Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 08:20 pm BST

PDB ID	:	6QPQ
Title	:	The structure of the cohesin head module elucidates the mechanism of ring opening
Authors	:	Li, Y.; Muir, K.W.; Panne, D.
Deposited on	:	2019-02-14
Resolution	:	2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

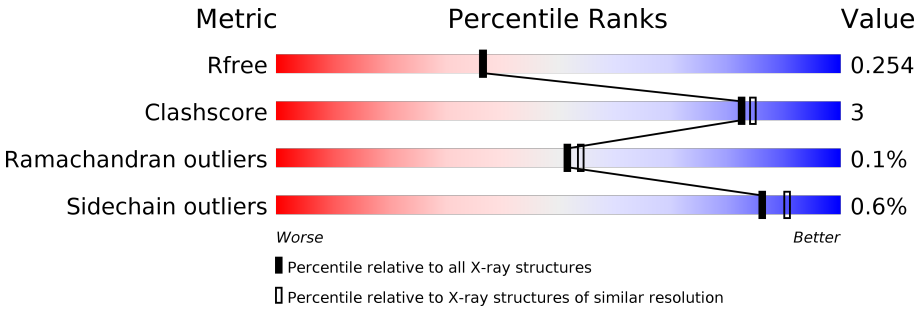
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.10 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}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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%

Mol	Chain	Length	Quality of chain
1	A	453	83% 5% 11%
1	C	453	79% 7% 13%
2	B	566	14% · 86%
2	D	566	14% · 86%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15820 atoms, of which 7723 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 Structural maintenance of chromosomes protein,Structural maintenance of chromosomes protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	402	Total	C	H	N	O	S	0	0	0
			6512	2068	3239	578	619	8			
1	C	392	Total	C	H	N	O	S	0	1	0
			6362	2021	3163	565	605	8			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	243	SER	-	linker	UNP G0SGH3
A	244	PRO	-	linker	UNP G0SGH3
A	245	GLY	-	linker	UNP G0SGH3
A	246	LEU	-	linker	UNP G0SGH3
A	247	GLU	-	linker	UNP G0SGH3
A	248	VAL	-	linker	UNP G0SGH3
A	249	LEU	-	linker	UNP G0SGH3
A	250	PHE	-	linker	UNP G0SGH3
A	1265	HIS	-	expression tag	UNP G0SGH3
A	1266	HIS	-	expression tag	UNP G0SGH3
C	1056	SER	-	linker	UNP G0SGH3
C	1057	PRO	-	linker	UNP G0SGH3
C	1058	GLY	-	linker	UNP G0SGH3
C	1059	LEU	-	linker	UNP G0SGH3
C	1060	GLU	-	linker	UNP G0SGH3
C	1061	VAL	-	linker	UNP G0SGH3
C	1062	LEU	-	linker	UNP G0SGH3
C	1063	PHE	-	linker	UNP G0SGH3
C	1265	HIS	-	expression tag	UNP G0SGH3
C	1266	HIS	-	expression tag	UNP G0SGH3

- Molecule 2 is a protein called Sister chromatid cohesion protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	81	Total	C	H	N	O	S	0	0	0
			1292	405	655	107	123	2			
2	D	82	Total	C	H	N	O	S	0	0	0
			1311	411	666	108	124	2			

- Molecule 3 is water.

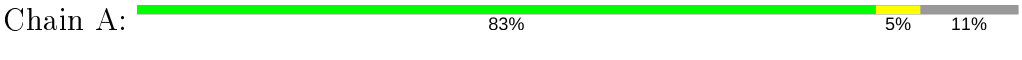
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	170	Total	O	0	0
			170	170		
3	B	13	Total	O	0	0
			13	13		
3	C	139	Total	O	0	0
			139	139		
3	D	21	Total	O	0	0
			21	21		

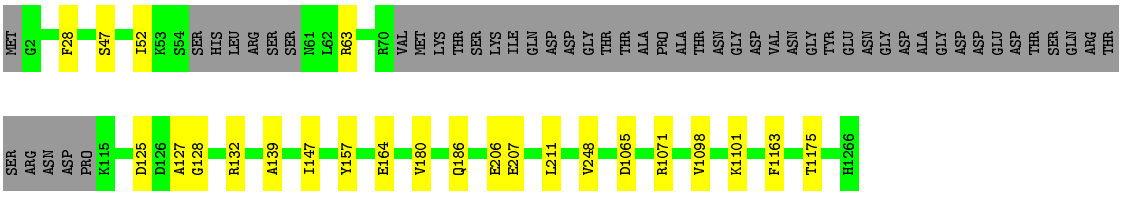
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Structural maintenance of chromosomes protein,Structural maintenance of chromosomes protein

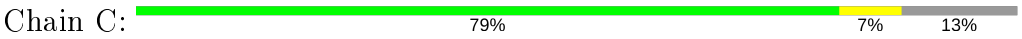
Chain A:

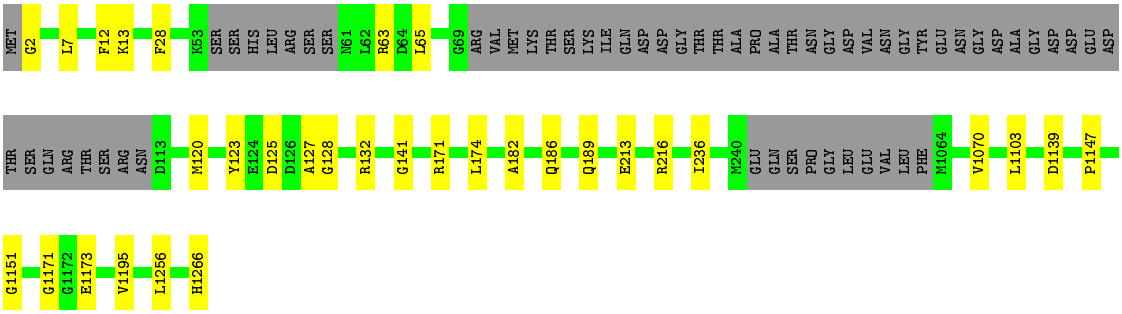




• Molecule 1: Structural maintenance of chromosomes protein,Structural maintenance of chromosomes protein

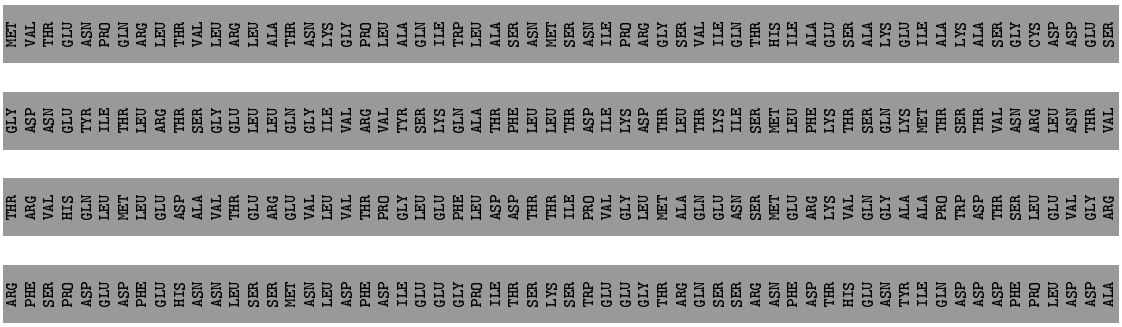
Chain C:

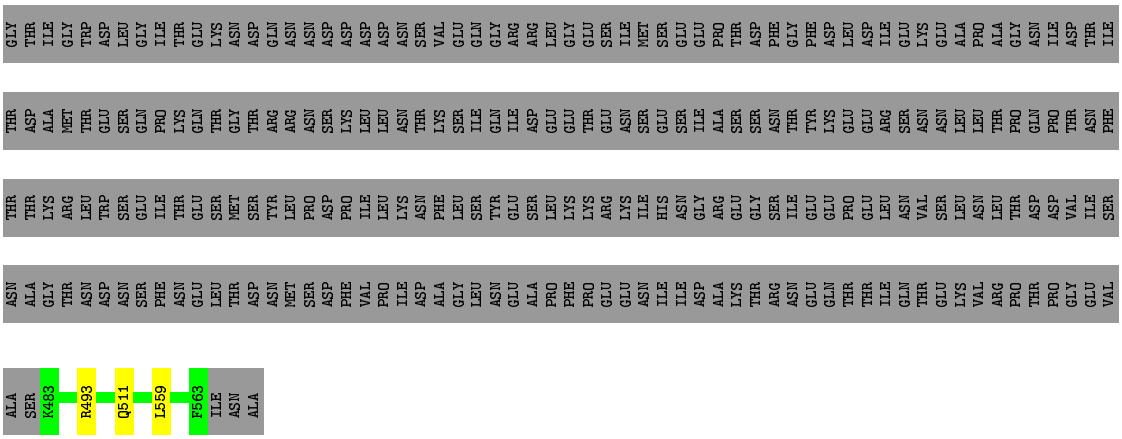




• Molecule 2: Sister chromatid cohesion protein 1

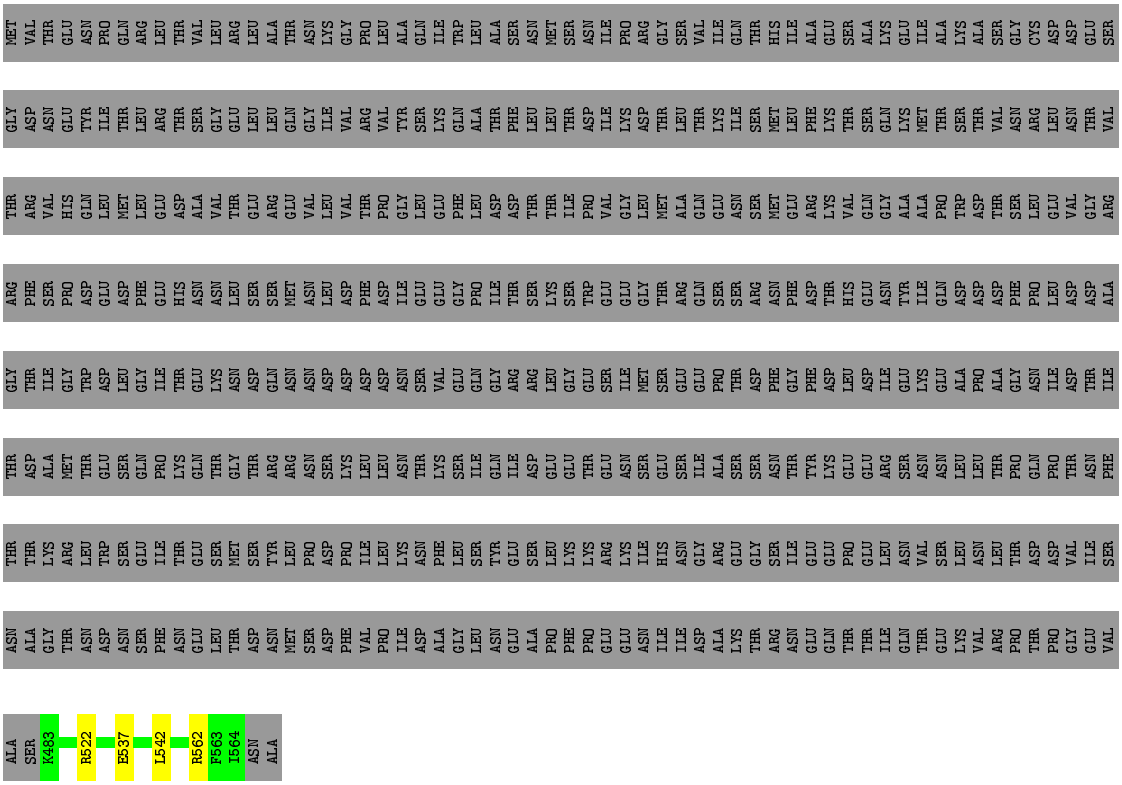
Chain B:





- Molecule 2: Sister chromatid cohesion protein 1

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	80.75Å 111.13Å 166.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.70 – 2.10 45.78 – 2.09	Depositor EDS
% Data completeness (in resolution range)	98.8 (45.70-2.10) 93.8 (45.78-2.09)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 2.10Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	(Not available) , (Not available) 0.223 , 0.254	Depositor DCC
R_{free} test set	1995 reflections (2.27%)	wwPDB-VP
Wilson B-factor (Å^2)	43.1	Xtriage
Anisotropy	0.216	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 45.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	15820	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.04% 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	0.26	0/3333	0.44	0/4481
1	C	0.26	0/3261	0.45	0/4385
2	B	0.25	0/644	0.43	0/862
2	D	0.26	0/652	0.44	0/873
All	All	0.26	0/7890	0.45	0/10601

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	3273	3239	3239	11	1
1	C	3199	3163	3165	24	0
2	B	637	655	655	3	0
2	D	645	666	666	3	0
3	A	170	0	0	3	0
3	B	13	0	0	2	0
3	C	139	0	0	5	0
3	D	21	0	0	0	0
All	All	8097	7723	7725	39	1

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 3.

All (39) 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:C:171:ARG:NH1	3:C:1303:HOH:O	2.18	0.76
1:C:1139:ASP:OD1	3:C:1301:HOH:O	2.05	0.74
1:C:1173:GLU:OE1	3:C:1302:HOH:O	2.06	0.73
1:A:164:GLU:OE2	3:A:1301:HOH:O	2.08	0.72
1:A:63:ARG:NH1	1:A:139:ALA:O	2.27	0.68
1:A:186:GLN:OE1	3:A:1302:HOH:O	2.12	0.66
1:C:189:GLN:OE1	1:C:1151:GLY:N	2.33	0.62
1:A:127:ALA:N	1:A:128:GLY:HA2	2.18	0.59
1:C:236:ILE:HD11	1:C:1070:VAL:CG2	2.33	0.58
1:C:127:ALA:N	1:C:128:GLY:HA2	2.18	0.58
1:C:7:LEU:HD12	1:C:120:MET:O	2.08	0.53
1:C:132:ARG:NH1	3:C:1308:HOH:O	2.34	0.53
1:C:182:ALA:O	1:C:186[B]:GLN:HG3	2.09	0.52
1:C:236:ILE:HD11	1:C:1070:VAL:HG21	1.92	0.52
1:C:1256:LEU:HD11	2:D:542:LEU:HD13	1.92	0.51
2:D:537:GLU:OE1	2:D:562:ARG:NH2	2.44	0.49
1:C:236:ILE:HD11	1:C:1070:VAL:CG1	2.44	0.48
1:C:174:LEU:C	1:C:174:LEU:HD12	2.34	0.47
2:B:511:GLN:NE2	3:B:602:HOH:O	2.47	0.47
1:C:132:ARG:NH2	3:C:1309:HOH:O	2.42	0.47
1:A:132:ARG:NH2	3:A:1309:HOH:O	2.49	0.46
1:A:147:ILE:HD11	1:A:157:TYR:HA	1.97	0.45
1:C:213:GLU:OE2	1:C:216:ARG:NH2	2.44	0.45
2:B:493:ARG:NH1	2:B:559:LEU:O	2.44	0.45
1:A:125:ASP:OD1	1:A:128:GLY:HA2	2.18	0.44
1:C:63:ARG:HG2	1:C:141:GLY:CA	2.48	0.43
1:C:125:ASP:OD1	1:C:128:GLY:HA2	2.18	0.43
1:C:1103:LEU:HD11	1:C:1147:PRO:HD2	2.00	0.43
1:A:47:SER:HB2	1:A:52:ILE:HB	2.02	0.42
1:C:174:LEU:HB2	1:C:1195:VAL:HB	2.01	0.42
1:C:2:GLY:N	1:C:123:TYR:HH	2.19	0.41
1:A:180:VAL:HG11	1:A:1175:THR:HG23	2.03	0.41
1:A:206:GLU:HG3	1:A:207:GLU:N	2.35	0.41
2:B:511:GLN:NE2	3:B:601:HOH:O	2.51	0.41
1:C:174:LEU:CB	1:C:1195:VAL:HB	2.50	0.41
1:A:211:LEU:HD12	1:A:1098:VAL:HG11	2.03	0.40
1:C:12:PHE:CE2	1:C:13:LYS:HG3	2.56	0.40
1:C:1266:HIS:CG	2:D:522:ARG:HH12	2.39	0.40
1:C:13:LYS:HD3	1:C:65:LEU:HD22	2.03	0.40

All (1) 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:248:VAL:O	1:A:1071:ARG:NH1[2_556]	2.17	0.03

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	394/453 (87%)	384 (98%)	10 (2%)	0	100 100
1	C	385/453 (85%)	377 (98%)	7 (2%)	1 (0%)	41 41
2	B	79/566 (14%)	79 (100%)	0	0	100 100
2	D	80/566 (14%)	79 (99%)	1 (1%)	0	100 100
All	All	938/2038 (46%)	919 (98%)	18 (2%)	1 (0%)	51 54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	1171	GLY

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	348/391 (89%)	344 (99%)	4 (1%)	73 79
1	C	340/391 (87%)	339 (100%)	1 (0%)	92 95

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	B	69/506 (14%)	69 (100%)	0	100 100
2	D	70/506 (14%)	70 (100%)	0	100 100
All	All	827/1794 (46%)	822 (99%)	5 (1%)	86 90

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	28	PHE
1	A	1065	ASP
1	A	1101	LYS
1	A	1163	PHE
1	C	28	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	250:PHE	C	1064:MET	N	8.47

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

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

6.4 Ligands

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

6.5 Other polymers

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