPRESSION TAG	UNP Q0ASS9
X	-18	GLY	-	EXPRESSION TAG	UNP Q0ASS9
X	-17	SER	-	EXPRESSION TAG	UNP Q0ASS9
X	-16	SER	-	EXPRESSION TAG	UNP Q0ASS9
X	-15	HIS	-	EXPRESSION TAG	UNP Q0ASS9

Continued on next page...

Continued from previous page...

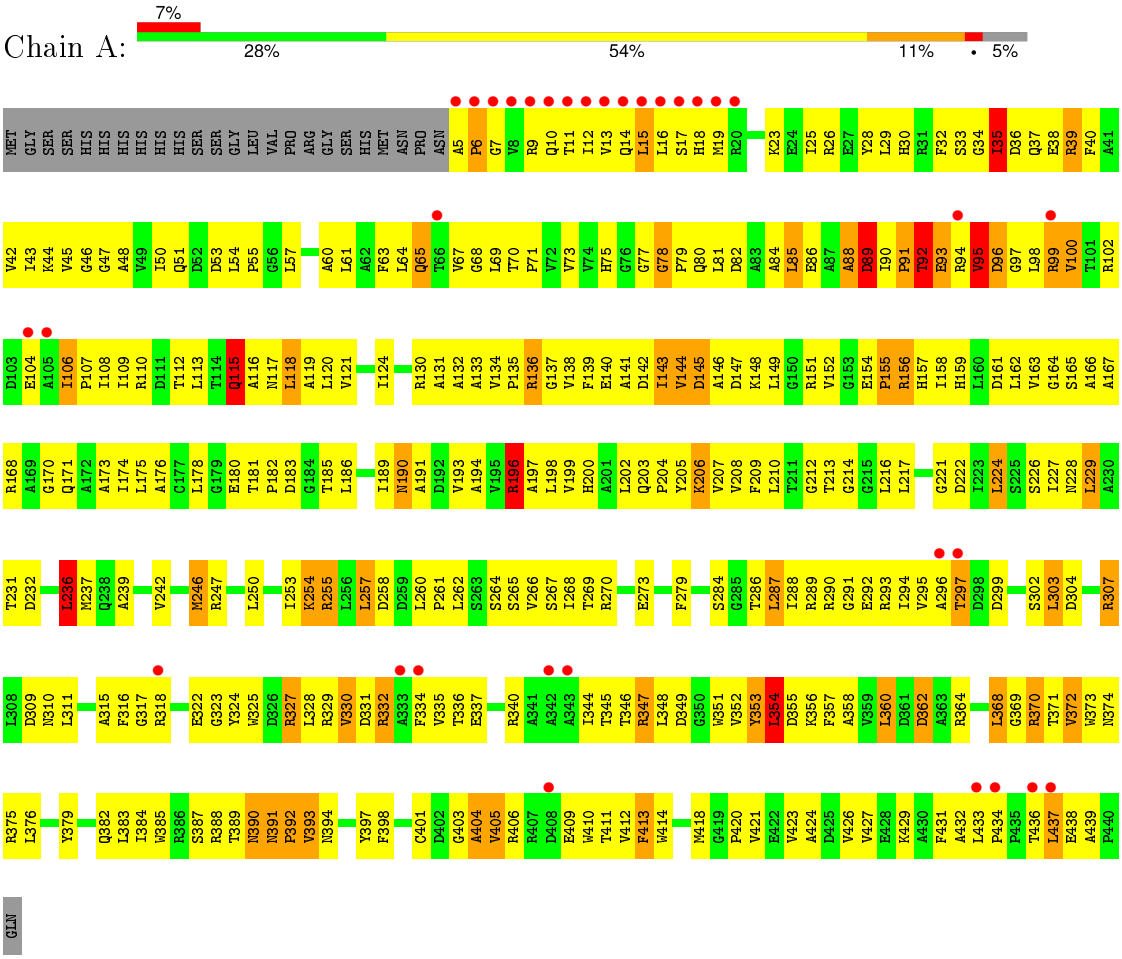
Chain	Residue	Modelled	Actual	Comment	Reference
X	-14	HIS	-	EXPRESSION TAG	UNP Q0ASS9
X	-13	HIS	-	EXPRESSION TAG	UNP Q0ASS9
X	-12	HIS	-	EXPRESSION TAG	UNP Q0ASS9
X	-11	HIS	-	EXPRESSION TAG	UNP Q0ASS9
X	-10	HIS	-	EXPRESSION TAG	UNP Q0ASS9
X	-9	SER	-	EXPRESSION TAG	UNP Q0ASS9
X	-8	SER	-	EXPRESSION TAG	UNP Q0ASS9
X	-7	GLY	-	EXPRESSION TAG	UNP Q0ASS9
X	-6	LEU	-	EXPRESSION TAG	UNP Q0ASS9
X	-5	VAL	-	EXPRESSION TAG	UNP Q0ASS9
X	-4	PRO	-	EXPRESSION TAG	UNP Q0ASS9
X	-3	ARG	-	EXPRESSION TAG	UNP Q0ASS9
X	-2	GLY	-	EXPRESSION TAG	UNP Q0ASS9
X	-1	SER	-	EXPRESSION TAG	UNP Q0ASS9
X	0	HIS	-	EXPRESSION TAG	UNP Q0ASS9

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.

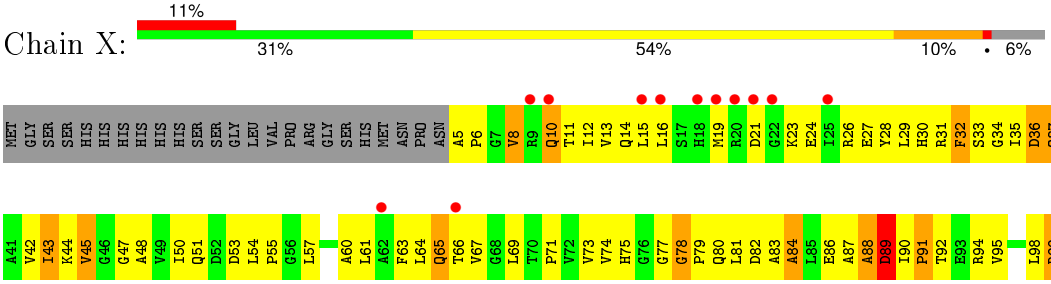
- Molecule 1: N-acetylglutamate kinase / N-acetylglutamate synthase

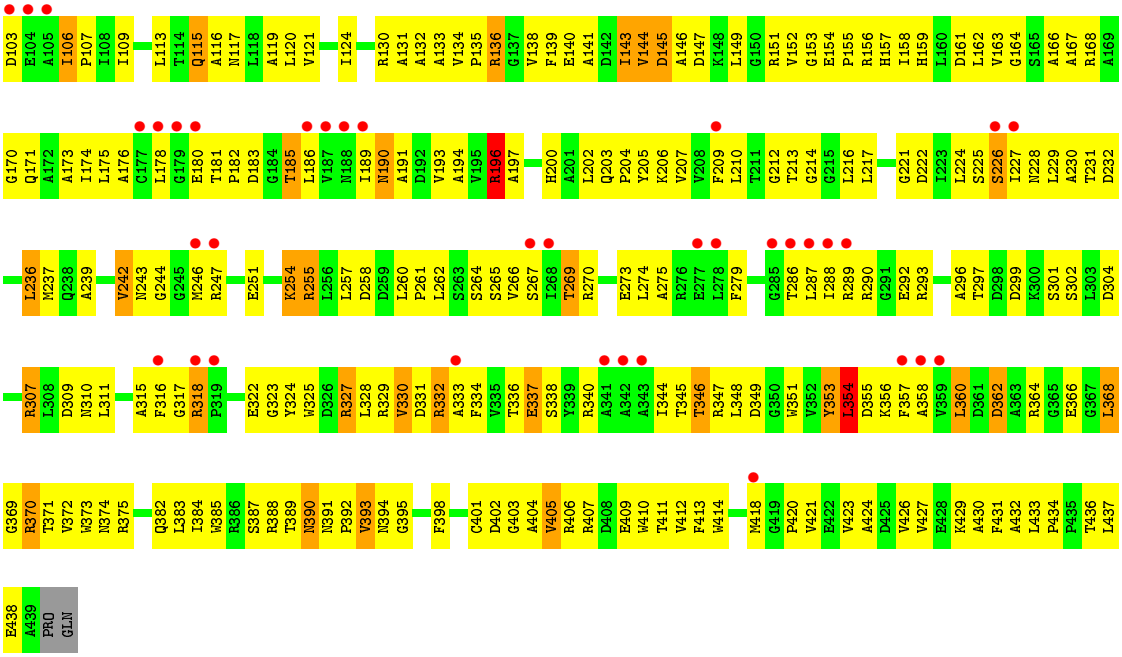
Chain A:



- Molecule 1: N-acetylglutamate kinase / N-acetylglutamate synthase

Chain X:




D103 E104 A105 I106 P107 I108 I109 L113 T114 Q115 A116 N117 L118 A119 L120 V121 I124 R130 A131 A132 A133 V134 P135 R136 G137 V138 F139 E140 A141 D142 I143 V144 D145 A146 D147 K148 L149 G150 R151 V152 G153 E154 P155 R156 H157 I158 H159 L160 D161 L162 V163 G164 S165 A166 A167 R168 A169
G170 Q171 A172 A173 I174 L175 A176 C177 L178 G179 E180 T181 P182 D183 G184 T185 L186 V187 N188 I189 N190 A191 D192 V193 A194 V195 R196 A197 H200 A201 L202 Q203 P204 Y205 K206 V207 V208 F209 L210 T211 G212 T213 G214 G215 L216 L217 G221 D222 I223 L224 S225 S226 I227 N228 L229 A230 T231 D232
L236 M237 Q238 A239 V242 N243 G244 G245 M246 R247 E251 K254 R255 L256 L257 D258 D259 L260 P261 L262 S263 S264 S265 V266 S267 I268 T269 R270 E273 L274 A275 R276 E277 L278 F279 G285 T286 L287 I288 R289 R290 G291 E292 R293 A296 T297 D298 D299 K300 S301 S302 L303 D304
R307 L308 D309 N310 L311 A315 F316 G317 R318 P319 E322 G323 Y324 W325 D326 R327 L328 R329 V330 D331 R332 A333 F334 V335 T336 E337 S338 Y339 R340 A341 A342 A343 T344 T345 T346 R347 L348 D349 G350 W351 V352 Y353 L354 D355 K356 F357 A358 V359 L360 D361 D362 A363 R364 G365 E366 G367 L368
G369 R370 T371 V372 W373 N374 R375 Q382 L383 I384 W385 R386 S387 R388 T389 N390 N391 P392 V393 N394 G395 F398 C401 D402 G403 A404 V405 R406 R407 D408 E409 W410 T411 V412 F413 W414 M418 G419 P420 V421 E422 V423 A424 D425 V426 V427 E428 K429 A430 F431 A432 L433 P434 F435 T436 L437
E438 A439 PRO GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	95.11Å 95.11Å 253.01Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	31.13 – 4.31 46.74 – 4.31	Depositor EDS
% Data completeness (in resolution range)	92.1 (31.13-4.31) 89.2 (46.74-4.31)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.99 (at 4.29Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.274 , 0.419 0.262 , 0.414	Depositor DCC
R_{free} test set	867 reflections (9.93%)	DCC
Wilson B-factor (Å^2)	192.6	Xtriage
Anisotropy	0.090	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 406.6	EDS
Estimated twinning fraction	0.067 for -h,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Outliers	0 of 9472 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6683	wwPDB-VP
Average B, all atoms (Å^2)	456.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% 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

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	0.77	1/3412 (0.0%)	1.29	26/4639 (0.6%)
1	X	0.66	0/3393	0.99	12/4613 (0.3%)
All	All	0.72	1/6805 (0.0%)	1.15	38/9252 (0.4%)

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 (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	404	ALA	CA-CB	-6.79	1.38	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	368	LEU	CB-CG-CD1	-9.50	94.86	111.00
1	A	236	LEU	CA-CB-CG	-9.06	94.45	115.30
1	A	260	LEU	CA-CB-CG	8.92	135.81	115.30
1	A	85	LEU	CB-CG-CD2	-8.54	96.49	111.00
1	A	354	LEU	CA-CB-CG	-7.52	98.01	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	115	GLN	Sidechain
1	A	413	PHE	Mainchain

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	3349	0	3352	533	0
1	X	3334	0	3332	462	0
All	All	6683	0	6684	956	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 72.

The worst 5 of 956 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:A:296:ALA:O	1:A:297:THR:HG22	1.21	1.26
1:A:398:PHE:CD1	1:X:398:PHE:CD1	2.33	1.15
1:A:95:VAL:H	1:A:100:VAL:CG1	1.59	1.14
1:A:295:VAL:HG22	1:A:335:VAL:O	1.46	1.14
1:A:266:VAL:O	1:A:287:LEU:HA	1.47	1.14

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	435/461 (94%)	360 (83%)	46 (11%)	29 (7%)	1 25
1	X	433/461 (94%)	349 (81%)	57 (13%)	27 (6%)	2 27
All	All	868/922 (94%)	709 (82%)	103 (12%)	56 (6%)	1 26

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	ILE
1	A	89	ASP
1	A	95	VAL
1	A	96	ASP
1	A	99	ARG

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	346/367 (94%)	310 (90%)	36 (10%)	9 40
1	X	344/367 (94%)	311 (90%)	33 (10%)	10 44
All	All	690/734 (94%)	621 (90%)	69 (10%)	10 41

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	390	ASN
1	X	37	GLN
1	X	364	ARG
1	A	393	VAL
1	A	437	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	65	GLN
1	X	80	GLN
1	X	14	GLN
1	A	30	HIS
1	A	80	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

There are no ligands in this entry.

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	436/461 (94%)	0.37	33 (7%) 17 12	185, 422, 690, 893	0
1	X	435/461 (94%)	0.48	50 (11%) 6 6	209, 472, 685, 784	0
All	All	871/922 (94%)	0.42	83 (9%) 10 8	185, 455, 688, 893	0

The worst 5 of 83 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	8	VAL	16.9
1	A	5	ALA	13.2
1	X	342	ALA	11.6
1	A	17	SER	11.5
1	A	18	HIS	11.2

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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.