Full wwPDB X-ray Structure Validation Report (i)

Nov 21, 2023 – 07:56 AM JST

PDB ID	:	7FFH
Title	:	Diarylpentanoid-producing polyketide synthase (N199L mutant)
Authors	:	Morita, H.; Wong, C.P.; Liu, Q.; Takeshi, K.; Lee, Y.; Nakashima, Y.
Deposited on	:	2021-07-23
Resolution	:	2.20 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 (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 (1)) were used in the production of this report:

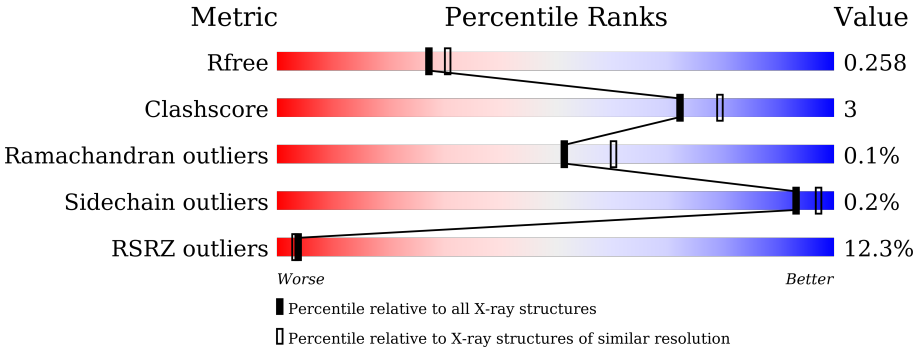
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	431	11% 81% 9% 11%
1	B	431	11% 81% 9% 10%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5966 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 Type III polyketide synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	385	2935	1872	499	551	13	0	0	0
1	B	386	2943	1878	501	551	13	0	1	0

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-33	MET	-	initiating methionine	UNP A0A385MEG6
A	-32	GLY	-	expression tag	UNP A0A385MEG6
A	-31	SER	-	expression tag	UNP A0A385MEG6
A	-30	SER	-	expression tag	UNP A0A385MEG6
A	-29	HIS	-	expression tag	UNP A0A385MEG6
A	-28	HIS	-	expression tag	UNP A0A385MEG6
A	-27	HIS	-	expression tag	UNP A0A385MEG6
A	-26	HIS	-	expression tag	UNP A0A385MEG6
A	-25	HIS	-	expression tag	UNP A0A385MEG6
A	-24	HIS	-	expression tag	UNP A0A385MEG6
A	-23	SER	-	expression tag	UNP A0A385MEG6
A	-22	SER	-	expression tag	UNP A0A385MEG6
A	-21	GLY	-	expression tag	UNP A0A385MEG6
A	-20	LEU	-	expression tag	UNP A0A385MEG6
A	-19	VAL	-	expression tag	UNP A0A385MEG6
A	-18	PRO	-	expression tag	UNP A0A385MEG6
A	-17	ARG	-	expression tag	UNP A0A385MEG6
A	-16	GLY	-	expression tag	UNP A0A385MEG6
A	-15	SER	-	expression tag	UNP A0A385MEG6
A	-14	HIS	-	expression tag	UNP A0A385MEG6
A	-13	MET	-	expression tag	UNP A0A385MEG6
A	-12	ALA	-	expression tag	UNP A0A385MEG6
A	-11	SER	-	expression tag	UNP A0A385MEG6
A	-10	MET	-	expression tag	UNP A0A385MEG6
A	-9	THR	-	expression tag	UNP A0A385MEG6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	GLY	-	expression tag	UNP A0A385MEG6
A	-7	GLY	-	expression tag	UNP A0A385MEG6
A	-6	GLN	-	expression tag	UNP A0A385MEG6
A	-5	GLN	-	expression tag	UNP A0A385MEG6
A	-4	MET	-	expression tag	UNP A0A385MEG6
A	-3	GLY	-	expression tag	UNP A0A385MEG6
A	-2	ARG	-	expression tag	UNP A0A385MEG6
A	-1	GLY	-	expression tag	UNP A0A385MEG6
A	0	SER	-	expression tag	UNP A0A385MEG6
A	199	LEU	ASN	engineered mutation	UNP A0A385MEG6
B	-33	MET	-	initiating methionine	UNP A0A385MEG6
B	-32	GLY	-	expression tag	UNP A0A385MEG6
B	-31	SER	-	expression tag	UNP A0A385MEG6
B	-30	SER	-	expression tag	UNP A0A385MEG6
B	-29	HIS	-	expression tag	UNP A0A385MEG6
B	-28	HIS	-	expression tag	UNP A0A385MEG6
B	-27	HIS	-	expression tag	UNP A0A385MEG6
B	-26	HIS	-	expression tag	UNP A0A385MEG6
B	-25	HIS	-	expression tag	UNP A0A385MEG6
B	-24	HIS	-	expression tag	UNP A0A385MEG6
B	-23	SER	-	expression tag	UNP A0A385MEG6
B	-22	SER	-	expression tag	UNP A0A385MEG6
B	-21	GLY	-	expression tag	UNP A0A385MEG6
B	-20	LEU	-	expression tag	UNP A0A385MEG6
B	-19	VAL	-	expression tag	UNP A0A385MEG6
B	-18	PRO	-	expression tag	UNP A0A385MEG6
B	-17	ARG	-	expression tag	UNP A0A385MEG6
B	-16	GLY	-	expression tag	UNP A0A385MEG6
B	-15	SER	-	expression tag	UNP A0A385MEG6
B	-14	HIS	-	expression tag	UNP A0A385MEG6
B	-13	MET	-	expression tag	UNP A0A385MEG6
B	-12	ALA	-	expression tag	UNP A0A385MEG6
B	-11	SER	-	expression tag	UNP A0A385MEG6
B	-10	MET	-	expression tag	UNP A0A385MEG6
B	-9	THR	-	expression tag	UNP A0A385MEG6
B	-8	GLY	-	expression tag	UNP A0A385MEG6
B	-7	GLY	-	expression tag	UNP A0A385MEG6
B	-6	GLN	-	expression tag	UNP A0A385MEG6
B	-5	GLN	-	expression tag	UNP A0A385MEG6
B	-4	MET	-	expression tag	UNP A0A385MEG6
B	-3	GLY	-	expression tag	UNP A0A385MEG6
B	-2	ARG	-	expression tag	UNP A0A385MEG6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP A0A385MEG6
B	0	SER	-	expression tag	UNP A0A385MEG6
B	199	LEU	ASN	engineered mutation	UNP A0A385MEG6

- Molecule 2 is water.

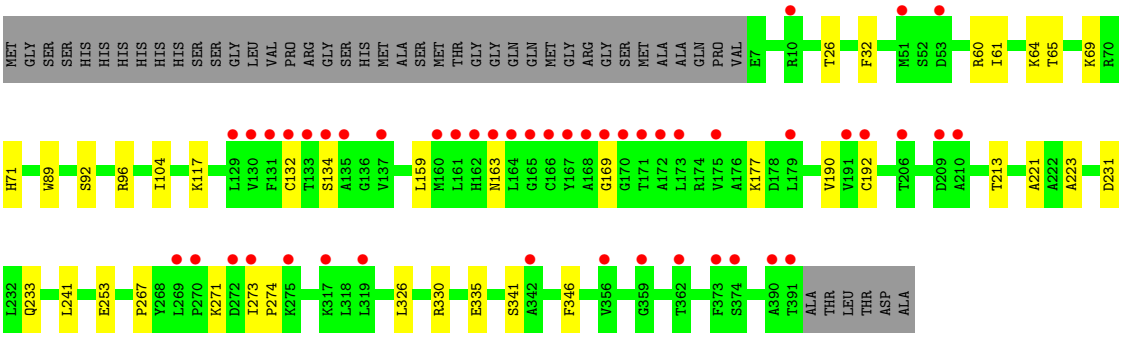
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	58	58	58	0	0
			Total	O		
2	B	30	30	30	0	0

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.

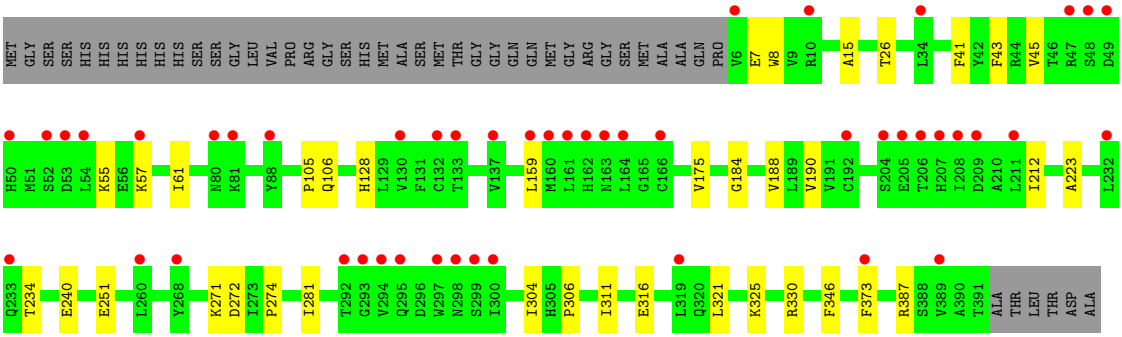
• Molecule 1: Type III polyketide synthase

Chain A:



• Molecule 1: Type III polyketide synthase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	146.41Å 65.20Å 108.39Å 90.00° 128.11° 90.00°	Depositor
Resolution (Å)	40.51 – 2.20 40.51 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.5 (40.51-2.20) 99.5 (40.51-2.20)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.19 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.213 , 0.258 0.213 , 0.258	Depositor DCC
R_{free} test set	2126 reflections (5.21%)	wwPDB-VP
Wilson B-factor (Å^2)	42.6	Xtriage
Anisotropy	0.280	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 42.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5966	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.97% 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)

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	0.46	1/2994 (0.0%)	0.59	0/4067
1	B	0.39	0/3003	0.54	0/4083
All	All	0.43	1/5997 (0.0%)	0.56	0/8150

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	132	CYS	CB-SG	-5.97	1.72	1.81

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2935	0	2964	19	0
1	B	2943	0	2955	22	0
2	A	58	0	0	0	0
2	B	30	0	0	0	0
All	All	5966	0	5919	40	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 3.

All (40) 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:117:LYS:HE3	1:B:234:THR:HA	1.80	0.62
1:B:304:ILE:HG23	1:B:373[A]:PHE:HD1	1.64	0.61
1:B:159:LEU:HD23	1:B:175:VAL:HG11	1.83	0.60
1:B:251:GLU:O	1:B:271:LYS:NZ	2.37	0.58
1:A:26:THR:HB	1:A:346:PHE:CZ	2.41	0.54
1:B:15:ALA:HB2	1:B:184:GLY:HA2	1.90	0.54
1:B:304:ILE:HG23	1:B:373[A]:PHE:CD1	2.42	0.53
1:B:316:GLU:HG3	1:B:321:LEU:HB2	1.91	0.53
1:B:274:PRO:HA	1:B:311:ILE:HG12	1.93	0.51
1:A:60:ARG:O	1:A:64:LYS:HD3	2.12	0.49
1:B:105:PRO:HD2	1:B:106:GLN:OE1	2.12	0.48
1:B:316:GLU:OE1	1:B:330:ARG:NH2	2.46	0.48
1:B:61:ILE:HD13	1:B:212:ILE:HD13	1.96	0.48
1:A:177:LYS:HE2	1:A:241:LEU:O	2.14	0.48
1:A:61:ILE:O	1:A:65:THR:HG23	2.14	0.47
1:B:26:THR:HB	1:B:346:PHE:CZ	2.49	0.47
1:A:190:VAL:O	1:A:223:ALA:HA	2.15	0.46
1:B:272:ASP:OD1	1:B:274:PRO:HD2	2.15	0.46
1:A:32:PHE:HA	1:A:71:HIS:HA	1.97	0.46
1:A:326:LEU:O	1:A:330:ARG:HG3	2.16	0.45
1:B:41:PHE:O	1:B:45:VAL:HG22	2.16	0.45
1:B:240:GLU:OE1	1:B:387:ARG:NH1	2.50	0.45
1:A:159:LEU:HD23	1:A:159:LEU:HA	1.86	0.44
1:B:43:PHE:CE2	1:B:55:LYS:HA	2.54	0.43
1:A:69:LYS:HG3	1:A:335:GLU:O	2.19	0.42
1:A:231:ASP:OD1	1:A:233:GLN:HG2	2.19	0.42
1:A:213:THR:HG21	1:A:267:PRO:HG3	2.01	0.42
1:A:273:ILE:HB	1:A:274:PRO:HD3	2.02	0.42
1:A:104:ILE:HD13	1:A:104:ILE:HA	1.97	0.41
1:B:281:ILE:HD11	1:B:373[B]:PHE:CE2	2.55	0.41
1:B:128:HIS:HB2	1:B:188:VAL:HG22	2.01	0.41
1:A:169:GLY:N	1:A:341:SER:HB2	2.36	0.41
1:B:7:GLU:HG2	1:B:8:TRP:H	1.86	0.41
1:B:57:LYS:HD2	1:B:57:LYS:HA	1.85	0.41
1:A:192:CYS:O	1:A:221:ALA:HA	2.21	0.41
1:A:253:GLU:O	1:A:271:LYS:HE2	2.20	0.41
1:B:325:LYS:HD3	1:B:325:LYS:HA	1.83	0.41
1:B:190:VAL:O	1:B:223:ALA:HA	2.20	0.41
1:A:134:SER:OG	1:A:163:ASN:HA	2.21	0.40
1:A:92:SER:O	1:A:96:ARG:HG3	2.20	0.40

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	383/431 (89%)	368 (96%)	15 (4%)	0	100	100
1	B	385/431 (89%)	365 (95%)	19 (5%)	1 (0%)	41	46
All	All	768/862 (89%)	733 (95%)	34 (4%)	1 (0%)	51	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	306	PRO

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	314/352 (89%)	313 (100%)	1 (0%)	92	97
1	B	313/352 (89%)	313 (100%)	0	100	100
All	All	627/704 (89%)	626 (100%)	1 (0%)	93	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

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 (i)

6.1 Protein, DNA and RNA chains (i)

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	A	385/431 (89%)	0.49	47 (12%)	4	3	33, 52, 85, 118	0
1	B	386/431 (89%)	0.53	48 (12%)	4	3	34, 61, 105, 133	0
All	All	771/862 (89%)	0.51	95 (12%)	4	3	33, 56, 100, 133	0

All (95) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	208	ILE	7.3
1	A	209	ASP	5.3
1	A	161	LEU	5.2
1	A	168	ALA	4.4
1	B	297	TRP	4.4
1	A	166	CYS	4.4
1	A	164	LEU	4.1
1	A	167	TYR	4.1
1	B	204	SER	4.1
1	B	54	LEU	3.9
1	B	132	CYS	3.9
1	B	299	SER	3.9
1	B	373[A]	PHE	3.7
1	B	53	ASP	3.7
1	A	269	LEU	3.7
1	A	162	HIS	3.6
1	A	10	ARG	3.6
1	B	49	ASP	3.5
1	A	160	MET	3.5
1	A	132	CYS	3.5
1	A	137	VAL	3.4
1	B	268	TYR	3.4
1	B	161	LEU	3.4
1	A	270	PRO	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	164	LEU	3.3
1	B	6	VAL	3.3
1	B	10	ARG	3.2
1	B	209	ASP	3.2
1	A	171	THR	3.2
1	A	169	GLY	3.1
1	B	211	LEU	2.9
1	A	390	ALA	2.9
1	A	272	ASP	2.9
1	A	173	LEU	2.9
1	B	298	ASN	2.9
1	A	130	VAL	2.9
1	A	373	PHE	2.8
1	A	374	SER	2.8
1	A	210	ALA	2.8
1	A	273	ILE	2.8
1	B	166	CYS	2.8
1	B	207	HIS	2.8
1	B	162	HIS	2.7
1	B	293	GLY	2.7
1	A	131	PHE	2.7
1	B	130	VAL	2.7
1	A	175	VAL	2.7
1	A	191	VAL	2.7
1	B	206	THR	2.7
1	A	134	SER	2.6
1	A	163	ASN	2.6
1	A	172	ALA	2.6
1	A	165	GLY	2.6
1	B	137	VAL	2.6
1	A	135	ALA	2.5
1	B	292	THR	2.5
1	B	192	CYS	2.5
1	A	133	THR	2.5
1	A	317	LYS	2.5
1	B	34	LEU	2.5
1	B	159	LEU	2.5
1	B	300	ILE	2.4
1	B	260	LEU	2.4
1	A	206	THR	2.4
1	B	47	ARG	2.4
1	A	275	LYS	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	53	ASP	2.4
1	B	232	LEU	2.4
1	B	319	LEU	2.3
1	B	133	THR	2.3
1	B	160	MET	2.3
1	B	80	ASN	2.3
1	B	389	VAL	2.3
1	A	192	CYS	2.3
1	A	391	THR	2.3
1	B	52	SER	2.3
1	B	81	LYS	2.2
1	B	57	LYS	2.2
1	A	362	THR	2.2
1	B	50	HIS	2.2
1	A	179	LEU	2.2
1	A	359	GLY	2.2
1	B	48	SER	2.2
1	B	163	ASN	2.2
1	B	294	VAL	2.2
1	B	295	GLN	2.2
1	B	88	TYR	2.1
1	A	356	VAL	2.1
1	B	233	GLN	2.1
1	A	170	GLY	2.1
1	A	129	LEU	2.1
1	A	51	MET	2.1
1	B	205	GLU	2.1
1	A	342	ALA	2.1
1	A	319	LEU	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 monosaccharides 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.