wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:37 PM GMT

PDB ID	:	1PV2
Title	:	Native Form 2 E.coli Chaperone Hsp31
Authors	:	Quigley, P.M.; Korotkov, K.; Baneyx, F.; Hol, W.G.J.
Deposited on	:	2003-06-26
Resolution	:	2.71 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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

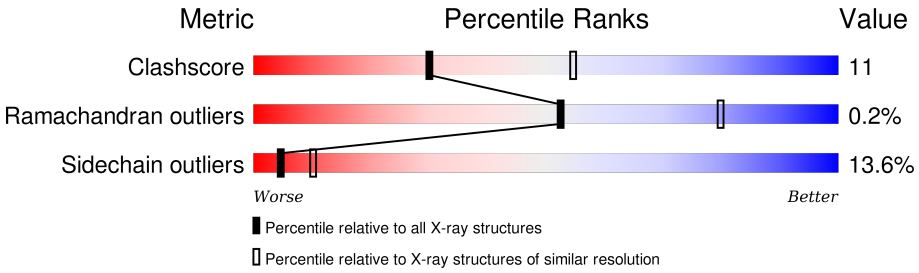
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.71 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	102246	2771 (2.74-2.70)
Ramachandran outliers	100387	2726 (2.74-2.70)
Sidechain outliers	100360	2727 (2.74-2.70)

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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	283	47% 31% 6% • 15%
1	B	283	65% 20% 5% 9%
1	C	283	64% 20% • 12%
1	D	283	63% 26% • 7%
1	E	283	64% 25% 5% 5%
1	F	283	62% 28% • 8%
1	G	283	66% 23% • 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	283	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 16233 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 Chaperone protein hchA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	241	Total	C	N	O	S	0	0	0
			1864	1203	309	344	8			
1	B	257	Total	C	N	O	S	0	0	0
			2008	1297	337	365	9			
1	C	249	Total	C	N	O	S	0	0	0
			1940	1252	323	356	9			
1	D	264	Total	C	N	O	S	0	0	0
			2048	1319	342	378	9			
1	E	270	Total	C	N	O	S	0	0	0
			2091	1341	352	389	9			
1	F	261	Total	C	N	O	S	0	0	0
			2031	1310	339	373	9			
1	G	260	Total	C	N	O	S	0	0	0
			2023	1305	338	371	9			
1	H	277	Total	C	N	O	S	0	0	0
			2156	1386	359	402	9			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	INITIATING MET	UNP P31658
B	1	MET	-	INITIATING MET	UNP P31658
C	1	MET	-	INITIATING MET	UNP P31658
D	1	MET	-	INITIATING MET	UNP P31658
E	1	MET	-	INITIATING MET	UNP P31658
F	1	MET	-	INITIATING MET	UNP P31658
G	1	MET	-	INITIATING MET	UNP P31658
H	1	MET	-	INITIATING MET	UNP P31658

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	7	Total O 7 7	0	0
2	B	7	Total O 7 7	0	0
2	C	13	Total O 13 13	0	0
2	D	7	Total O 7 7	0	0
2	E	6	Total O 6 6	0	0
2	F	8	Total O 8 8	0	0
2	G	11	Total O 11 11	0	0
2	H	13	Total O 13 13	0	0

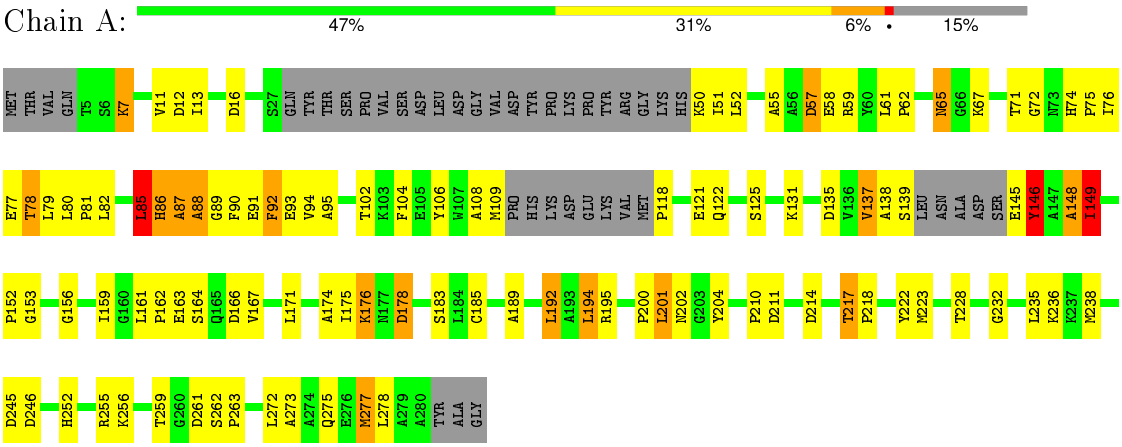
3 Residue-property plots (i)

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

Note EDS was not executed.

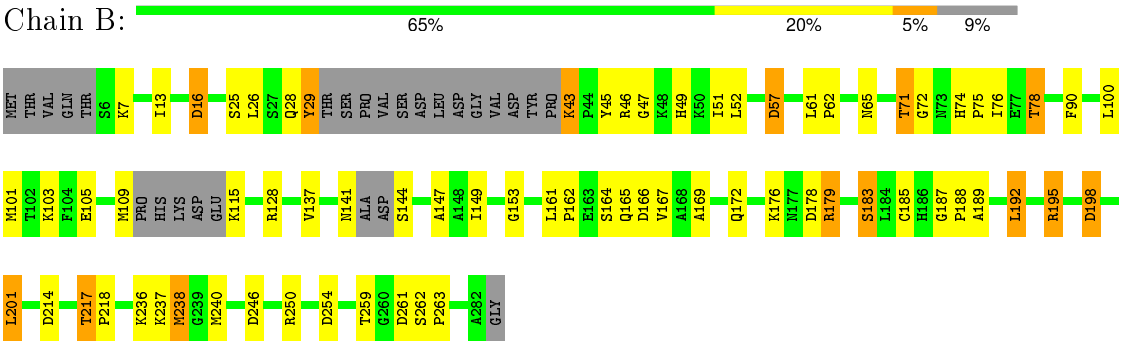
- Molecule 1: Chaperone protein hchA

Chain A:



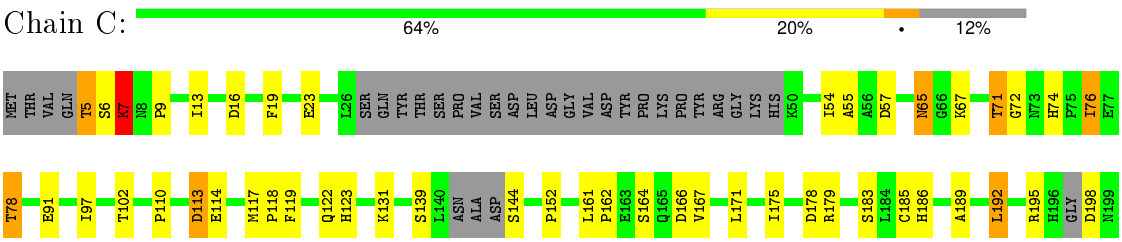
- Molecule 1: Chaperone protein hchA

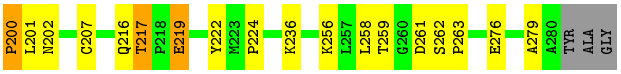
Chain B:



- Molecule 1: Chaperone protein hchA

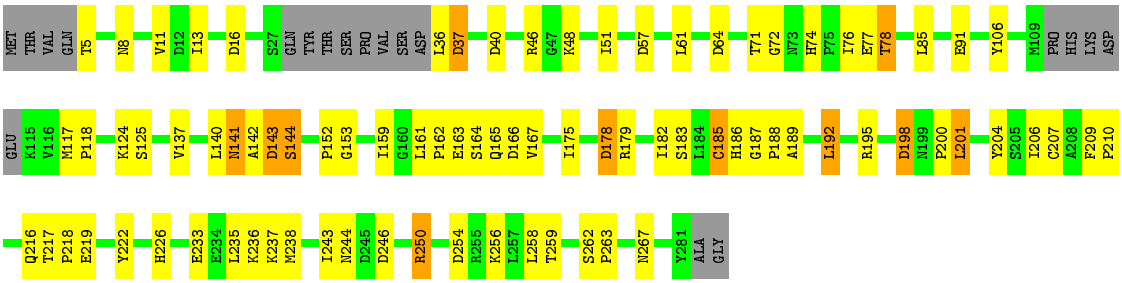
Chain C:





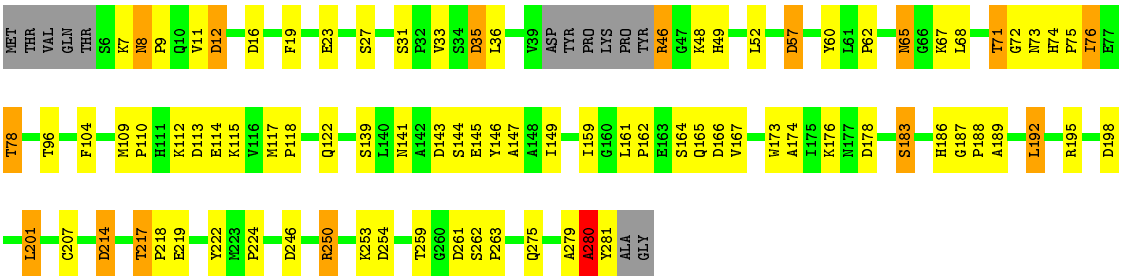
• Molecule 1: Chaperone protein hchA

Chain D:



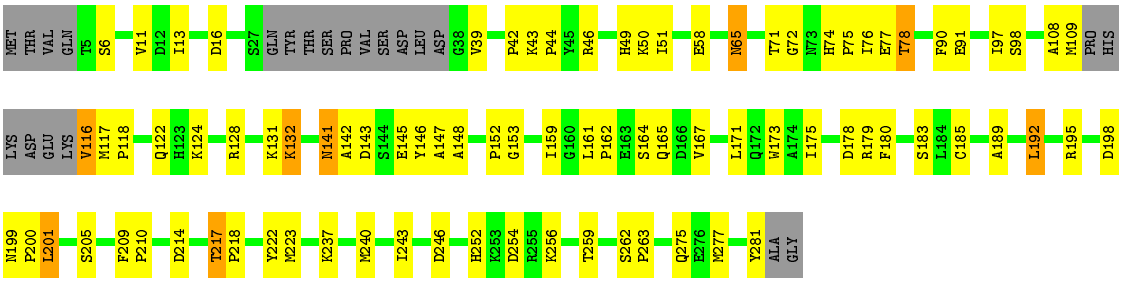
• Molecule 1: Chaperone protein hchA

Chain E:



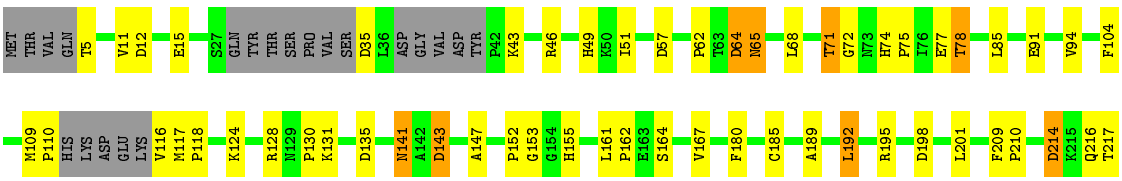
• Molecule 1: Chaperone protein hchA

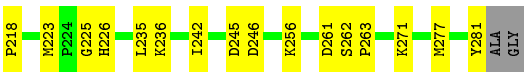
Chain F:



• Molecule 1: Chaperone protein hchA

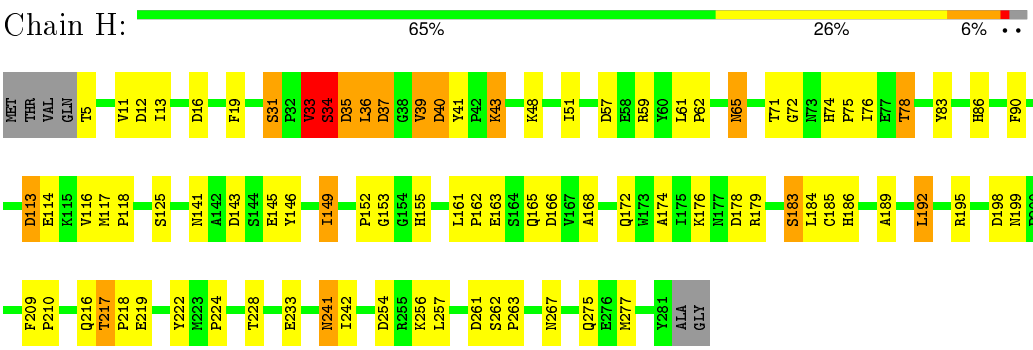
Chain G:





- Molecule 1: Chaperone protein hchA

Chain H:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	54.49Å 99.02Å 116.80Å 102.95° 101.52° 94.19°	Depositor
Resolution (Å)	48.80 – 2.71	Depositor
% Data completeness (in resolution range)	94.0 (48.80-2.71)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.221 , 0.286	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	16233	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

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	0.42	0/1912	0.98	24/2591 (0.9%)
1	B	0.26	0/2062	0.60	8/2792 (0.3%)
1	C	0.39	2/1991 (0.1%)	0.61	6/2698 (0.2%)
1	D	0.26	0/2104	0.62	10/2853 (0.4%)
1	E	0.32	0/2147	0.68	16/2912 (0.5%)
1	F	0.26	0/2087	0.61	6/2831 (0.2%)
1	G	0.26	0/2078	0.61	11/2817 (0.4%)
1	H	0.30	1/2218 (0.0%)	0.65	12/3014 (0.4%)
All	All	0.31	3/16599 (0.0%)	0.68	93/22508 (0.4%)

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	C	0	1
1	E	0	2
1	H	0	1
All	All	0	8

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	7	LYS	C-N	-10.12	1.10	1.34
1	C	200	PRO	C-N	8.60	1.53	1.34
1	H	34	SER	C-N	-6.46	1.19	1.34

The worst 5 of 93 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	146	TYR	CB-CG-CD2	18.61	132.17	121.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	148	ALA	C-N-CA	12.80	153.70	121.70
1	A	146	TYR	CB-CG-CD1	-12.66	113.40	121.00
1	A	87	ALA	N-CA-CB	8.19	121.56	110.10
1	A	146	TYR	CZ-CE2-CD2	-7.67	112.89	119.80

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	146	TYR	Sidechain
1	A	148	ALA	Mainchain,Peptide
1	A	85	LEU	Peptide
1	C	7	LYS	Mainchain
1	E	280	ALA	Mainchain

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	1864	0	1818	76	0
1	B	2008	0	1965	45	0
1	C	1940	0	1898	37	0
1	D	2048	0	1987	41	0
1	E	2091	0	2034	37	0
1	F	2031	0	1979	44	0
1	G	2023	0	1976	32	0
1	H	2156	0	2087	48	0
2	A	7	0	0	1	0
2	B	7	0	0	0	0
2	C	13	0	0	1	0
2	D	7	0	0	0	0
2	E	6	0	0	0	0
2	F	8	0	0	0	0
2	G	11	0	0	0	0
2	H	13	0	0	0	0
All	All	16233	0	15744	354	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 11.

The worst 5 of 354 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:A:74:HIS:O	1:A:78:THR:HG22	1.59	1.00
1:A:50:LYS:N	1:A:51:ILE:HA	1.72	0.99
1:A:85:LEU:O	1:A:90:PHE:HB2	1.74	0.86
1:B:49:HIS:HB3	1:B:147:ALA:HB2	1.59	0.85
1:A:75:PRO:HG2	1:A:109:MET:HB3	1.59	0.85

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	233/283 (82%)	204 (88%)	27 (12%)	2 (1%)	21 47
1	B	249/283 (88%)	232 (93%)	17 (7%)	0	100 100
1	C	241/283 (85%)	226 (94%)	14 (6%)	1 (0%)	39 68
1	D	258/283 (91%)	248 (96%)	10 (4%)	0	100 100
1	E	266/283 (94%)	251 (94%)	14 (5%)	1 (0%)	39 68
1	F	255/283 (90%)	242 (95%)	13 (5%)	0	100 100
1	G	252/283 (89%)	242 (96%)	10 (4%)	0	100 100
1	H	275/283 (97%)	256 (93%)	18 (6%)	1 (0%)	39 68
All	All	2029/2264 (90%)	1901 (94%)	123 (6%)	5 (0%)	52 80

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	36	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	86	HIS
1	A	87	ALA
1	E	280	ALA
1	C	279	ALA

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	192/230 (84%)	166 (86%)	26 (14%)	5 10
1	B	207/230 (90%)	178 (86%)	29 (14%)	4 9
1	C	202/230 (88%)	177 (88%)	25 (12%)	6 13
1	D	211/230 (92%)	185 (88%)	26 (12%)	6 13
1	E	218/230 (95%)	182 (84%)	36 (16%)	3 6
1	F	210/230 (91%)	179 (85%)	31 (15%)	4 9
1	G	210/230 (91%)	191 (91%)	19 (9%)	12 26
1	H	225/230 (98%)	190 (84%)	35 (16%)	3 8
All	All	1675/1840 (91%)	1448 (86%)	227 (14%)	5 10

5 of 227 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	236	LYS
1	E	141	ASN
1	H	145	GLU
1	D	250	ARG
1	E	46	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	267	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	196	HIS
1	H	196	HIS
1	E	65	ASN
1	E	141	ASN

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

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.