wwPDB X-ray Structure Validation Summary Report (i)

Jun 22, 2024 – 06:34 PM EDT

PDB ID	:	5CJS
Title	:	Crystal structure of a monomeric influenza hemagglutinin stem in complex with an broadly neutralizing antibody CR9114
Authors	:	Zhu, X.; Wilson, I.A.
Deposited on	:	2015-07-15
Resolution	:	4.30 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 (i) 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 (i)) were used in the production of this report:

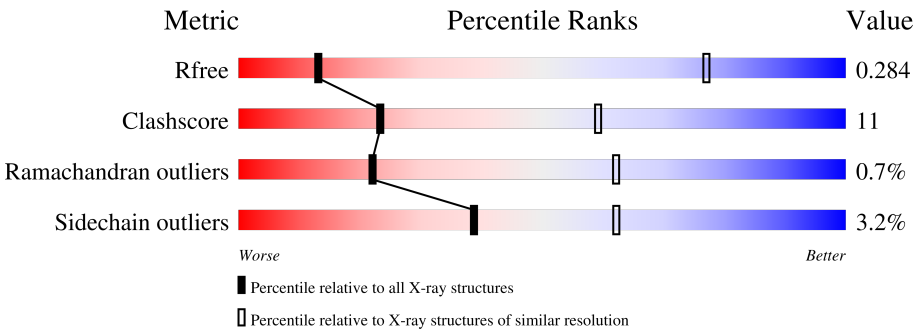
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 4.30 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	1014 (4.80-3.80)
Clashscore	141614	1077 (4.80-3.80)
Ramachandran outliers	138981	1029 (4.80-3.80)
Sidechain outliers	138945	1012 (4.80-3.80)

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 <=5%

Mol	Chain	Length	Quality of chain
1	E	215	77% 18% • •
1	L	215	77% 20% • •
2	F	230	73% 21% 6%
2	H	230	73% 18% • • 6%
3	C	62	44% 37% • 18%
3	J	62	56% 19% 8% 16%
4	D	191	42% 29% • • 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	K	191	59% 19% • 18%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9576 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 CR9114 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	211	1568	978	266	320	4	0	0	0
1	E	211	1568	978	266	320	4	0	0	0

- Molecule 2 is a protein called CR9114 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	217	1613	1017	269	320	7	0	0	0
2	F	217	1613	1017	269	320	7	0	0	0

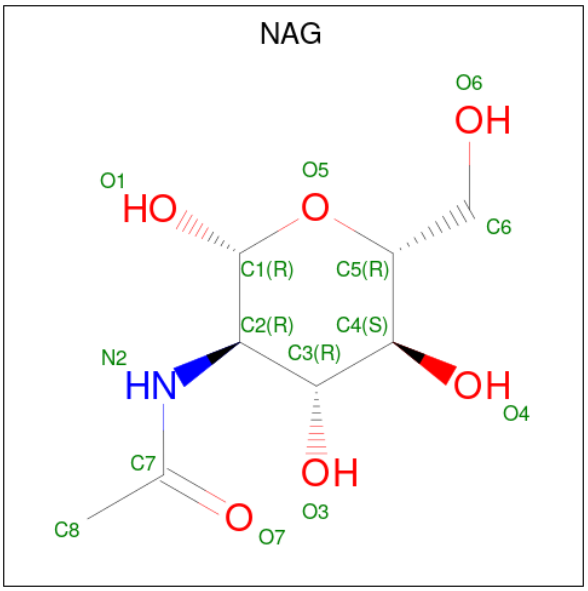
- Molecule 3 is a protein called Designed influenza hemagglutinin stem #4454, HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	51	387	235	70	79	3	0	0	0
3	J	52	397	242	72	80	3	0	0	0

- Molecule 4 is a protein called Designed influenza hemagglutinin stem #4454, HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	143	1154	717	195	235	7	0	0	0
4	K	156	1234	766	209	252	7	0	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



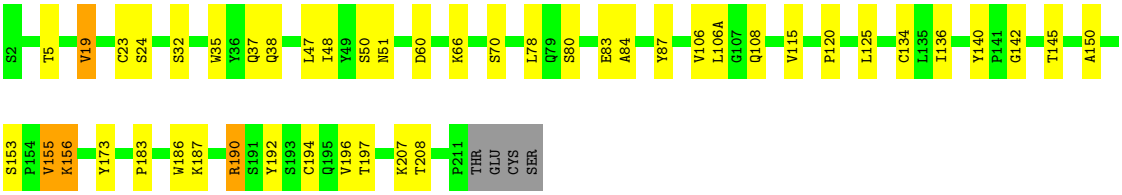
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	C	1	14	8	1	5	0	0
			Total	C	N	O		
5	J	1	14	8	1	5	0	0
			Total	C	N	O		
5	K	1	14	8	1	5	0	0

3 Residue-property plots (i)

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.

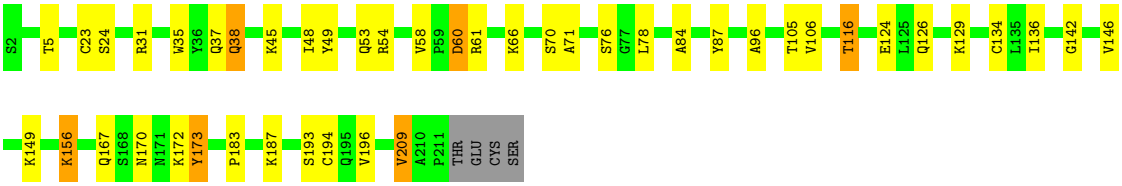
- Molecule 1: CR9114 light chain

Chain L:



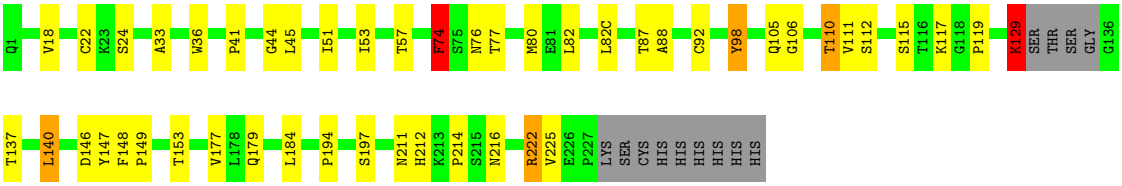
- Molecule 1: CR9114 light chain

Chain E:



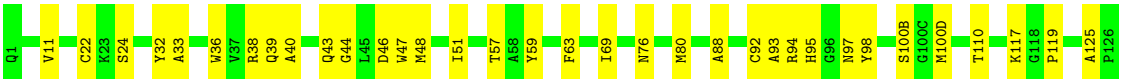
- Molecule 2: CR9114 heavy chain

Chain H:



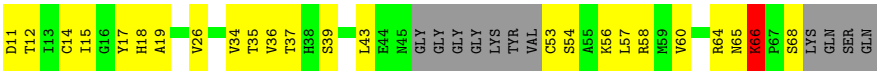
- Molecule 2: CR9114 heavy chain

Chain F:



• Molecule 3: Designed influenza hemagglutinin stem #4454, HA1

Chain C:

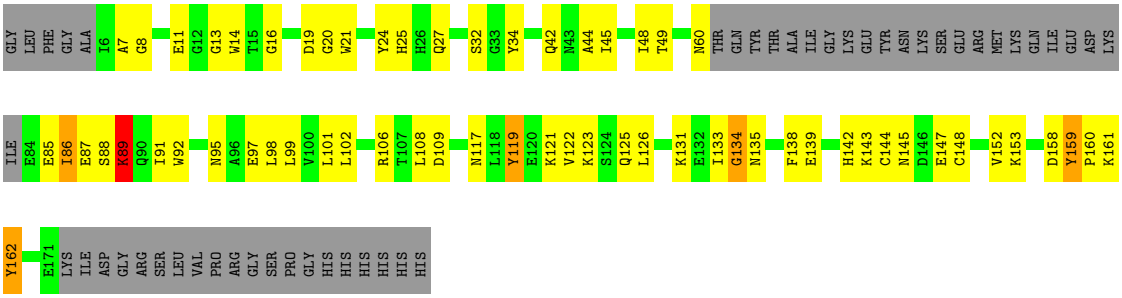


• Molecule 3: Designed influenza hemagglutinin stem #4454, HA1

Chain J:

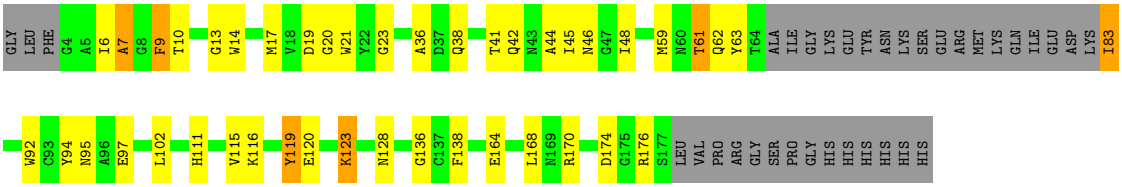
• Molecule 4: Designed influenza hemagglutinin stem #4454, HA2

Chain D:



• Molecule 4: Designed influenza hemagglutinin stem #4454, HA2

Chain K:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	110.87Å 110.87Å 359.16Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.01 – 4.30 48.01 – 4.29	Depositor EDS
% Data completeness (in resolution range)	99.5 (48.01-4.30) 99.6 (48.01-4.29)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.21	Depositor
$< I/\sigma(I) >$ [1]	1.96 (at 4.29Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.262 , 0.303 0.277 , 0.284	Depositor DCC
R_{free} test set	933 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	147.6	Xtriage
Anisotropy	0.615	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 165.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.069 for -h,-k,l	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	9576	wwPDB-VP
Average B, all atoms (Å^2)	179.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.15% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	E	0.34	0/1606	0.71	2/2193 (0.1%)
1	L	0.39	0/1606	0.85	6/2193 (0.3%)
2	F	0.38	0/1652	0.69	0/2251
2	H	0.47	1/1652 (0.1%)	0.81	5/2251 (0.2%)
3	C	0.51	0/389	1.08	2/527 (0.4%)
3	J	0.47	0/399	0.98	2/539 (0.4%)
4	D	0.61	2/1176 (0.2%)	1.08	10/1581 (0.6%)
4	K	0.50	0/1257	0.92	3/1692 (0.2%)
All	All	0.45	3/9737 (0.0%)	0.85	30/13227 (0.2%)

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
4	D	0	1
4	K	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	129	LYS	CE-NZ	8.88	1.71	1.49
4	D	143	LYS	CE-NZ	6.59	1.65	1.49
4	D	11	GLU	CG-CD	5.08	1.59	1.51

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	L	190	ARG	NE-CZ-NH2	-11.53	114.53	120.30
4	K	83	ILE	CG1-CB-CG2	11.24	136.12	111.40
1	L	19	VAL	CG1-CB-CG2	8.04	123.76	110.90
4	D	131	LYS	CB-CA-C	-7.86	94.68	110.40
1	E	209	VAL	CG1-CB-CG2	7.55	122.97	110.90

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	159	TYR	Sidechain
4	K	59	MET	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	E	1568	0	1521	29	1
1	L	1568	0	1521	26	0
2	F	1613	0	1572	32	0
2	H	1613	0	1572	34	1
3	C	387	0	386	20	0
3	J	397	0	401	11	0
4	D	1154	0	1073	52	0
4	K	1234	0	1134	33	0
5	C	14	0	13	0	0
5	J	14	0	13	0	0
5	K	14	0	13	0	0
All	All	9576	0	9219	212	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 11.

The worst 5 of 212 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)
2:H:129:LYS:CE	2:H:129:LYS:NZ	1.71	1.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:133:ILE:HD11	4:D:139:GLU:HB2	1.45	0.94
3:J:26:VAL:HG21	3:J:60:VAL:HG22	1.54	0.87
1:E:142:GLY:HA3	1:E:173:TYR:CD2	2.17	0.80
3:C:26:VAL:HG21	3:C:60:VAL:HG22	1.65	0.78

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)
2:H:115:SER:O	1:E:126:GLN:NE2[2_454]	2.05	0.15

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	E	209/215 (97%)	203 (97%)	6 (3%)	0	100	100
1	L	209/215 (97%)	203 (97%)	6 (3%)	0	100	100
2	F	213/230 (93%)	208 (98%)	4 (2%)	1 (0%)	29	68
2	H	213/230 (93%)	207 (97%)	5 (2%)	1 (0%)	29	68
3	C	47/62 (76%)	42 (89%)	5 (11%)	0	100	100
3	J	48/62 (77%)	45 (94%)	1 (2%)	2 (4%)	3	25
4	D	139/191 (73%)	138 (99%)	1 (1%)	0	100	100
4	K	152/191 (80%)	142 (93%)	6 (4%)	4 (3%)	5	35
All	All	1230/1396 (88%)	1188 (97%)	34 (3%)	8 (1%)	22	62

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	K	6	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	K	7	ALA
4	K	174	ASP
4	K	176	ARG
3	J	69	LYS

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	E	175/179 (98%)	168 (96%)	7 (4%)	31	56
1	L	175/179 (98%)	171 (98%)	4 (2%)	50	70
2	F	181/193 (94%)	180 (99%)	1 (1%)	86	92
2	H	181/193 (94%)	175 (97%)	6 (3%)	38	62
3	C	47/54 (87%)	45 (96%)	2 (4%)	29	55
3	J	48/54 (89%)	44 (92%)	4 (8%)	11	37
4	D	125/165 (76%)	119 (95%)	6 (5%)	25	52
4	K	130/165 (79%)	126 (97%)	4 (3%)	40	63
All	All	1062/1182 (90%)	1028 (97%)	34 (3%)	39	62

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	63	LEU
3	J	66	LYS
4	K	83	ILE
4	D	60	ASN
3	C	66	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	L	189	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	D	95	ASN
1	E	38	GLN
2	F	216	ASN

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)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	NAG	J	501	3	14,14,15	0.40	0	17,19,21	0.65	0
5	NAG	K	201	4	14,14,15	0.67	1 (7%)	17,19,21	0.82	1 (5%)
5	NAG	C	101	3	14,14,15	0.53	0	17,19,21	0.93	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	J	501	3	-	2/6/23/26	0/1/1/1
5	NAG	K	201	4	-	0/6/23/26	0/1/1/1
5	NAG	C	101	3	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	K	201	NAG	C1-C2	2.21	1.55	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	101	NAG	C1-O5-C5	3.31	116.68	112.19
5	K	201	NAG	C1-O5-C5	2.65	115.79	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	J	501	NAG	O5-C5-C6-O6
5	J	501	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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.