wwPDB X-ray Structure Validation Summary Report (i)

Oct 10, 2021 – 06:49 PM EDT

PDB ID	:	3ER6
Title	:	Crystal structure of a putative transcriptional regulator protein from Vibrio parahaemolyticus
Authors	:	Bonanno, J.B.; Freeman, J.; Bain, K.T.; Chang, S.; Ozyurt, S.; Wasserman, S.; Sauder, J.M.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2008-10-01
Resolution	:	1.90 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 following versions of software and data (see references (i)) were used in the production of this report:

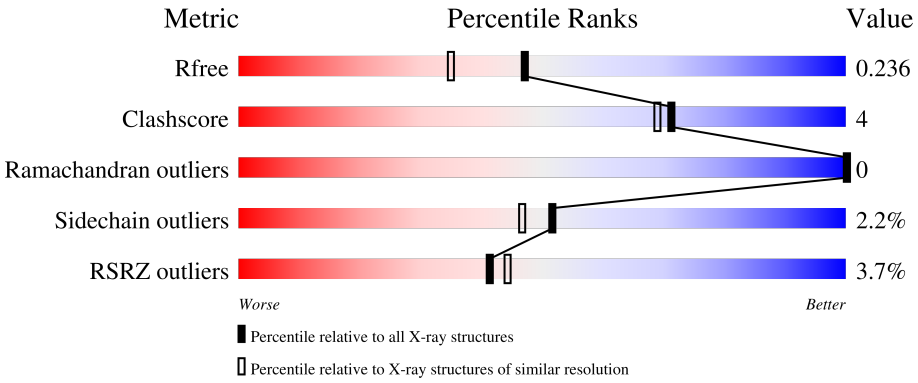
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.90 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.



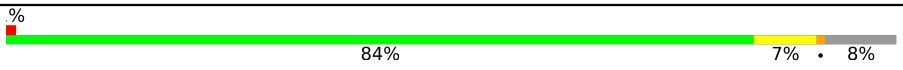
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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% 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.

Mol	Chain	Length	Quality of chain
1	A	209	86% 5% 8%
1	B	209	2% 82% 5% 12%
1	C	209	6% 78% 8% • 12%
1	D	209	6% 86% 6% 9%
1	E	209	6% 80% 11% • 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	209	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9362 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 Putative transcriptional regulator protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	192	Total	C	N	O	S	0	3	0
			1511	979	247	280	5			
1	B	184	Total	C	N	O	S	0	0	0
			1425	927	229	264	5			
1	C	183	Total	C	N	O	S	0	0	0
			1426	925	234	262	5			
1	D	191	Total	C	N	O	S	0	1	0
			1476	956	245	270	5			
1	E	191	Total	C	N	O	S	0	0	0
			1466	947	241	273	5			
1	F	192	Total	C	N	O	S	0	1	0
			1501	973	246	277	5			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q87I90
A	2	SER	-	expression tag	UNP Q87I90
A	3	LEU	-	expression tag	UNP Q87I90
A	92	LYS	ASN	engineered mutation	UNP Q87I90
A	202	GLU	-	expression tag	UNP Q87I90
A	203	GLY	-	expression tag	UNP Q87I90
A	204	HIS	-	expression tag	UNP Q87I90
A	205	HIS	-	expression tag	UNP Q87I90
A	206	HIS	-	expression tag	UNP Q87I90
A	207	HIS	-	expression tag	UNP Q87I90
A	208	HIS	-	expression tag	UNP Q87I90
A	209	HIS	-	expression tag	UNP Q87I90
B	1	MET	-	expression tag	UNP Q87I90
B	2	SER	-	expression tag	UNP Q87I90
B	3	LEU	-	expression tag	UNP Q87I90
B	92	LYS	ASN	engineered mutation	UNP Q87I90
B	202	GLU	-	expression tag	UNP Q87I90

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	203	GLY	-	expression tag	UNP Q87I90
B	204	HIS	-	expression tag	UNP Q87I90
B	205	HIS	-	expression tag	UNP Q87I90
B	206	HIS	-	expression tag	UNP Q87I90
B	207	HIS	-	expression tag	UNP Q87I90
B	208	HIS	-	expression tag	UNP Q87I90
B	209	HIS	-	expression tag	UNP Q87I90
C	1	MET	-	expression tag	UNP Q87I90
C	2	SER	-	expression tag	UNP Q87I90
C	3	LEU	-	expression tag	UNP Q87I90
C	92	LYS	ASN	engineered mutation	UNP Q87I90
C	202	GLU	-	expression tag	UNP Q87I90
C	203	GLY	-	expression tag	UNP Q87I90
C	204	HIS	-	expression tag	UNP Q87I90
C	205	HIS	-	expression tag	UNP Q87I90
C	206	HIS	-	expression tag	UNP Q87I90
C	207	HIS	-	expression tag	UNP Q87I90
C	208	HIS	-	expression tag	UNP Q87I90
C	209	HIS	-	expression tag	UNP Q87I90
D	1	MET	-	expression tag	UNP Q87I90
D	2	SER	-	expression tag	UNP Q87I90
D	3	LEU	-	expression tag	UNP Q87I90
D	92	LYS	ASN	engineered mutation	UNP Q87I90
D	202	GLU	-	expression tag	UNP Q87I90
D	203	GLY	-	expression tag	UNP Q87I90
D	204	HIS	-	expression tag	UNP Q87I90
D	205	HIS	-	expression tag	UNP Q87I90
D	206	HIS	-	expression tag	UNP Q87I90
D	207	HIS	-	expression tag	UNP Q87I90
D	208	HIS	-	expression tag	UNP Q87I90
D	209	HIS	-	expression tag	UNP Q87I90
E	1	MET	-	expression tag	UNP Q87I90
E	2	SER	-	expression tag	UNP Q87I90
E	3	LEU	-	expression tag	UNP Q87I90
E	92	LYS	ASN	engineered mutation	UNP Q87I90
E	202	GLU	-	expression tag	UNP Q87I90
E	203	GLY	-	expression tag	UNP Q87I90
E	204	HIS	-	expression tag	UNP Q87I90
E	205	HIS	-	expression tag	UNP Q87I90
E	206	HIS	-	expression tag	UNP Q87I90
E	207	HIS	-	expression tag	UNP Q87I90
E	208	HIS	-	expression tag	UNP Q87I90

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	209	HIS	-	expression tag	UNP Q87I90
F	1	MET	-	expression tag	UNP Q87I90
F	2	SER	-	expression tag	UNP Q87I90
F	3	LEU	-	expression tag	UNP Q87I90
F	92	LYS	ASN	engineered mutation	UNP Q87I90
F	202	GLU	-	expression tag	UNP Q87I90
F	203	GLY	-	expression tag	UNP Q87I90
F	204	HIS	-	expression tag	UNP Q87I90
F	205	HIS	-	expression tag	UNP Q87I90
F	206	HIS	-	expression tag	UNP Q87I90
F	207	HIS	-	expression tag	UNP Q87I90
F	208	HIS	-	expression tag	UNP Q87I90
F	209	HIS	-	expression tag	UNP Q87I90

- Molecule 2 is water.

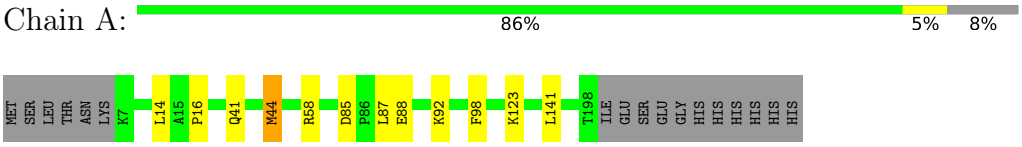
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	114	Total O 114 114	0	0
2	B	100	Total O 100 100	0	0
2	C	74	Total O 74 74	0	0
2	D	66	Total O 66 66	0	0
2	E	103	Total O 103 103	0	0
2	F	100	Total O 100 100	0	0

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 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.

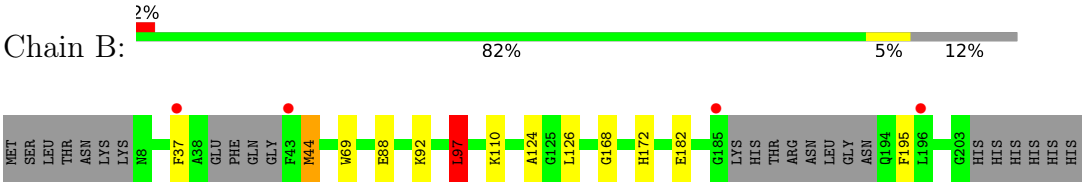
• Molecule 1: Putative transcriptional regulator protein

Chain A:



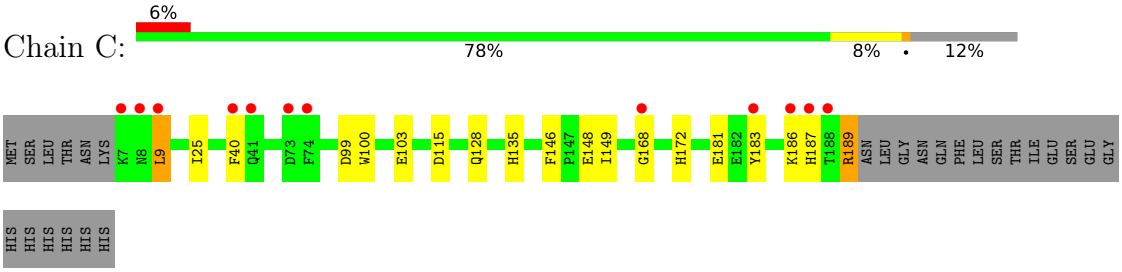
• Molecule 1: Putative transcriptional regulator protein

Chain B:



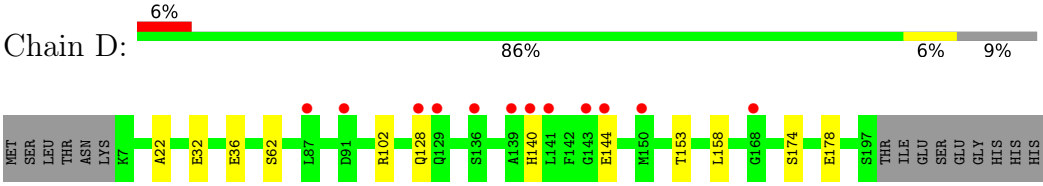
• Molecule 1: Putative transcriptional regulator protein

Chain C:



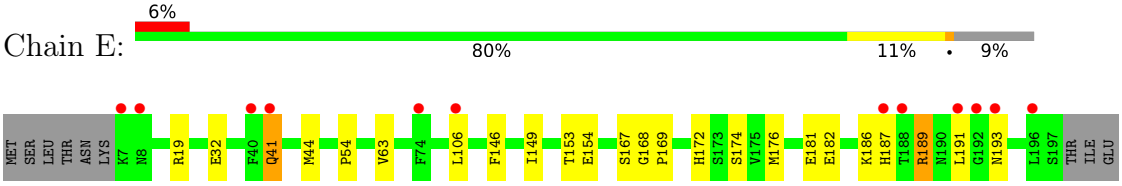
• Molecule 1: Putative transcriptional regulator protein

Chain D:



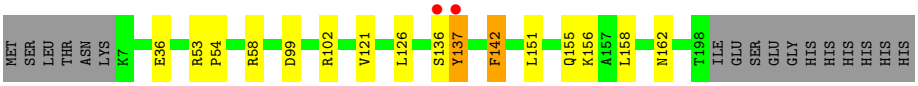
• Molecule 1: Putative transcriptional regulator protein

Chain E:



• Molecule 1: Putative transcriptional regulator protein

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	105.27Å 105.27Å 101.75Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 1.90 46.75 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-1.90) 100.0 (46.75-1.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	2.23 (at 1.90Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.192 , 0.231 0.197 , 0.236	Depositor DCC
R_{free} test set	4961 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	27.4	Xtriage
Anisotropy	0.269	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 55.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.005 for -h,-k,l 0.026 for h,-h-k,-l 0.037 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9362	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% 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	0.96	0/1555	0.83	3/2109 (0.1%)
1	B	0.99	0/1456	0.85	3/1974 (0.2%)
1	C	0.77	0/1460	0.74	0/1982
1	D	0.76	0/1513	0.73	1/2054 (0.0%)
1	E	0.96	0/1499	0.86	3/2037 (0.1%)
1	F	0.97	0/1539	0.79	0/2088
All	All	0.91	0/9022	0.80	10/12244 (0.1%)

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	44	MET	CG-SD-CE	-7.57	88.08	100.20
1	B	44	MET	CA-CB-CG	7.18	125.51	113.30
1	E	176	MET	CG-SD-CE	5.94	109.71	100.20
1	A	85	ASP	CB-CG-OD1	5.82	123.54	118.30
1	E	19	ARG	NE-CZ-NH1	-5.78	117.41	120.30

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	1511	0	1509	7	0
1	B	1425	0	1420	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1426	0	1416	13	0
1	D	1476	0	1465	6	0
1	E	1466	0	1444	13	0
1	F	1501	0	1493	15	0
2	A	114	0	0	0	0
2	B	100	0	0	0	0
2	C	74	0	0	0	0
2	D	66	0	0	0	0
2	E	103	0	0	0	0
2	F	100	0	0	1	0
All	All	9362	0	8747	62	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 4.

The worst 5 of 62 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:E:41:GLN:H	1:E:41:GLN:NE2	1.67	0.92
1:F:156:LYS:HD2	1:F:156:LYS:N	1.87	0.89
1:F:137:TYR:O	1:F:137:TYR:CD2	2.32	0.82
1:F:137:TYR:O	1:F:137:TYR:HD2	1.64	0.80
1:D:140:HIS:O	1:D:144:GLU:HG3	1.89	0.72

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	193/209 (92%)	191 (99%)	2 (1%)	0	100	100
1	B	178/209 (85%)	175 (98%)	3 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	181/209 (87%)	176 (97%)	5 (3%)	0	100	100
1	D	190/209 (91%)	188 (99%)	2 (1%)	0	100	100
1	E	189/209 (90%)	184 (97%)	5 (3%)	0	100	100
1	F	191/209 (91%)	188 (98%)	3 (2%)	0	100	100
All	All	1122/1254 (90%)	1102 (98%)	20 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	165/179 (92%)	162 (98%)	3 (2%)	59	55
1	B	154/179 (86%)	152 (99%)	2 (1%)	69	68
1	C	153/179 (86%)	150 (98%)	3 (2%)	55	51
1	D	158/179 (88%)	154 (98%)	4 (2%)	47	41
1	E	156/179 (87%)	152 (97%)	4 (3%)	46	39
1	F	162/179 (90%)	157 (97%)	5 (3%)	40	32
All	All	948/1074 (88%)	927 (98%)	21 (2%)	52	47

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	189	ARG
1	F	137	TYR
1	F	162	ASN
1	F	142	PHE
1	F	136	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	135	HIS
1	D	128	GLN
1	E	41	GLN
1	D	155	GLN
1	B	162	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 monosaccharides 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

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	192/209 (91%)	-0.07	0 100 100	25, 33, 46, 53	0
1	B	184/209 (88%)	0.15	4 (2%) 62 64	25, 34, 52, 61	0
1	C	183/209 (87%)	0.27	12 (6%) 18 20	32, 41, 63, 78	0
1	D	191/209 (91%)	0.19	12 (6%) 20 22	31, 42, 63, 70	0
1	E	191/209 (91%)	0.20	12 (6%) 20 22	26, 34, 68, 82	0
1	F	192/209 (91%)	0.08	2 (1%) 82 84	26, 34, 54, 60	0
All	All	1133/1254 (90%)	0.13	42 (3%) 41 44	25, 36, 60, 82	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	168	GLY	5.4
1	F	137	TYR	5.1
1	D	139	ALA	4.1
1	D	150	MET	4.0
1	D	140	HIS	4.0

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.