



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 1, 2026 – 09:57 PM UTC

PDB ID : 2E1C / pdb_00002e1c
Title : Structure of Putative HTH-type transcriptional regulator PH1519/DNA Complex
Authors : Koike, H.; Suzuki, M.
Deposited on : 2006-10-24
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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

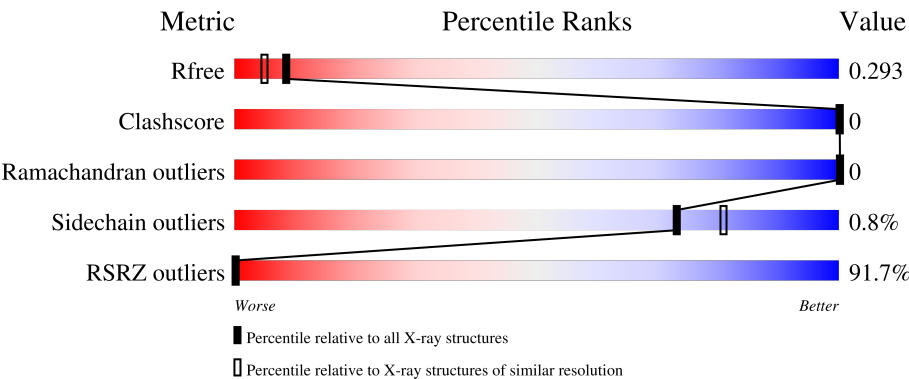
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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}	180053	6658 (2.10-2.10)
Clashscore	190562	7164 (2.10-2.10)
Ramachandran outliers	187476	7099 (2.10-2.10)
Sidechain outliers	187428	7100 (2.10-2.10)
RSRZ outliers	180081	6662 (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 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 $\leq 5\%$. 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	1-B	17	65% 94%6%
1	2-B	17	94%6%
2	1-D	17	59% 94%6%
2	2-D	17	94%6%
3	1-A	171	85% 84%.14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	2-A	171	 A horizontal bar chart showing the quality of the chain. The bar is divided into two segments: a green segment representing 84% and a yellow segment representing 14%. The total length of the bar is 100%.

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3958 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 DNA chain called DNA (5'-D(*DAP*DGP*DTP*DGP*DAP*DAP*DAP*DAP*DTP*DTP*DTP*DTP*DTP*DTP*DCP*DAP*DCP*DA)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1-B	17	Total	C	N	O	P	0	0	0
			346	168	63	99	16			
1	2-B	17	Total	C	N	O	P	0	0	0
			346	168	63	99	16			

- Molecule 2 is a DNA chain called DNA (5'-D(*DTP*DGP*DTP*DGP*DAP*DAP*DAP*DAP*DAP*DTP*DTP*DTP*DTP*DTP*DCP*DAP*DCP*DT)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	1-D	17	Total	C	N	O	P	0	0	0
			345	168	60	101	16			
2	2-D	17	Total	C	N	O	P	0	0	0
			345	168	60	101	16			

- Molecule 3 is a protein called Putative HTH-type transcriptional regulator PH1519.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	1-A	147	Total	C	N	O	S	0	0	0
			1155	741	190	221	3			
3	2-A	147	Total	C	N	O	S	0	0	0
			1155	741	190	221	3			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP O59188
A	2	GLY	-	expression tag	UNP O59188
A	3	SER	-	expression tag	UNP O59188
A	4	SER	-	expression tag	UNP O59188
A	5	HIS	-	expression tag	UNP O59188
A	6	HIS	-	expression tag	UNP O59188

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	HIS	-	expression tag	UNP O59188
A	8	HIS	-	expression tag	UNP O59188
A	9	HIS	-	expression tag	UNP O59188
A	10	HIS	-	expression tag	UNP O59188
A	11	SER	-	expression tag	UNP O59188
A	12	SER	-	expression tag	UNP O59188
A	13	GLY	-	expression tag	UNP O59188
A	14	LEU	-	expression tag	UNP O59188
A	15	VAL	-	expression tag	UNP O59188
A	16	PRO	-	expression tag	UNP O59188
A	17	ARG	-	expression tag	UNP O59188
A	18	GLY	-	expression tag	UNP O59188
A	19	SER	-	expression tag	UNP O59188
A	20	HIS	-	expression tag	UNP O59188

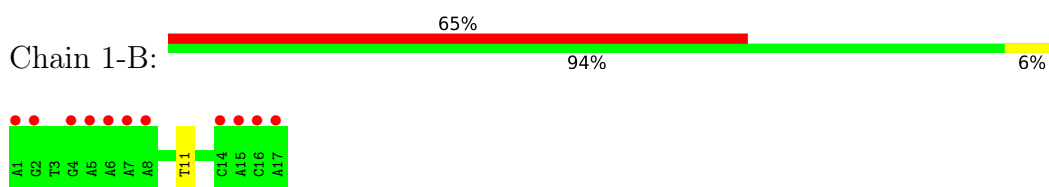
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	1-B	36	Total O 36 36	0	0
4	2-B	4	Total O 4 4	0	0
4	1-D	5	Total O 5 5	0	0
4	2-D	39	Total O 39 39	0	0
4	1-A	92	Total O 92 92	0	0
4	2-A	90	Total O 90 90	0	0

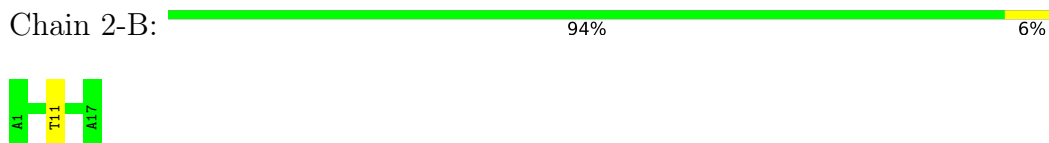
3 Residue-property plots [i](#)

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.

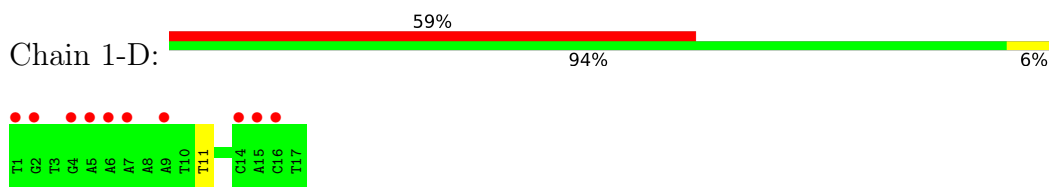
- Molecule 1: DNA (5'-D(*DAP*DGP*DTP*DGP*DAP*DAP*DAP*DAP*DTP*DTP*DTP*DTP*DTP*DCP*DAP*DCP*DA)-3')



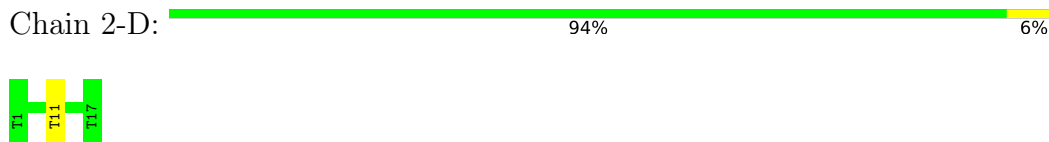
- Molecule 1: DNA (5'-D(*DAP*DGP*DTP*DGP*DAP*DAP*DAP*DAP*DTP*DTP*DTP*DTP*DTP*DCP*DAP*DCP*DA)-3')



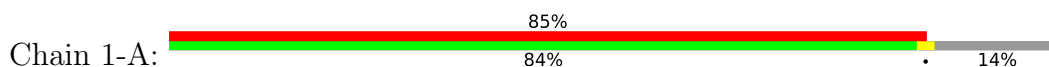
- Molecule 2: DNA (5'-D(*DTP*DGP*DTP*DGP*DAP*DAP*DAP*DAP*DAP*DTP*DTP*DTP*DTP*DCP*DAP*DCP*DT)-3')

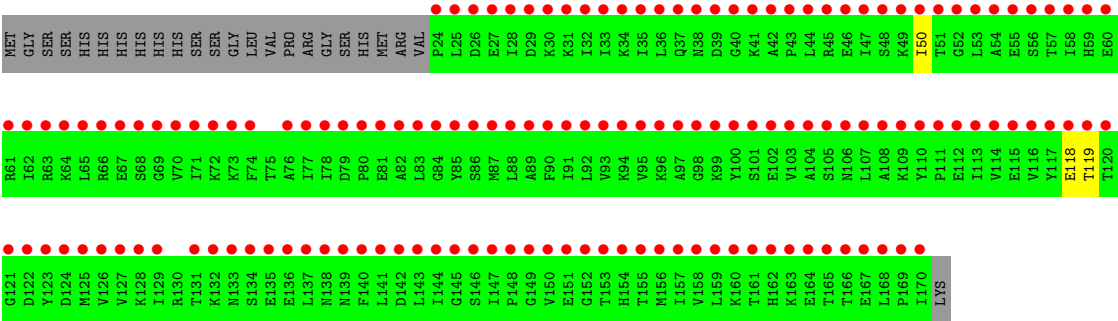


- Molecule 2: DNA (5'-D(*DTP*DGP*DTP*DGP*DAP*DAP*DAP*DAP*DAP*DTP*DTP*DTP*DTP*DCP*DAP*DCP*DT)-3')



- Molecule 3: Putative HTH-type transcriptional regulator PH1519



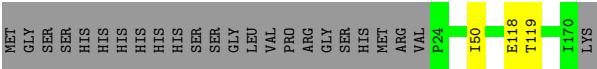


● Molecule 3: Putative HTH-type transcriptional regulator PH1519

Chain 2-A:

84%

14%



4 Data and refinement statistics

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	68.41Å 92.35Å 109.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	55.00 – 2.10 55.00 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.2 (55.00-2.10) 99.2 (55.00-2.10)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.64 (at 2.00Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.206 , 0.230 0.288 , 0.293	Depositor DCC
R_{free} test set	1215 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å ²)	41.1	Xtriage
Anisotropy	0.612	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 107.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.78	EDS
Total number of atoms	3958	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.62% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	1-B	0.25	0/388	0.71	0/597
1	2-B	0.25	0/388	0.71	0/597
2	1-D	0.26	0/386	0.71	0/594
2	2-D	0.26	0/386	0.71	0/594
3	1-A	0.36	0/1171	0.87	2/1578 (0.1%)
3	2-A	0.36	0/1171	0.87	2/1578 (0.1%)
All	All	0.32	0/3890	0.81	4/5538 (0.1%)

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	1-B	0	1
1	2-B	0	1
2	1-D	0	1
2	2-D	0	1
All	All	0	4

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	1-A	119	THR	N-CA-C	5.70	117.64	109.14
3	2-A	119	THR	N-CA-C	5.70	117.64	109.14
3	1-A	118	GLU	N-CA-C	-5.14	100.87	109.24
3	2-A	118	GLU	N-CA-C	-5.14	100.87	109.24

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	1-B	11	DT	Sidechain
2	1-D	11	DT	Sidechain
1	2-B	11	DT	Sidechain
2	2-D	11	DT	Sidechain

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	1-B	346	0	195	0	0
1	2-B	346	0	195	0	0
2	1-D	345	0	196	0	0
2	2-D	345	0	196	0	0
3	1-A	1155	0	1217	0	0
3	2-A	1155	0	1217	0	0
4	1-A	92	0	0	0	0
4	1-B	36	0	0	0	0
4	1-D	5	0	0	0	0
4	2-A	90	0	0	0	0
4	2-B	4	0	0	0	0
4	2-D	39	0	0	0	0
All	All	3958	0	3216	0	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 0.

There are no clashes within the asymmetric unit.

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	
3	1-A	145/171 (85%)	142 (98%)	3 (2%)	0	100	100
3	2-A	145/171 (85%)	142 (98%)	3 (2%)	0	100	100
All	All	290/342 (85%)	284 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	
3	1-A	130/151 (86%)	129 (99%)	1 (1%)	73	81
3	2-A	130/151 (86%)	129 (99%)	1 (1%)	73	81
All	All	260/302 (86%)	258 (99%)	2 (1%)	73	81

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	1-A	50	ILE
3	2-A	50	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	1-A	37	GLN
3	2-A	37	GLN

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 [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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

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, 95th 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	1-B	17/17 (100%)	2.32	11 (64%) 0 0	17, 24, 30, 33	17 (100%)
1	2-B	0/17	-	-	-	-
2	1-D	17/17 (100%)	2.29	10 (58%) 0 0	16, 24, 30, 32	17 (100%)
2	2-D	0/17	-	-	-	-
3	1-A	147/171 (85%)	4.88	145 (98%) 0 0	16, 22, 31, 39	147 (100%)
3	2-A	0/171	-	-	-	-
All	All	181/410 (44%)	4.40	166 (91%) 0 0	16, 23, 31, 39	181 (100%)

All (166) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	1-A	105	SER	14.9
3	1-A	24	PRO	14.2
3	1-A	135	GLU	11.6
3	1-A	81	GLU	10.8
3	1-A	25	LEU	10.6
3	1-A	170	ILE	9.4
3	1-A	98	GLY	9.2
3	1-A	82	ALA	9.1
3	1-A	67	GLU	8.6
3	1-A	66	ARG	8.4
3	1-A	27	GLU	8.3
3	1-A	97	ALA	8.3
3	1-A	101	SER	7.9
3	1-A	160	LYS	7.7
3	1-A	138	ASN	7.6
3	1-A	108	ALA	7.4
3	1-A	83	LEU	7.3
3	1-A	28	ILE	7.0
3	1-A	53	LEU	6.9
3	1-A	95	VAL	6.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	1-A	84	GLY	6.7
3	1-A	99	LYS	6.6
3	1-A	102	GLU	6.6
3	1-A	68	SER	6.6
3	1-A	48	SER	6.5
3	1-A	96	LYS	6.5
3	1-A	103	VAL	6.5
3	1-A	151	GLU	6.5
3	1-A	100	TYR	6.2
3	1-A	104	ALA	6.1
3	1-A	119	THR	6.1
3	1-A	72	LYS	6.1
3	1-A	168	LEU	6.1
3	1-A	64	LYS	6.0
3	1-A	85	TYR	6.0
3	1-A	42	ALA	6.0
3	1-A	134	SER	5.9
3	1-A	132	LYS	5.9
3	1-A	106	ASN	5.8
3	1-A	79	ASP	5.8
3	1-A	123	TYR	5.7
3	1-A	133	ASN	5.6
3	1-A	125	MET	5.6
3	1-A	94	LYS	5.6
3	1-A	142	ASP	5.5
3	1-A	86	SER	5.5
3	1-A	169	PRO	5.5
3	1-A	159	LEU	5.5
3	1-A	40	GLY	5.4
3	1-A	141	LEU	5.3
3	1-A	122	ASP	5.3
3	1-A	39	ASP	5.2
3	1-A	51	THR	5.2
3	1-A	60	GLU	5.2
3	1-A	30	LYS	5.2
3	1-A	34	LYS	5.2
3	1-A	87	MET	5.1
3	1-A	152	GLY	5.1
3	1-A	153	THR	5.1
3	1-A	158	VAL	5.0
3	1-A	63	ARG	5.0
3	1-A	150	VAL	5.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	1-A	167	GLU	4.9
3	1-A	41	LYS	4.8
3	1-A	52	GLY	4.8
3	1-A	109	LYS	4.8
3	1-A	162	HIS	4.7
3	1-A	55	GLU	4.7
3	1-A	59	HIS	4.6
3	1-A	129	ILE	4.6
3	1-A	50	ILE	4.5
3	1-A	161	THR	4.5
3	1-A	139	ASN	4.4
3	1-A	147	ILE	4.4
3	1-A	165	THR	4.3
3	1-A	140	PHE	4.3
3	1-A	107	LEU	4.3
3	1-A	78	ILE	4.3
3	1-A	113	ILE	4.3
3	1-A	26	ASP	4.2
3	1-A	121	GLY	4.2
3	1-A	92	LEU	4.1
3	1-A	131	THR	4.0
3	1-A	136	GLU	4.0
3	1-A	127	VAL	4.0
3	1-A	120	THR	3.9
3	1-A	166	THR	3.9
3	1-A	35	ILE	3.9
3	1-A	74	PHE	3.9
3	1-A	69	GLY	3.9
3	1-A	43	PRO	3.9
3	1-A	29	ASP	3.8
3	1-A	126	VAL	3.7
3	1-A	91	ILE	3.7
3	1-A	65	LEU	3.7
3	1-A	80	PRO	3.7
3	1-A	148	PRO	3.7
3	1-A	145	GLY	3.7
3	1-A	71	ILE	3.7
3	1-A	154	HIS	3.6
3	1-A	76	ALA	3.6
3	1-A	149	GLY	3.6
1	1-B	15	DA	3.5
3	1-A	49	LYS	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	1-A	117	TYR	3.5
3	1-A	56	SER	3.5
2	1-D	15	DA	3.5
3	1-A	157	ILE	3.4
3	1-A	110	TYR	3.4
3	1-A	33	ILE	3.4
3	1-A	137	LEU	3.4
3	1-A	70	VAL	3.4
2	1-D	14	DC	3.3
1	1-B	1	DA	3.3
1	1-B	14	DC	3.3
3	1-A	155	THR	3.3
3	1-A	128	LYS	3.3
3	1-A	114	VAL	3.3
3	1-A	54	ALA	3.3
3	1-A	57	THR	3.2
3	1-A	93	VAL	3.2
3	1-A	62	ILE	3.1
3	1-A	111	PRO	3.1
3	1-A	36	LEU	3.1
3	1-A	32	ILE	3.1
3	1-A	31	LYS	3.1
3	1-A	156	MET	3.1
3	1-A	38	ASN	3.1
1	1-B	4	DG	3.1
2	1-D	4	DG	3.0
3	1-A	88	LEU	3.0
3	1-A	73	LYS	3.0
3	1-A	164	GLU	2.9
3	1-A	47	ILE	2.9
3	1-A	61	ARG	2.9
1	1-B	5	DA	2.8
1	1-B	6	DA	2.8
2	1-D	5	DA	2.8
2	1-D	1	DT	2.8
2	1-D	6	DA	2.8
3	1-A	112	GLU	2.7
3	1-A	44	LEU	2.7
3	1-A	89	ALA	2.6
3	1-A	58	ILE	2.6
3	1-A	144	ILE	2.6
3	1-A	118	GLU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	1-B	7	DA	2.5
2	1-D	7	DA	2.5
3	1-A	115	GLU	2.5
3	1-A	45	ARG	2.5
3	1-A	77	ILE	2.5
2	1-D	9	DA	2.5
3	1-A	90	PHE	2.5
3	1-A	146	SER	2.4
3	1-A	37	GLN	2.4
3	1-A	124	ASP	2.4
3	1-A	116	VAL	2.4
3	1-A	143	LEU	2.3
1	1-B	2	DG	2.3
2	1-D	2	DG	2.3
1	1-B	17	DA	2.2
3	1-A	46	GLU	2.2
1	1-B	16	DC	2.2
2	1-D	16	DC	2.2
3	1-A	163	LYS	2.2
1	1-B	8	DA	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.