WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 06:33 pm BST

PDB ID	:	2GID
Title	:	Crystal structures of trypanosoma bruciei MRP1/MRP2
Authors	:	Schumacher, M.A.; Karamooz, E.; Zikova, A.; Trantirek, L.; Lukes, J.
Deposited on	:	2006-03-28
Resolution	:	3.35 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 following versions of software and data (see references ⓘ) were used in the production of this report:

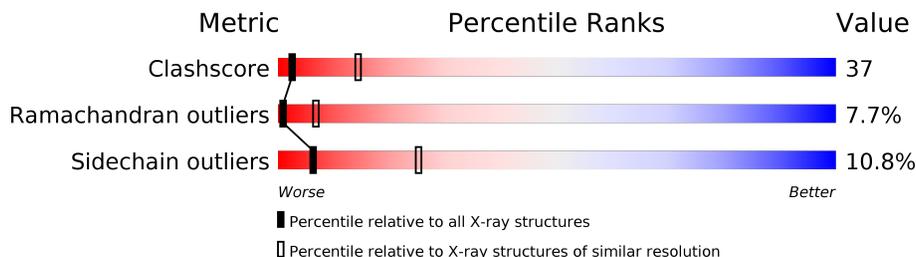|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.35 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	Percentile Ranks	Value
Clashscore		37
Ramachandran outliers		7.7%
Sidechain outliers		10.8%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	195	22% 46% 9% • 23%
1	G	195	32% 34% 11% • 23%
1	H	195	28% 42% 8% 23%
1	P	195	23% 48% 7% • 23%
2	B	187	26% 42% 10% • 22%
2	D	187	39% 35% 5% 22%
2	J	187	37% 37% • 22%
2	K	187	30% 40% 7% • 22%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9747 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 mitochondrial RNA-binding protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	150	Total	C	N	O	S	Se	0	0	0
			1224	774	223	223	1	3			
1	G	151	Total	C	N	O	S	Se	0	0	0
			1228	778	224	222	1	3			
1	H	151	Total	C	N	O	S	Se	0	0	0
			1231	779	224	224	1	3			
1	P	151	Total	C	N	O	S	Se	0	0	0
			1228	778	224	222	1	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
A	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
A	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
G	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
G	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
G	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
H	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
H	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
H	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
P	110	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
P	125	MSE	MET	MODIFIED RESIDUE	UNP Q952G2
P	204	MSE	MET	MODIFIED RESIDUE	UNP Q952G2

- Molecule 2 is a protein called mitochondrial RNA-binding protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	146	Total	C	N	O	S	Se	0	0	0
			1195	754	220	217	1	3			
2	D	146	Total	C	N	O	S	Se	0	0	0
			1195	754	220	217	1	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	K	146	Total 1195	C 754	N 220	O 217	S 1	Se 3	0	0	0
2	J	146	Total 1195	C 754	N 220	O 217	S 1	Se 3	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	43	GLU	LEU	CONFLICT	UNP P90629
B	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
B	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
B	139	MSE	MET	MODIFIED RESIDUE	UNP P90629
D	43	GLU	LEU	CONFLICT	UNP P90629
D	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
D	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
D	139	MSE	MET	MODIFIED RESIDUE	UNP P90629
K	43	GLU	LEU	CONFLICT	UNP P90629
K	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
K	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
K	139	MSE	MET	MODIFIED RESIDUE	UNP P90629
J	43	GLU	LEU	CONFLICT	UNP P90629
J	46	MSE	MET	MODIFIED RESIDUE	UNP P90629
J	109	MSE	MET	MODIFIED RESIDUE	UNP P90629
J	139	MSE	MET	MODIFIED RESIDUE	UNP P90629

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total 4	O 4	0	0
3	B	11	Total 11	O 11	0	0
3	D	6	Total 6	O 6	0	0
3	G	1	Total 1	O 1	0	0
3	H	4	Total 4	O 4	0	0
3	K	8	Total 8	O 8	0	0
3	J	7	Total 7	O 7	0	0

Continued on next page...

Continued from previous page...

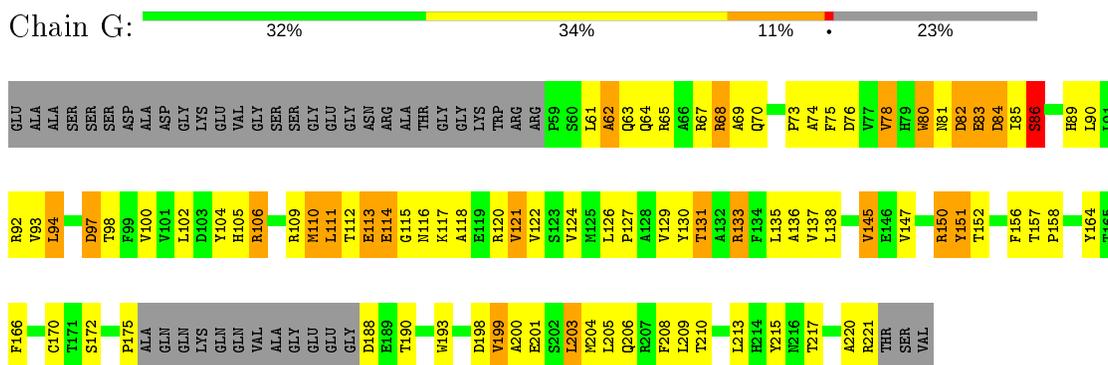| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	15	Total 15	O 15	0	0

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. 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.
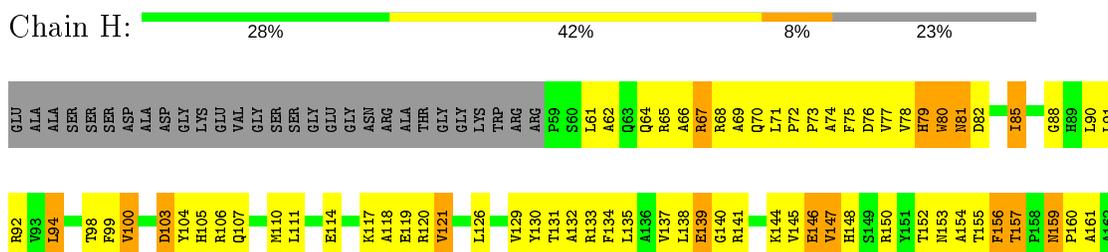
Note EDS was not executed.
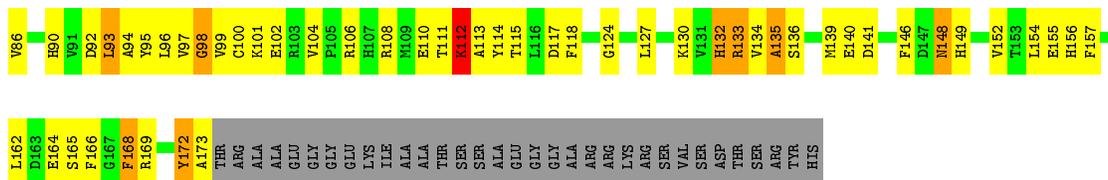
• Molecule 1: mitochondrial RNA-binding protein 2

Chain A:
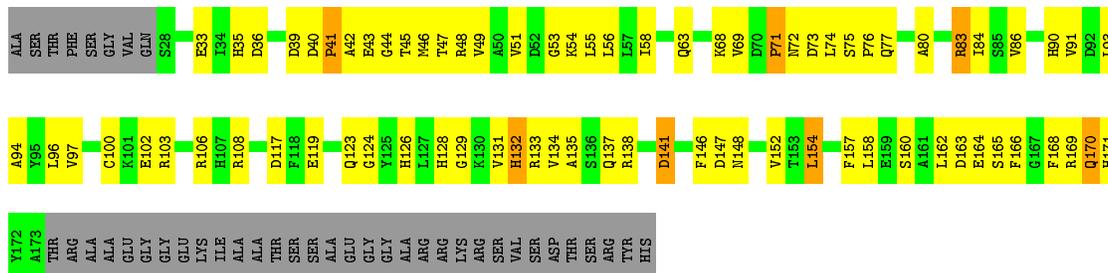

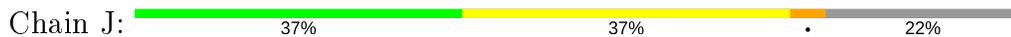• Molecule 1: mitochondrial RNA-binding protein 2

Chain G:

• Molecule 1: mitochondrial RNA-binding protein 2

Chain H:

P163 Y164 T165 F166 T167 L168 K169 C170 T171 S172 P175 ALA GLN GLN LYS GLN GLN VAL ALA GLY GLU GLU GLY D188 F191 E192 W193 T194 V195 A200 E201 S202 L203 M204 L205 Q206 R207 Q211 A212 L213 H214 Y215 N216 T217 G218 F219 A220 R221 THR SER VAL

• Molecule 1: mitochondrial RNA-binding protein 2

Chain P: 23% 48% 7% • 23%

GLU ALA ALA SER SER ASP ALA ASP GLY LYS GLU VAL GLY SER SER GLY GLU GLY ASN ARG ALA THR GLY GLY LYS TRP ARG ARG P59 S60 Q63 Q64 R65 A66 R67 R68 A69 Q70 L71 P72 P73 A74 F75 D76 V77 V78 H79 W80 N81 D82 E83 D84 I85 S86 H89 L90 L91

R92 V93 L94 H95 R96 D97 V100 V101 L102 R106 Q107 A108 R109 M110 L111 T112 E113 E114 G115 N116 K117 A118 E119 R120 V121 V122 S123 V124 M125 L126 P127 Y130 R133 F134 L135 A136 V137 E143 K144 V145 E146 V147 R150 Y151 T152 T155 F156 T157 P158 N159 P160 A161 A162

P163 Y164 T165 F166 T167 L168 K169 C170 T171 S172 T173 R174 P175 ALA GLN GLN LYS GLN GLN VAL ALA GLY GLU GLU GLY D188 E189 E192 W193 E196 F197 V199 D198 A200 E201 S202 L203 M204 L205 Q206 R207 F208 L209 T210 Q211 A212 L213 H214 Y215 N216 A220 R221 THR SER VAL

• Molecule 2: mitochondrial RNA-binding protein 1

Chain B: 26% 42% 10% • 22%

ALA SER THR PHE SER GLY VAL GLN S28 L29 P30 K31 F32 E33 I34 H35 D36 D39 D40 P41 A42 E43 G44 T45 M46 T47 R48 V49 A50 V51 K54 L55 L56 L57 Y61 P62 Q63 L64 G65 P66 R67 K68 V69 D70 D73 Q77 F78 D79 A80 D81 R82 R83 R89 H90

V91 D92 L93 A94 Y95 L96 V97 G98 V99 C100 K101 E102 R103 V104 P105 R106 H107 R108 M109 E110 T111 K112 A113 Y114 T115 L116 D117 F118 E119 K120 S121 G124 L127 H128 G129 K130 V131 V134 A135 S136 V144 K145 F146 D147 N148 H149 F150 A151 V152 T153 L154 E155 H156 F157 L158 E159

S160 A161 L162 D163 E164 S165 F166 G167 F168 R169 Q170 H171 Y172 A173 THR ARG ALA ALA GLU GLY GLY GLU GLY LYS ILE ALA ALA THR SER SER ALA GLU GLY GLY ALA ARG ARG LYS ARG SER VAL ASP THR SER ARG TYR HIS

• Molecule 2: mitochondrial RNA-binding protein 1

Chain D: 39% 35% 5% 22%

ALA SER THR PHE SER GLY VAL GLN S28 F32 H35 R38 D39 D40 P41 A42 E43 G44 T45 M46 T47 R48 V49 A50 V51 K54 L55 S59 Q60 Y61 P62 Q63 R67 K68 V69 D70 L74 S75 P76 Q77 F78 D81 R82 R83 I84 S85 V86 R87 L88 R89 H90 V91

D92 L93 A94 Y95 L96 V99 R103 V104 P105 R106 H107 R108 M109 E110 T111 Y114 T115 L116 D117 F118 E119 K120 S121 A122 Q123 G124 Y125 H126 L127 H128 G129 K130 V131 H132 R133 W142 S143 V144 D147 N148 H149 F150 T153 L154 E155 H156 F157 L162 D163 E164 S165 F168

R169 A173 THR ARG ALA ALA GLU GLY GLY GLU GLY LYS ILE ALA ALA THR SER SER ALA GLU GLY GLY ALA ARG ARG LYS ARG SER VAL SER ASP THR SER ARG TYR HIS

• Molecule 2: mitochondrial RNA-binding protein 1

Chain K: 30% 40% 7% • 22%

ALA SER THR PHE SER GLY VAL GLN S28 L29 P30 K31 F32 D36 D39 D40 P41 A42 E43 G44 T45 M46 T47 R48 V49 A50 V51 D52 G53 K54 L55 L56 L57 I58 P62 Q63 L64 G65 P66 R67 K68 V69 D70 P71 N72 D73 L74 S75 P76 Q77 F78 D79 R82 R83 I84 S85

● Molecule 2: mitochondrial RNA-binding protein 1

Chain J:

37% 37% 22%

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	235.99Å 235.99Å 85.45Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	78.84 – 3.35	Depositor
% Data completeness (in resolution range)	99.2 (78.84-3.35)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.233 , 0.299	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9747	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

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	0.44	0/1250	0.73	0/1693
1	G	0.44	0/1255	0.70	1/1700 (0.1%)
1	H	0.45	0/1258	0.72	0/1704
1	P	0.47	0/1255	0.72	0/1700
2	B	0.47	0/1222	0.79	1/1645 (0.1%)
2	D	0.49	0/1222	0.74	0/1645
2	J	0.48	0/1222	0.76	0/1645
2	K	0.49	0/1222	0.77	0/1645
All	All	0.47	0/9906	0.74	2/13377 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	29	LEU	CA-CB-CG	9.37	136.84	115.30
1	G	86	SER	N-CA-C	-5.39	96.45	111.00

There are no chirality outliers.

There are no planarity outliers.

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	1224	0	1195	120	0
1	G	1228	0	1201	89	0
1	H	1231	0	1203	103	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	P	1228	0	1201	120	0
2	B	1195	0	1164	99	0
2	D	1195	0	1164	65	0
2	J	1195	0	1164	72	0
2	K	1195	0	1164	86	0
3	A	4	0	0	0	0
3	B	11	0	0	0	0
3	D	6	0	0	0	0
3	G	1	0	0	0	0
3	H	4	0	0	0	0
3	J	7	0	0	0	0
3	K	8	0	0	0	0
3	P	15	0	0	0	0
All	All	9747	0	9456	711	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 37.

The worst 5 of 711 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:H:90:LEU:HD13	1:H:91:LEU:N	1.71	1.05
1:G:100:VAL:HB	1:G:126:LEU:HB2	1.44	0.99
1:G:113:GLU:HG3	1:G:114:GLU:H	1.27	0.99
1:A:82:ASP:HA	1:A:111:LEU:HD22	1.45	0.99
1:H:90:LEU:HD13	1:H:91:LEU:H	1.19	0.98

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	146/195 (75%)	104 (71%)	32 (22%)	10 (7%)	1	9
1	G	147/195 (75%)	111 (76%)	19 (13%)	17 (12%)	0	2
1	H	147/195 (75%)	98 (67%)	32 (22%)	17 (12%)	0	2
1	P	147/195 (75%)	105 (71%)	33 (22%)	9 (6%)	1	11
2	B	144/187 (77%)	111 (77%)	24 (17%)	9 (6%)	1	10
2	D	144/187 (77%)	121 (84%)	18 (12%)	5 (4%)	3	23
2	J	144/187 (77%)	116 (81%)	20 (14%)	8 (6%)	2	12
2	K	144/187 (77%)	102 (71%)	28 (19%)	14 (10%)	0	4
All	All	1163/1528 (76%)	868 (75%)	206 (18%)	89 (8%)	1	7

5 of 89 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	87	ARG
1	A	113	GLU
1	A	128	ALA
1	A	219	PHE
2	B	70	ASP

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	131/159 (82%)	115 (88%)	16 (12%)	5	19
1	G	131/159 (82%)	116 (88%)	15 (12%)	5	22
1	H	132/159 (83%)	118 (89%)	14 (11%)	6	26
1	P	131/159 (82%)	119 (91%)	12 (9%)	9	32
2	B	129/155 (83%)	112 (87%)	17 (13%)	4	17
2	D	129/155 (83%)	114 (88%)	15 (12%)	5	21
2	J	129/155 (83%)	121 (94%)	8 (6%)	18	49
2	K	129/155 (83%)	114 (88%)	15 (12%)	5	21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1041/1256 (83%)	929 (89%)	112 (11%)	6 25

5 of 112 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	78	VAL
1	G	213	LEU
1	P	106	ARG
1	G	82	ASP
1	G	133	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	171	HIS
1	G	206	GLN
1	P	64	GLN
1	G	64	GLN
1	G	70	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 (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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.