



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 12, 2024 – 05:55 PM EDT

PDB ID : 3IKM
Title : Crystal structure of human mitochondrial DNA polymerase holoenzyme
Authors : Lee, Y.-S.; Kennedy, W.D.; Yin, Y.W.
Deposited on : 2009-08-06
Resolution : 3.24 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
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)	:	2.36.2

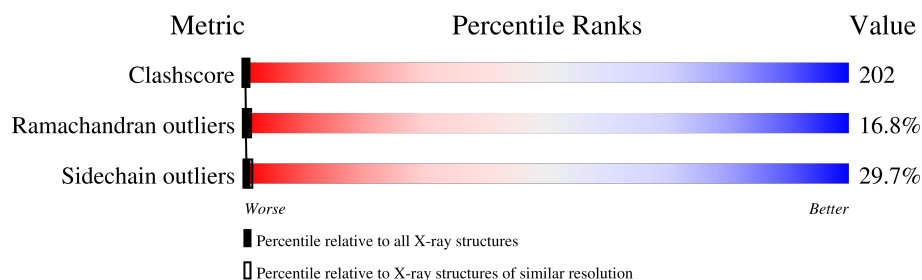
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.24 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))
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1172	
1	D	1172	
2	B	427	
2	C	427	
2	E	427	
2	F	427	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 29480 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 DNA polymerase subunit gamma-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1093	Total	C	N	O	S	0	0	0
			8681	5505	1537	1589	50			
1	D	1094	Total	C	N	O	S	0	0	0
			8695	5515	1540	1590	50			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	354	ASN	-	SEE REMARK 999	UNP P54098
A	355	SER	-	SEE REMARK 999	UNP P54098
D	354	ASN	-	SEE REMARK 999	UNP P54098
D	354A	SER	-	SEE REMARK 999	UNP P54098

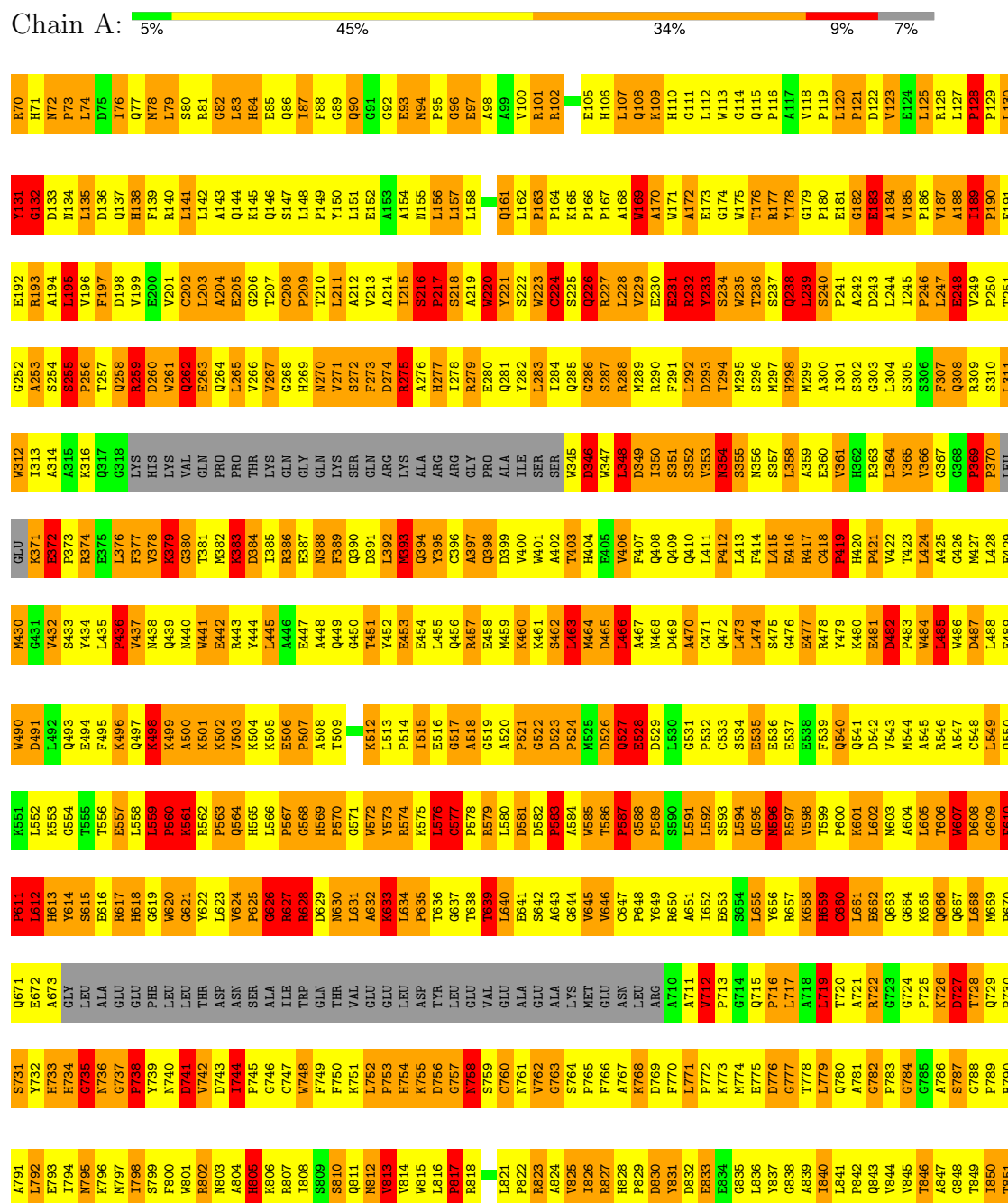
- Molecule 2 is a protein called DNA polymerase subunit gamma-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	355	Total	C	N	O	S	0	0	0
			2871	1840	504	511	16			
2	C	396	Total	C	N	O	S	0	0	0
			3181	2031	563	571	16			
2	E	355	Total	C	N	O	S	0	0	0
			2871	1840	504	511	16			
2	F	396	Total	C	N	O	S	0	0	0
			3181	2031	563	571	16			

3 Residue-property plots

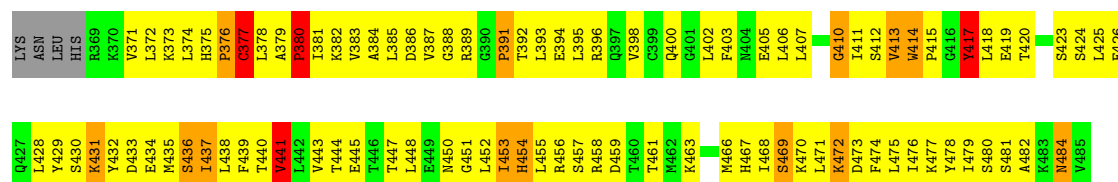
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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: DNA polymerase subunit gamma-1



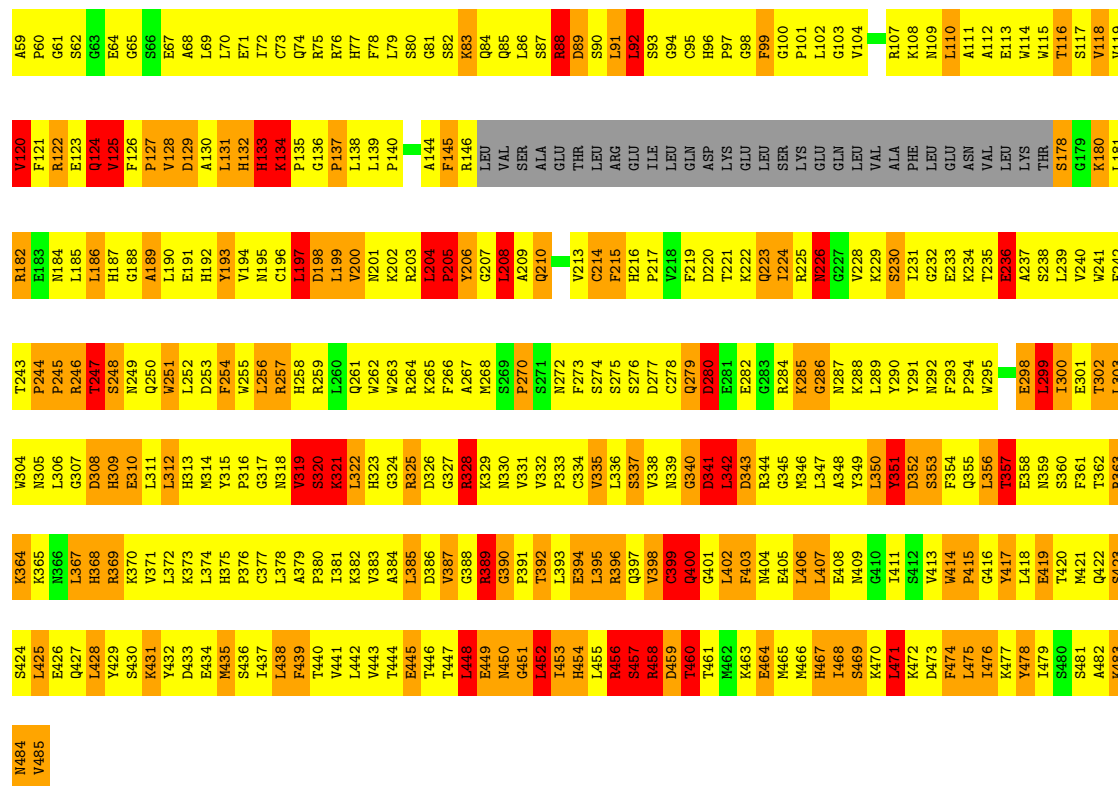
[illegible]





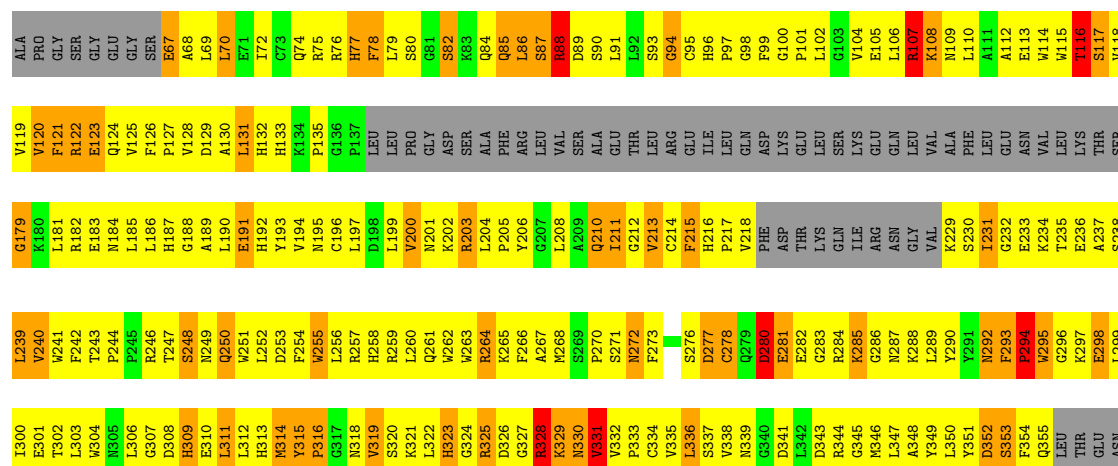
• Molecule 2: DNA polymerase subunit gamma-2

Chain C: 6% 54% 25% 8% 7%



• Molecule 2: DNA polymerase subunit gamma-2

Chain E: 6% 55% 19% 17%



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	138.39Å 138.39Å 226.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.92 – 3.24 46.92 – 3.25	Depositor EDS
% Data completeness (in resolution range)	89.5 (46.92-3.24) 89.5 (46.92-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.35 (at 3.25Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.284 , 0.303 0.362 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	91.9	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 64.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l 0.468 for h,-h-k,-l 0.023 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	29480	wwPDB-VP
Average B, all atoms (Å ²)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.11% 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

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	0.72	10/8913 (0.1%)	1.33	149/12100 (1.2%)
1	D	0.81	18/8927 (0.2%)	1.45	161/12118 (1.3%)
2	B	0.60	2/2944 (0.1%)	0.99	16/3981 (0.4%)
2	C	0.68	0/3262	1.31	47/4411 (1.1%)
2	E	0.67	6/2944 (0.2%)	1.11	23/3981 (0.6%)
2	F	0.64	0/3262	1.28	37/4411 (0.8%)
All	All	0.72	36/30252 (0.1%)	1.31	433/41002 (1.1%)

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	4
1	D	1	3
2	B	0	1
2	C	0	1
2	E	1	2
2	F	0	1
All	All	2	12

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	358	GLU	CG-CD	12.87	1.71	1.51
2	E	385	LEU	CA-CB	-11.49	1.27	1.53
1	D	352	SER	N-CA	9.45	1.65	1.46
1	A	232	ARG	N-CA	-8.85	1.28	1.46
2	E	382	LYS	N-CA	-8.69	1.28	1.46

The worst 5 of 433 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	382	LYS	N-CA-C	16.56	155.72	111.00
2	E	385	LEU	N-CA-C	16.34	155.11	111.00
1	D	248	GLU	CA-C-N	-14.80	84.64	117.20
2	B	280	ASP	C-N-CA	14.32	157.51	121.70
1	A	1212	ILE	C-N-CD	-12.99	92.03	120.60

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	D	249	VAL	CA
2	E	385	LEU	CA

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1179	PHE	Sidechain
1	A	221	TYR	Sidechain
1	A	239	LEU	Mainchain
1	A	395	TYR	Sidechain
2	B	417	TYR	Sidechain

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	8681	0	8521	3592	8
1	D	8695	0	8540	3900	7
2	B	2871	0	2862	809	0
2	C	3181	0	3168	1419	1
2	E	2871	0	2862	945	0
2	F	3181	0	3168	1467	4
All	All	29480	0	29121	11814	10

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 202.

The worst 5 of 11814 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:352:SER:CA	1:D:352:SER:C	1.74	1.55
1:D:460:LYS:HE2	1:D:460:LYS:N	1.20	1.48
2:E:432:TYR:HB2	2:E:437:ILE:CD1	1.51	1.39
1:D:352:SER:CA	1:D:358:GLU:OE2	1.73	1.36
1:D:914:THR:O	1:D:918:TRP:HB3	1.24	1.35

The worst 5 of 10 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:673:ALA:O	2:F:318:ASN:O[2_734]	1.91	0.29
1:A:673:ALA:C	2:F:318:ASN:O[2_734]	1.98	0.22
1:A:1034:ARG:NE	2:F:450:ASN:ND2[1_545]	2.04	0.16
1:A:994:TRP:NE1	1:D:522:GLY:N[1_545]	2.06	0.14
1:D:1232:GLU:OE2	2:F:223:GLN:OE1[2_844]	2.11	0.09

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	1085/1172 (93%)	629 (58%)	254 (23%)	202 (19%)	0	0
1	D	1086/1172 (93%)	645 (59%)	234 (22%)	207 (19%)	0	0
2	B	349/427 (82%)	205 (59%)	94 (27%)	50 (14%)	0	1
2	C	394/427 (92%)	254 (64%)	84 (21%)	56 (14%)	0	1
2	E	349/427 (82%)	219 (63%)	79 (23%)	51 (15%)	0	1
2	F	394/427 (92%)	251 (64%)	93 (24%)	50 (13%)	0	1
All	All	3657/4052 (90%)	2203 (60%)	838 (23%)	616 (17%)	0	0

5 of 616 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	ASN
1	A	76	ILE
1	A	96	GLY
1	A	163	PRO
1	A	169	TRP

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	918/988 (93%)	611 (67%)	307 (33%)	0	0
1	D	919/988 (93%)	584 (64%)	335 (36%)	0	0
2	B	317/380 (83%)	259 (82%)	58 (18%)	1	7
2	C	350/380 (92%)	260 (74%)	90 (26%)	0	1
2	E	317/380 (83%)	270 (85%)	47 (15%)	3	13
2	F	350/380 (92%)	245 (70%)	105 (30%)	0	1
All	All	3171/3496 (91%)	2229 (70%)	942 (30%)	0	1

5 of 942 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	87	ILE
2	F	299	LEU
1	D	479	TYR
2	F	251	TRP
2	E	131	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 119 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	110	HIS
2	F	187	HIS
1	D	456	GLN
2	F	84	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	454	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.