Full wwPDB NMR Structure Validation Report (i)

Mar 6, 2022 – 03:20 PM EST

PDB ID	:	2ODX
Title	:	Solution structure of Zn(II)Cox4
Authors	:	Coyne III, J.H.; Ciofi-Baffoni, S.
Deposited on	:	2006-12-27

This is a Full wwPDB NMR 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/NMRValidationReportHelp
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:

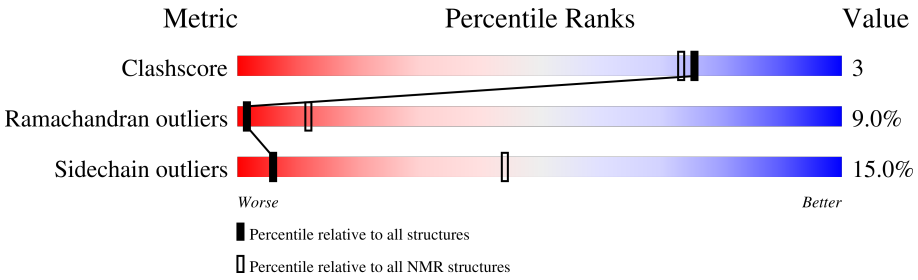
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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	80	51% 16% 32%

2 Ensemble composition and analysis (i)

This entry contains 21 models. Model 1 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:94-A:147 (54)	0.40	1

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 5 clusters and 4 single-model clusters were found.

Cluster number	Models
1	1, 3, 7, 11, 18
2	2, 4, 6, 14, 21
3	5, 9, 19
4	12, 13
5	15, 17
Single-model clusters	8; 10; 16; 20

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 828 atoms, of which 405 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cytochrome c oxidase polypeptide IV.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	54	827	269	405	69	79	5	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	76	MET	-	cloning artifact	UNP P04037
A	77	MET	-	cloning artifact	UNP P04037
A	78	ALA	-	cloning artifact	UNP P04037

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
2	A	1	1	1

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A: 51% 16% 32%

4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

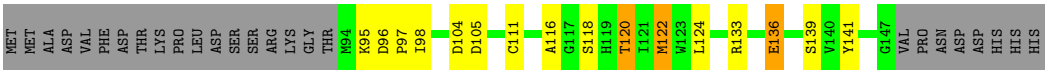
• Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A: 50% 16% • 32%

4.2.2 Score per residue for model 2

• Molecule 1: Cytochrome c oxidase polypeptide IV

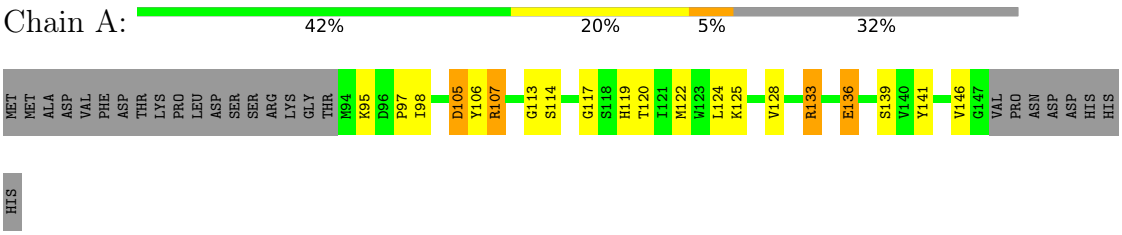
Chain A: 48% 16% • 32%



4.2.3 Score per residue for model 3

- Molecule 1: Cytochrome c oxidase polypeptide IV

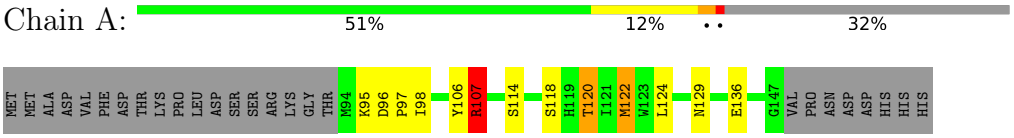
Chain A:



4.2.4 Score per residue for model 4

- Molecule 1: Cytochrome c oxidase polypeptide IV

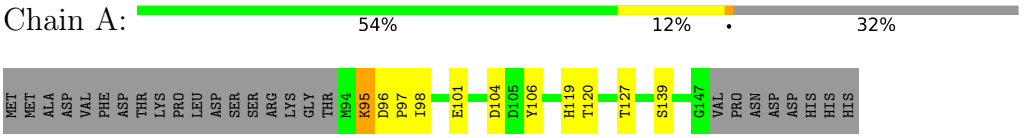
Chain A:



4.2.5 Score per residue for model 5

- Molecule 1: Cytochrome c oxidase polypeptide IV

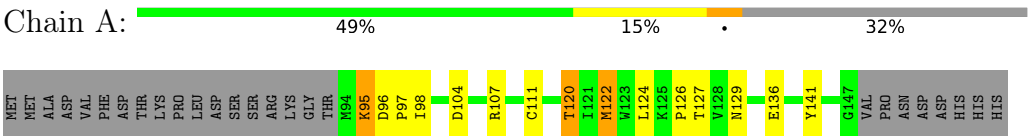
Chain A:



4.2.6 Score per residue for model 6

- Molecule 1: Cytochrome c oxidase polypeptide IV

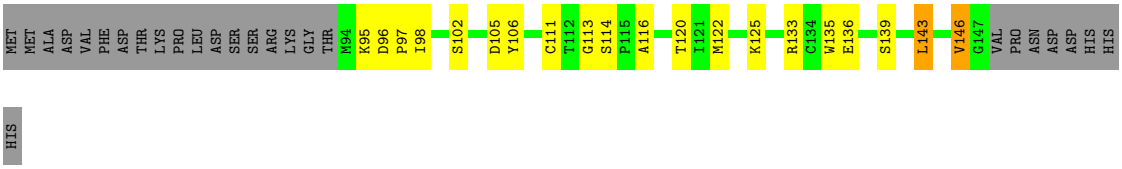
Chain A:



4.2.7 Score per residue for model 7

- Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A:



4.2.8 Score per residue for model 8

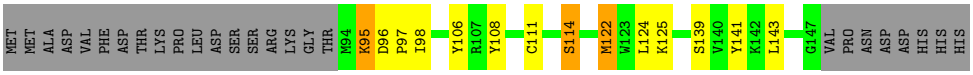
- Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A:

4.2.9 Score per residue for model 9

- Molecule 1: Cytochrome c oxidase polypeptide IV

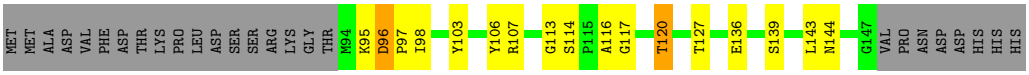
Chain A:



4.2.10 Score per residue for model 10

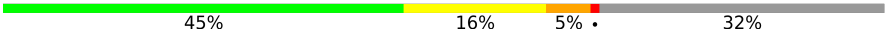
- Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A:



4.2.11 Score per residue for model 11

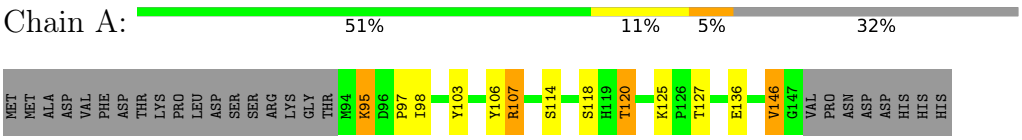
- Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A: 

4.2.12 Score per residue for model 12

- Molecule 1: Cytochrome c oxidase polypeptide IV

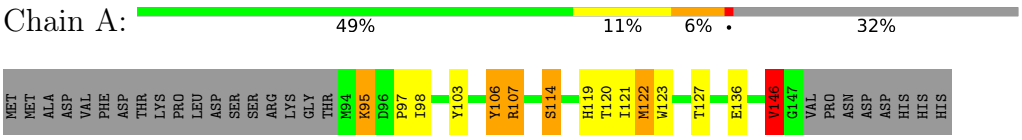
Chain A:



4.2.13 Score per residue for model 13

- Molecule 1: Cytochrome c oxidase polypeptide IV

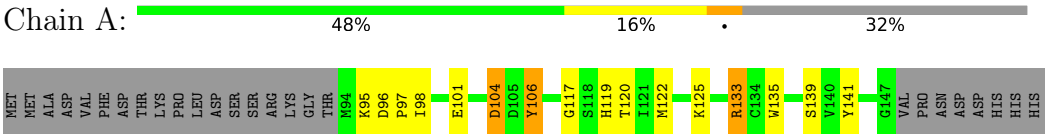
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: Cytochrome c oxidase polypeptide IV

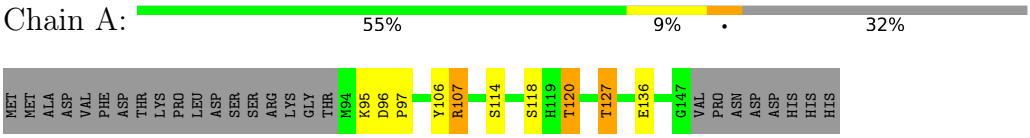
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Cytochrome c oxidase polypeptide IV

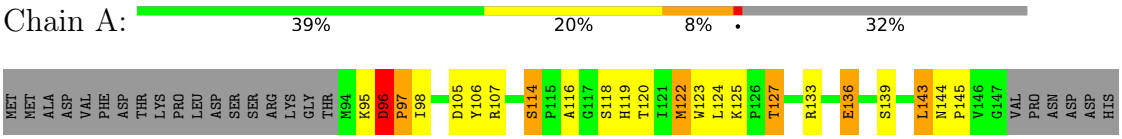
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A:



HIS
HIS

4.2.17 Score per residue for model 17

• Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A: 49% 12% 6% 32%

MET MET ALA ASP VAL PHE ASP THR LYS PRO LEU ASP SER SER ARG LYS GLY THR M94 K95 D96 P97 I98 D105 Y106 S114 T120 I121 M122 W123 L124 K125 P126 T127 V128 N129 E130 E136 G147 VAL PRO ASN ASP ASP HIS HIS HIS

4.2.18 Score per residue for model 18

• Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A: 54% 9% • • 32%

MET MET ALA ASP VAL PHE ASP THR LYS PRO LEU ASP SER SER ARG LYS GLY THR M94 K95 D96 P97 Y103 S114 H119 T120 I121 M122 W123 L124 E136 L143 V146 G147 VAL PRO ASN ASP ASP HIS HIS HIS

4.2.19 Score per residue for model 19

• Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A: 46% 16% 5% 32%

MET MET ALA ASP VAL PHE ASP THR LYS PRO LEU ASP SER SER ARG LYS GLY THR M94 K95 D96 P97 I98 D104 D105 Y106 R107 C111 S114 T120 I121 M122 W123 L124 K125 C134 W135 E136 S139 G147 VAL PRO ASN ASP ASP HIS HIS HIS

4.2.20 Score per residue for model 20

• Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A: 48% 20% 32%

MET MET ALA ASP VAL PHE ASP THR LYS PRO LEU ASP SER SER ARG LYS GLY THR M94 K95 D96 P97 I98 D105 Y106 R107 T120 I121 M122 K125 P126 T127 V128 N129 E130 R133 E136 L143 V146 G147 VAL PRO ASN ASP ASP HIS HIS HIS

4.2.21 Score per residue for model 21

• Molecule 1: Cytochrome c oxidase polypeptide IV

Chain A: 51% 12% • 32%

MET
MET
ALA
ASP
VAL
PHE
ASP
THR
LYS
PRO
LEU
ASP
SER
SER
ARG
LYS
GLY
THR
M94
K95
D96
P97
Y106
R107
C111
T112
G113
S114
T120
I121
M122
K125
E136
Y141
N144
G147
VAL
PRO
ASN
ASP
ASP
HIS
HIS
HIS

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics*.

Of the 100 calculated structures, 21 were deposited, based on the following criterion: *target function*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	structure solution	
Amber	refinement	8.0

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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

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 (average) 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±0.01	0±0/434 (0.0± 0.0%)	1.31±0.05	1±1/591 (0.2± 0.2%)
All	All	0.72	0/9114 (0.0%)	1.31	30/12411 (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	Planarity
1	A	0.0±0.0	1.9±1.4
All	All	0	39

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	107	ARG	NE-CZ-NH1	10.24	125.42	120.30	4	9
1	A	107	ARG	NE-CZ-NH2	-7.42	116.59	120.30	20	4
1	A	146	VAL	CG1-CB-CG2	-7.39	99.07	110.90	18	2
1	A	133	ARG	NE-CZ-NH1	6.56	123.58	120.30	8	5
1	A	141	TYR	CB-CG-CD1	-6.07	117.36	121.00	14	2
1	A	146	VAL	CA-CB-CG2	5.79	119.58	110.90	18	1
1	A	96	ASP	N-CA-CB	-5.50	100.70	110.60	5	3
1	A	122	MET	CG-SD-CE	5.47	108.95	100.20	20	1
1	A	107	ARG	CD-NE-CZ	5.25	130.95	123.60	4	1
1	A	107	ARG	NH1-CZ-NH2	-5.13	113.75	119.40	4	1
1	A	130	GLU	C-N-CA	5.04	134.29	121.70	17	1

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	106	TYR	Sidechain	11
1	A	141	TYR	Sidechain	5
1	A	117	GLY	Peptide	3
1	A	133	ARG	Sidechain	3
1	A	113	GLY	Peptide	3
1	A	146	VAL	Peptide	3
1	A	103	TYR	Peptide,Sidechain	3
1	A	129	ASN	Peptide	2
1	A	119	HIS	Sidechain,Peptide	2
1	A	107	ARG	Peptide,Sidechain	2
1	A	116	ALA	Peptide	1
1	A	108	TYR	Sidechain	1

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	422	405	404	3±2
All	All	8883	8505	8484	53

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 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:120:THR:HG22	1:A:136:GLU:OE1	0.66	1.91	11	5
1:A:120:THR:HG22	1:A:136:GLU:OE2	0.62	1.93	7	11
1:A:105:ASP:OD1	1:A:128:VAL:HG12	0.57	1.98	20	3
1:A:143:LEU:HD23	1:A:144:ASN:N	0.56	2.15	10	1
1:A:103:TYR:CD1	1:A:146:VAL:HG23	0.53	2.39	18	1
1:A:143:LEU:HD23	1:A:144:ASN:H	0.53	1.63	10	2
1:A:122:MET:CE	1:A:124:LEU:HD21	0.52	2.34	3	3
1:A:122:MET:SD	1:A:124:LEU:HD21	0.51	2.46	17	8

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:119:HIS:CE1	1:A:122:MET:SD	0.47	3.07	13	2
1:A:103:TYR:H	1:A:146:VAL:CG2	0.47	2.23	12	3
1:A:102:SER:HB2	1:A:143:LEU:HD21	0.45	1.88	7	1
1:A:111:CYS:H	1:A:122:MET:CE	0.45	2.25	19	1
1:A:105:ASP:CG	1:A:128:VAL:HG12	0.43	2.34	20	1
1:A:119:HIS:CD2	1:A:136:GLU:HB3	0.42	2.49	18	2
1:A:113:GLY:CA	1:A:119:HIS:CD2	0.42	3.03	3	1
1:A:122:MET:SD	1:A:135:TRP:NE1	0.42	2.92	7	2
1:A:111:CYS:H	1:A:122:MET:HE2	0.42	1.74	19	1
1:A:103:TYR:CE1	1:A:146:VAL:HG22	0.42	2.50	1	1
1:A:122:MET:CE	1:A:135:TRP:CD1	0.41	3.03	1	1
1:A:107:ARG:HH11	1:A:107:ARG:CG	0.41	2.28	4	1
1:A:96:ASP:N	1:A:97:PRO:HD3	0.41	2.31	16	1
1:A:121:ILE:C	1:A:122:MET:SD	0.41	2.99	13	1

6.3 Torsion angles (i)

6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	52/80 (65%)	39±2 (74±4%)	9±2 (17±4%)	5±1 (9±2%)	1 12
All	All	1092/1680 (65%)	810 (74%)	184 (17%)	98 (9%)	1 12

All 15 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	97	PRO	21
1	A	120	THR	15
1	A	114	SER	13
1	A	96	ASP	10
1	A	95	LYS	8
1	A	146	VAL	8
1	A	127	THR	8
1	A	105	ASP	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	116	ALA	4
1	A	118	SER	2
1	A	126	PRO	1
1	A	104	ASP	1
1	A	107	ARG	1
1	A	135	TRP	1
1	A	130	GLU	1

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	47/71 (66%)	40±3 (85±6%)	7±3 (15±6%)	6 44
All	All	987/1491 (66%)	839 (85%)	148 (15%)	6 44

All 25 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	95	LYS	20
1	A	98	ILE	17
1	A	125	LYS	13
1	A	106	TYR	12
1	A	139	SER	10
1	A	122	MET	9
1	A	114	SER	9
1	A	111	CYS	7
1	A	136	GLU	7
1	A	143	LEU	6
1	A	104	ASP	6
1	A	107	ARG	6
1	A	127	THR	4
1	A	118	SER	3
1	A	123	TRP	3
1	A	133	ARG	2
1	A	101	GLU	2
1	A	105	ASP	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	108	TYR	2
1	A	96	ASP	2
1	A	124	LEU	2
1	A	146	VAL	1
1	A	145	PRO	1
1	A	134	CYS	1
1	A	144	ASN	1

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided