



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 25, 2017 – 08:13 PM EDT

PDB ID : 3PQ1
Title : Crystal structure of human mitochondrial poly(A) polymerase (PAPD1)
Authors : Bai, Y.; Srivastava, S.K.; Chang, J.H.; Tong, L.
Deposited on : unknown
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

<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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
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-20029824

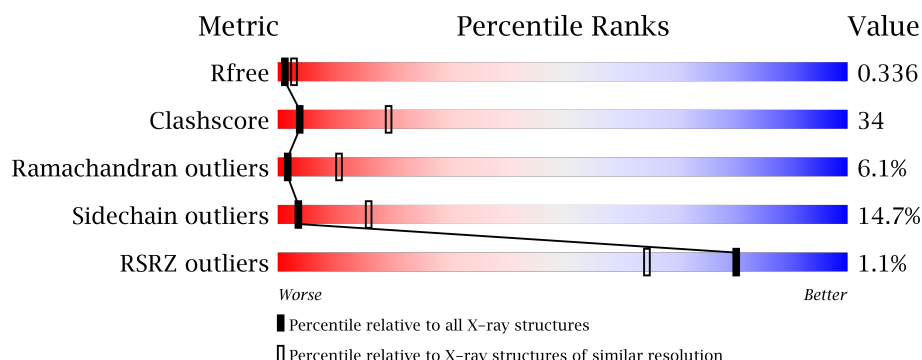
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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}	100719	1001 (3.12-3.08)
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)

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	464	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 36% 31% 8% 23% </div> </div>
1	B	464	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 29% 33% 7% 31% </div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5351 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 Poly(A) RNA polymerase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	355	Total	C	N	O	S	Se	0	0	0
			2793	1782	474	520	11	6			
1	B	321	Total	C	N	O	S	Se	0	0	0
			2558	1629	438	475	10	6			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MSE	-	EXPRESSION TAG	UNP Q9NVV4
A	24	GLY	-	EXPRESSION TAG	UNP Q9NVV4
A	25	SER	-	EXPRESSION TAG	UNP Q9NVV4
A	26	SER	-	EXPRESSION TAG	UNP Q9NVV4
A	27	HIS	-	EXPRESSION TAG	UNP Q9NVV4
A	28	HIS	-	EXPRESSION TAG	UNP Q9NVV4
A	29	HIS	-	EXPRESSION TAG	UNP Q9NVV4
A	30	HIS	-	EXPRESSION TAG	UNP Q9NVV4
A	31	HIS	-	EXPRESSION TAG	UNP Q9NVV4
A	32	HIS	-	EXPRESSION TAG	UNP Q9NVV4
A	33	SER	-	EXPRESSION TAG	UNP Q9NVV4
A	34	SER	-	EXPRESSION TAG	UNP Q9NVV4
A	35	GLY	-	EXPRESSION TAG	UNP Q9NVV4
A	36	LEU	-	EXPRESSION TAG	UNP Q9NVV4
A	37	VAL	-	EXPRESSION TAG	UNP Q9NVV4
A	38	PRO	-	EXPRESSION TAG	UNP Q9NVV4
A	39	ARG	-	EXPRESSION TAG	UNP Q9NVV4
A	40	GLY	-	EXPRESSION TAG	UNP Q9NVV4
A	41	SER	-	EXPRESSION TAG	UNP Q9NVV4
A	42	HIS	-	EXPRESSION TAG	UNP Q9NVV4
A	43	MSE	-	EXPRESSION TAG	UNP Q9NVV4
A	151	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	152	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	153	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	154	UNK	-	SEE REMARK 999	UNP Q9NVV4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	155	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	156	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	157	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	325	ALA	ASP	ENGINEERED MUTATION	UNP Q9NVV4
A	464	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	465	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	466	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	467	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	468	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	469	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	470	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	471	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	472	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	473	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	474	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	475	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	476	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	477	UNK	-	SEE REMARK 999	UNP Q9NVV4
A	478	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	23	MSE	-	EXPRESSION TAG	UNP Q9NVV4
B	24	GLY	-	EXPRESSION TAG	UNP Q9NVV4
B	25	SER	-	EXPRESSION TAG	UNP Q9NVV4
B	26	SER	-	EXPRESSION TAG	UNP Q9NVV4
B	27	HIS	-	EXPRESSION TAG	UNP Q9NVV4
B	28	HIS	-	EXPRESSION TAG	UNP Q9NVV4
B	29	HIS	-	EXPRESSION TAG	UNP Q9NVV4
B	30	HIS	-	EXPRESSION TAG	UNP Q9NVV4
B	31	HIS	-	EXPRESSION TAG	UNP Q9NVV4
B	32	HIS	-	EXPRESSION TAG	UNP Q9NVV4
B	33	SER	-	EXPRESSION TAG	UNP Q9NVV4
B	34	SER	-	EXPRESSION TAG	UNP Q9NVV4
B	35	GLY	-	EXPRESSION TAG	UNP Q9NVV4
B	36	LEU	-	EXPRESSION TAG	UNP Q9NVV4
B	37	VAL	-	EXPRESSION TAG	UNP Q9NVV4
B	38	PRO	-	EXPRESSION TAG	UNP Q9NVV4
B	39	ARG	-	EXPRESSION TAG	UNP Q9NVV4
B	40	GLY	-	EXPRESSION TAG	UNP Q9NVV4
B	41	SER	-	EXPRESSION TAG	UNP Q9NVV4
B	42	HIS	-	EXPRESSION TAG	UNP Q9NVV4
B	43	MSE	-	EXPRESSION TAG	UNP Q9NVV4
B	151	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	152	UNK	-	SEE REMARK 999	UNP Q9NVV4

Continued on next page...

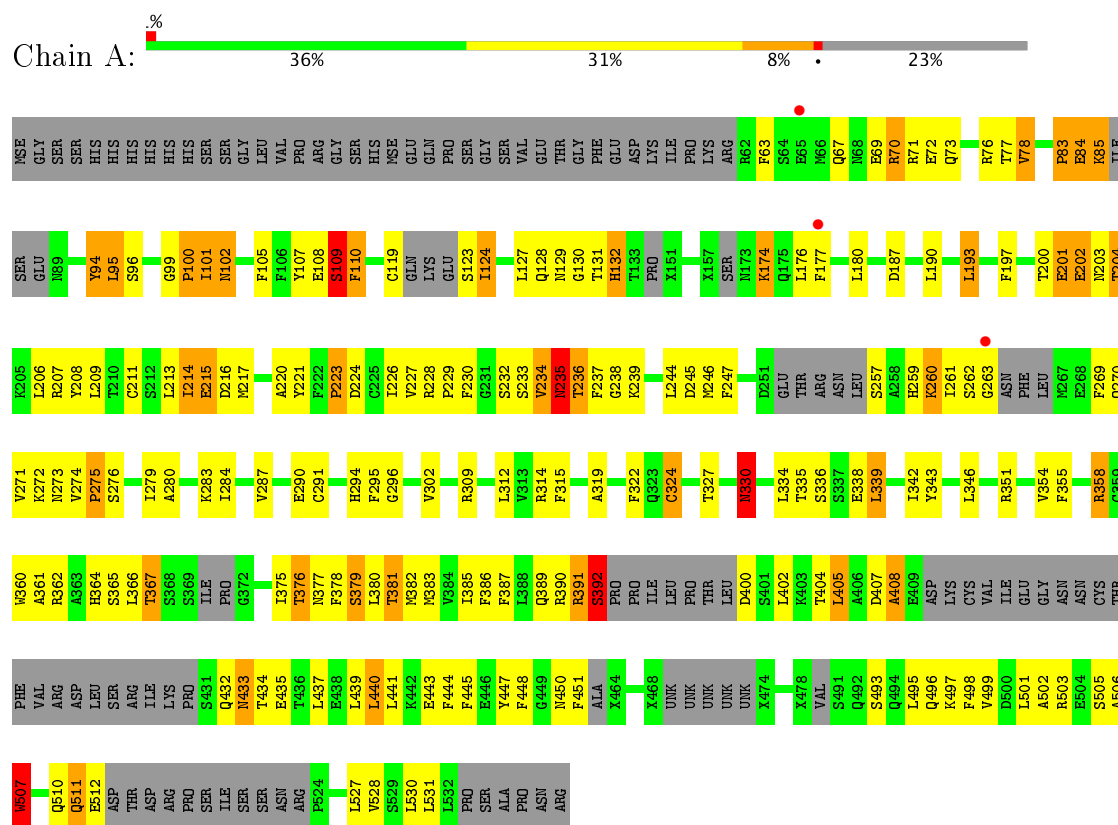
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	153	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	154	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	155	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	156	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	157	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	325	ALA	ASP	ENGINEERED MUTATION	UNP Q9NVV4
B	464	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	465	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	466	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	467	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	468	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	469	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	470	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	471	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	472	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	473	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	474	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	475	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	476	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	477	UNK	-	SEE REMARK 999	UNP Q9NVV4
B	478	UNK	-	SEE REMARK 999	UNP Q9NVV4

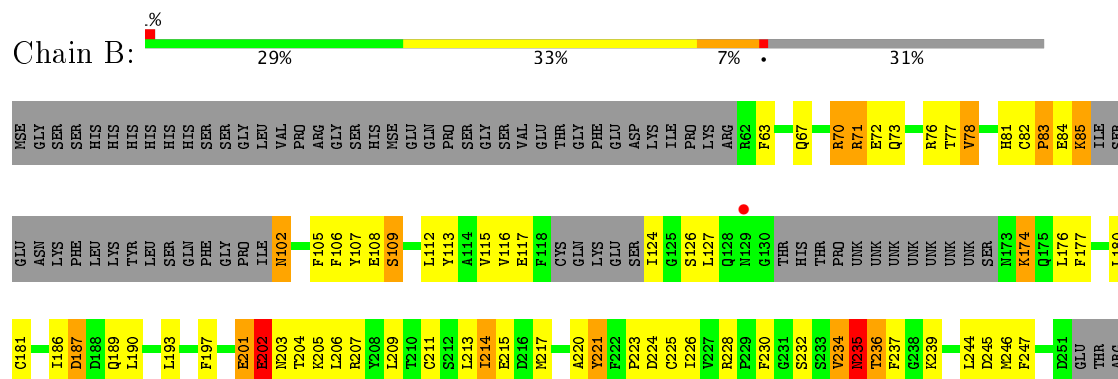
3 Residue-property plots

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: Poly(A) RNA polymerase



- Molecule 1: Poly(A) RNA polymerase



PHE	ALA	UNK	UNK	UNK	UNK	UNK	I469	X476	X477	X478	VAL	S491	Q492	S493	Q494	I495	Q496	X497	F498	V499	D500	I501	A502	B503	R507	I508	I509	O510	O511	B512	D515	ARG	PRO	SER	ILE	S520	S521	N522	R523	P524	I527	V528	S529	L530	L531	L532	PRO	SER	ALA	PRO	ASN	ARG					
1255	1256	S257	A258	K259	K260	I261	S262	S263	ASN	PHE	L266	N267	E268	F269	V271	K272	N273	V274	P275	S276	I279	A280	T281	Q282	K283	I284	L285	S286	V287	E290	H294	F295	G296	F297	G298	C299	V300	G301	V302	L306	N307	A308	R309	C310	F311	L312	N313	R314	F315	S316	R317	Q318	A319				
F322	Q323	C324	T327	R330	L334	T335	S336	S337	E338	L339	L340	Y341	I342	L346	D347	S348	R349	V350	R351	V354	F355	R358	C359	K360	A361	R362	A363	H364	SER	LEU	THR	SER	SER	ILE	PRO	SER	GLY	ALA	N374	I375	T376	N377	F378	S379	L380	T381	L382	N383	F384	I385	F386	F387	Q388	Q389			
R390	R391	S392	PRO	ILE	LEU	PRO	THR	LEU	ASP	SER	LEU	LYS	THR	LEU	ALA	ASP	ALA	GLU	ASP	LYS	CYS	VAL	ILE	GLU	GLY	ASN	ASN	CYS	THR	PHE	VAL	ARG	ASP	LEU	SER	ARG	ILE	LYS	PRO	SER	Q432	E435	T436	L437	E438	L439	L440	L441	K442	E443	F444	F445	E446	Y447	F448	G449	N450

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	68.68Å 76.88Å 87.63Å 90.00° 103.39° 90.00°	Depositor
Resolution (Å)	28.91 – 3.10 28.91 – 2.98	Depositor EDS
% Data completeness (in resolution range)	91.7 (28.91-3.10) 92.6 (28.91-2.98)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 3.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.247 , 0.328 0.255 , 0.336	Depositor DCC
R_{free} test set	772 reflections (4.91%)	DCC
Wilson B-factor (Å ²)	50.8	Xtriage
Anisotropy	0.261	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 57.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	5351	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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	0.53	0/2750	0.69	3/3690 (0.1%)
1	B	0.50	0/2544	0.69	1/3413 (0.0%)
All	All	0.51	0/5294	0.69	4/7103 (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	A	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	507	TRP	CA-CB-CG	10.60	133.84	113.70
1	A	507	TRP	CA-CB-CG	9.49	131.73	113.70
1	A	392	SER	N-CA-C	-5.20	96.96	111.00
1	A	507	TRP	N-CA-CB	-5.10	101.41	110.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	391	ARG	Peptide

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	2793	0	2688	202	0
1	B	2558	0	2478	195	0
All	All	5351	0	5166	355	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 34.

The worst 5 of 355 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:507:TRP:HZ3	1:B:511:GLN:OE1	1.32	1.10
1:A:131:THR:HA	1:B:267:MSE:O	1.55	1.03
1:A:84:GLU:O	1:A:85:LYS:HD3	1.64	0.97
1:A:507:TRP:CZ3	1:A:511:GLN:OE1	2.18	0.96
1:B:507:TRP:CZ3	1:B:511:GLN:OE1	2.18	0.96

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	316/464 (68%)	241 (76%)	53 (17%)	22 (7%)	1	8
1	B	291/464 (63%)	232 (80%)	44 (15%)	15 (5%)	2	14
All	All	607/928 (65%)	473 (78%)	97 (16%)	37 (6%)	2	11

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	94	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	100	PRO
1	A	101	ILE
1	A	102	ASN
1	A	201	GLU

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	300/388 (77%)	257 (86%)	43 (14%)	4	17
1	B	278/388 (72%)	236 (85%)	42 (15%)	3	15
All	All	578/776 (74%)	493 (85%)	85 (15%)	3	16

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	450	ASN
1	B	102	ASN
1	B	432	GLN
1	A	451	PHE
1	B	70	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	433	ASN
1	B	67	GLN
1	B	389	GLN
1	A	511	GLN
1	B	128	GLN

5.3.3 RNA ⓘ

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 carbohydrates 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	A	332/464 (71%)	-0.31	3 (0%) 84 69	17, 43, 83, 108	0
1	B	305/464 (65%)	-0.23	4 (1%) 77 59	17, 45, 85, 119	0
All	All	637/928 (68%)	-0.27	7 (1%) 80 65	17, 44, 85, 119	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	177	PHE	3.0
1	B	493	SER	2.8
1	A	263	GLY	2.4
1	B	374	TRP	2.4
1	B	491	SER	2.4

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.