Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 11:56 AM GMT

PDB ID	:	3QIT
Title	:	Thioesterase Domain From Curacin Biosynthetic Pathway
Authors	:	Gehret, J.J.; Smith, J.L.
Deposited on	:	2011-01-27
Resolution	:	1.68 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
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 (1)) were used in the production of this report:

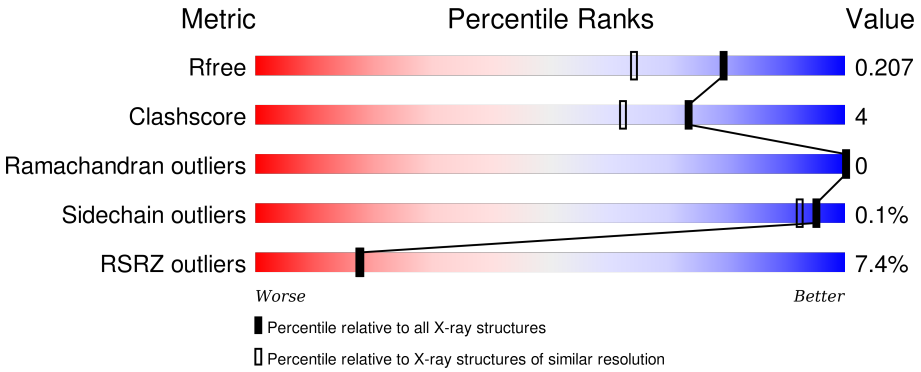
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.68 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}	91344	4802 (1.70-1.66)
Clashscore	102246	5317 (1.70-1.66)
Ramachandran outliers	100387	5225 (1.70-1.66)
Sidechain outliers	100360	5224 (1.70-1.66)
RSRZ outliers	91569	4813 (1.70-1.66)

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.

Mol	Chain	Length	Quality of chain
1	A	286	3% 90% 8% ·
1	B	286	4% 87% 12% ·
1	C	286	15% 89% 6% 6%
1	D	286	7% 90% 9% ·

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 10020 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 Polyketide synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	281	Total	C	N	O	S	0	12	0
			2218	1412	380	415	11			
1	B	283	Total	C	N	O	S	0	14	0
			2244	1438	380	414	12			
1	C	270	Total	C	N	O	S	0	5	0
			2112	1343	366	392	11			
1	D	283	Total	C	N	O	S	0	13	0
			2243	1432	384	415	12			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP D0E8E2
A	-1	ASN	-	EXPRESSION TAG	UNP D0E8E2
A	0	ALA	-	EXPRESSION TAG	UNP D0E8E2
B	-2	SER	-	EXPRESSION TAG	UNP D0E8E2
B	-1	ASN	-	EXPRESSION TAG	UNP D0E8E2
B	0	ALA	-	EXPRESSION TAG	UNP D0E8E2
C	-2	SER	-	EXPRESSION TAG	UNP D0E8E2
C	-1	ASN	-	EXPRESSION TAG	UNP D0E8E2
C	0	ALA	-	EXPRESSION TAG	UNP D0E8E2
D	-2	SER	-	EXPRESSION TAG	UNP D0E8E2
D	-1	ASN	-	EXPRESSION TAG	UNP D0E8E2
D	0	ALA	-	EXPRESSION TAG	UNP D0E8E2

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	371	Total	O	0	0
			371	371		
2	B	339	Total	O	0	0
			339	339		

Continued on next page...

Continued from previous page...

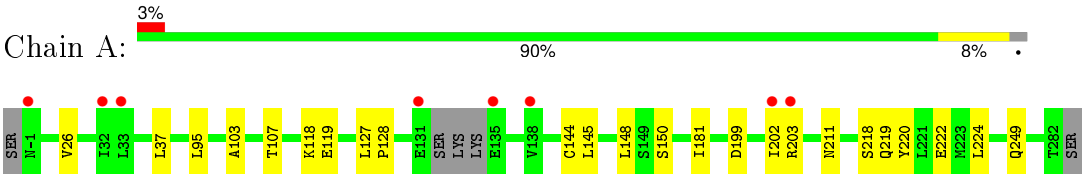
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	193	Total O 193 193	0	0
2	D	300	Total O 300 300	0	0

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.

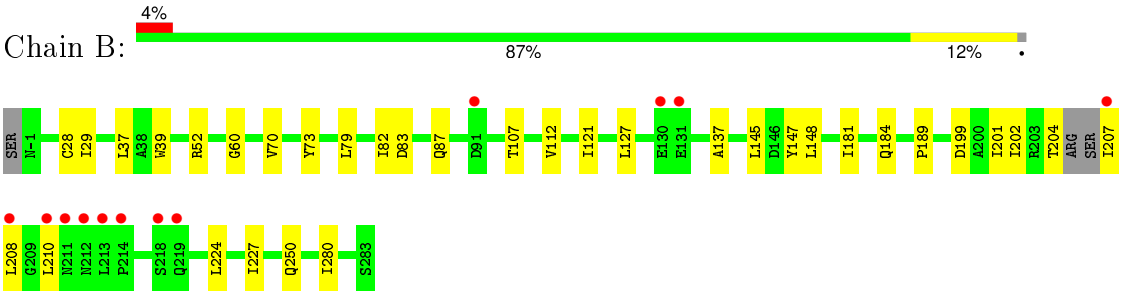
- Molecule 1: Polyketide synthase

Chain A:



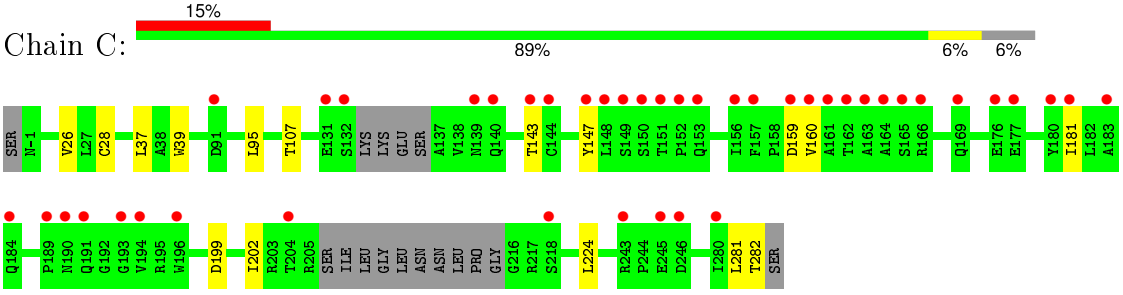
- Molecule 1: Polyketide synthase

Chain B:



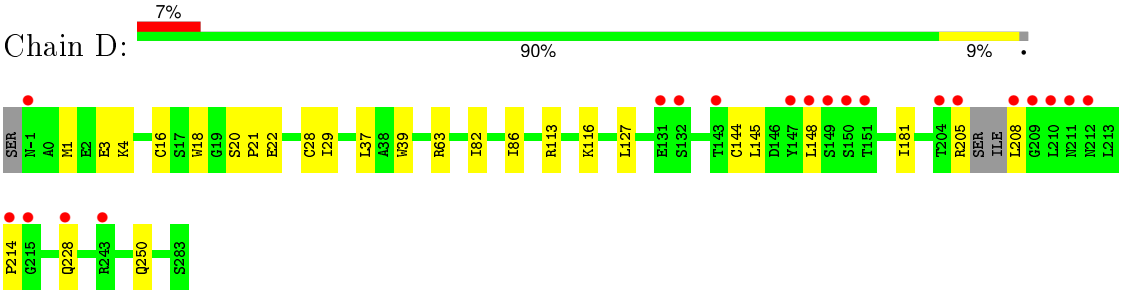
- Molecule 1: Polyketide synthase

Chain C:



- Molecule 1: Polyketide synthase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	74.56Å 86.91Å 87.59Å 90.00° 90.77° 90.00°	Depositor
Resolution (Å)	39.10 – 1.68 37.54 – 1.68	Depositor EDS
% Data completeness (in resolution range)	90.2 (39.10-1.68) 90.3 (37.54-1.68)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	1.37 (at 1.68Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.164 , 0.208 0.162 , 0.207	Depositor DCC
R_{free} test set	5806 reflections (5.34%)	DCC
Wilson B-factor (Å^2)	23.8	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 49.2	EDS
Estimated twinning fraction	0.000 for -h,l,k 0.012 for -h,-l,-k 0.019 for h,-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 114524 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	10020	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.96% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.58	0/2298	0.66	0/3123
1	B	0.55	0/2330	0.62	0/3165
1	C	0.48	0/2169	0.58	0/2946
1	D	0.54	0/2326	0.62	0/3155
All	All	0.54	0/9123	0.62	0/12389

There are no bond length outliers.

There are no bond angle outliers.

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	2218	0	2276	20	0
1	B	2244	0	2334	25	0
1	C	2112	0	2152	8	0
1	D	2243	0	2325	26	0
2	A	371	0	0	3	0
2	B	339	0	0	1	0
2	C	193	0	0	0	0
2	D	300	0	0	1	0
All	All	10020	0	9087	74	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.

All (74) 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:203[B]:ARG:HG3	1:A:203[B]:ARG:HH11	1.22	1.04
1:D:1[B]:MET:HE3	1:D:18:TRP:CZ2	1.96	1.00
1:D:1[B]:MET:CE	1:D:18:TRP:CZ2	2.45	1.00
1:D:1[B]:MET:HE2	1:D:18:TRP:CH2	2.00	0.97
1:D:3:GLU:OE2	1:D:63:ARG:NH1	2.06	0.88
1:D:1[B]:MET:CE	1:D:18:TRP:CH2	2.58	0.86
1:A:203[B]:ARG:HG3	1:A:203[B]:ARG:NH1	1.87	0.85
1:D:1[B]:MET:HE1	1:D:37:LEU:HA	1.59	0.84
1:D:1[B]:MET:HE3	1:D:18:TRP:HZ2	1.43	0.79
1:D:1[B]:MET:CE	1:D:37:LEU:HA	2.14	0.78
1:A:107[B]:THR:HG21	1:A:224:LEU:HD23	1.66	0.78
1:D:1[B]:MET:HE2	1:D:18:TRP:CZ2	2.15	0.77
1:B:107[B]:THR:HG21	1:B:224:LEU:HD23	1.69	0.73
1:D:86:ILE:HG21	1:D:116[B]:LYS:HD3	1.70	0.72
1:D:144:CYS:O	1:D:148:LEU:HG	1.93	0.69
1:A:203[B]:ARG:CG	1:A:203[B]:ARG:HH11	2.03	0.67
1:A:150:SER:O	2:A:306:HOH:O	2.10	0.67
1:B:121:ILE:HD13	1:B:280:ILE:HG12	1.75	0.67
1:A:37:LEU:HD21	1:A:181:ILE:HG22	1.77	0.66
1:B:73:TYR:HB3	1:B:210:LEU:HD21	1.79	0.65
1:B:145[B]:LEU:HD23	1:B:148:LEU:HD12	1.78	0.65
1:B:107[B]:THR:HG23	1:B:227[B]:ILE:HD11	1.78	0.65
1:D:29:ILE:HD12	1:D:82[B]:ILE:HD12	1.79	0.64
1:C:199:ASP:O	1:C:202:ILE:HG12	2.00	0.62
1:B:199:ASP:O	1:B:202[B]:ILE:HG12	1.99	0.61
1:D:1[B]:MET:HE2	1:D:16:CYS:HB3	1.84	0.60
1:B:204:THR:HG21	1:B:208:LEU:HG	1.86	0.58
1:C:107[B]:THR:HG21	1:C:224:LEU:HD23	1.85	0.58
1:A:103:ALA:O	1:A:107[A]:THR:HG23	2.04	0.57
1:B:127:LEU:HG	1:B:250:GLN:HE21	1.69	0.57
1:B:107[B]:THR:HG21	1:B:224:LEU:CD2	2.36	0.56
1:C:37:LEU:HD21	1:C:181:ILE:HG22	1.89	0.55
1:D:1[B]:MET:CE	1:D:16:CYS:HB3	2.37	0.54
1:B:29[B]:ILE:HD13	1:B:82:ILE:HD13	1.89	0.54
1:B:204:THR:OG1	1:B:207:ILE:HB	2.08	0.53
1:B:107[B]:THR:CG2	1:B:227[B]:ILE:HD11	2.39	0.52
1:D:113:ARG:HB3	1:D:116[B]:LYS:HD2	1.92	0.52
1:D:127:LEU:HG	1:D:250:GLN:HE21	1.76	0.51
1:D:1[B]:MET:HE1	1:D:37:LEU:CA	2.36	0.51
1:A:219:GLN:HG3	1:B:137:ALA:HB1	1.93	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:20:SER:OG	1:D:22:GLU:HG2	2.11	0.51
1:A:218:SER:O	1:A:222[A]:GLU:HG3	2.11	0.50
1:C:281:LEU:O	1:C:282:THR:HB	2.13	0.49
1:A:95:LEU:HD22	1:A:119:GLU:HB2	1.95	0.49
1:D:37:LEU:HD21	1:D:181:ILE:HG22	1.96	0.48
1:B:147:TYR:CE2	1:B:204:THR:HG21	2.50	0.47
1:B:60:GLY:HA3	1:B:202[A]:ILE:HD11	1.97	0.47
1:A:144:CYS:O	1:A:148:LEU:HG	2.15	0.46
1:A:145:LEU:HD12	1:B:112:VAL:HG11	1.97	0.46
1:B:70:VAL:HG13	1:B:201:ILE:HG12	1.98	0.45
1:D:1[B]:MET:HE3	1:D:37:LEU:HA	1.94	0.45
1:A:211:ASN:HB2	2:A:348:HOH:O	2.17	0.44
1:B:189:PRO:HB3	1:D:21:PRO:HB2	2.00	0.44
1:D:228:GLN:HB2	2:D:536:HOH:O	2.17	0.44
1:A:199:ASP:O	1:A:202:ILE:HG12	2.18	0.44
1:D:28:CYS:HB3	1:D:39:TRP:CE2	2.53	0.43
1:C:28:CYS:HB3	1:C:39:TRP:CE2	2.53	0.43
1:B:83:ASP:O	1:B:87:GLN:HG2	2.19	0.43
1:D:145:LEU:HD23	1:D:148:LEU:HD12	1.99	0.42
1:A:145:LEU:HD23	1:A:148:LEU:HD12	2.01	0.42
1:C:26:VAL:HG22	1:C:95:LEU:HD12	2.01	0.42
1:B:184:GLN:HA	1:D:4[B]:LYS:HE2	2.02	0.42
1:B:37:LEU:HD21	1:B:181:ILE:HG22	2.02	0.42
1:A:118:LYS:HE3	2:A:1173:HOH:O	2.20	0.42
1:A:128:PRO:HB3	1:A:220:TYR:HD2	1.83	0.41
1:C:159:ASP:CG	1:C:160:VAL:H	2.24	0.41
1:B:52:ARG:NH1	2:B:534:HOH:O	2.52	0.41
1:B:28:CYS:HB3	1:B:39:TRP:CE2	2.55	0.41
1:A:127:LEU:HD11	1:A:249:GLN:HB3	2.03	0.41
1:D:208:LEU:HD23	1:D:214:PRO:HD2	2.03	0.41
1:C:143:THR:O	1:C:147:TYR:HD2	2.04	0.40
1:B:204:THR:CG2	1:B:208:LEU:HG	2.51	0.40
1:A:26:VAL:HG22	1:A:95:LEU:HD12	2.04	0.40
1:A:145:LEU:HD11	1:B:79[B]:LEU:HG	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	289/286 (101%)	283 (98%)	6 (2%)	0	100	100
1	B	293/286 (102%)	284 (97%)	9 (3%)	0	100	100
1	C	269/286 (94%)	260 (97%)	9 (3%)	0	100	100
1	D	292/286 (102%)	286 (98%)	6 (2%)	0	100	100
All	All	1143/1144 (100%)	1113 (97%)	30 (3%)	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	
1	A	250/243 (103%)	250 (100%)	0	100	100
1	B	254/243 (104%)	254 (100%)	0	100	100
1	C	234/243 (96%)	234 (100%)	0	100	100
1	D	253/243 (104%)	252 (100%)	1 (0%)	93	90
All	All	991/972 (102%)	990 (100%)	1 (0%)	95	92

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

Mol	Chain	Res	Type
1	D	205	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no

such sidechains identified.

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

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	281/286 (98%)	0.02	8 (2%) 56 59	12, 21, 41, 91	0
1	B	283/286 (98%)	0.08	12 (4%) 40 41	14, 24, 50, 111	0
1	C	270/286 (94%)	0.77	43 (15%) 3 2	18, 34, 69, 98	0
1	D	283/286 (98%)	0.18	20 (7%) 19 19	14, 24, 58, 85	0
All	All	1117/1144 (97%)	0.26	83 (7%) 17 18	12, 25, 58, 111	0

All (83) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	148	LEU	12.5
1	C	150	SER	8.5
1	A	135	GLU	7.3
1	D	208	LEU	7.1
1	C	163	ALA	7.1
1	D	210	LEU	6.7
1	B	211	ASN	6.6
1	B	207	ILE	6.4
1	B	208	LEU	6.4
1	D	209	GLY	6.2
1	D	215	GLY	6.0
1	C	149	SER	5.8
1	C	147	TYR	5.8
1	C	180	TYR	5.5
1	A	131	GLU	5.3
1	B	210	LEU	5.2
1	C	162	THR	5.2
1	C	132	SER	5.1
1	B	212	ASN	4.8
1	C	156	ILE	4.7
1	C	183	ALA	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	161	ALA	4.4
1	D	148	LEU	4.3
1	C	193	GLY	4.0
1	B	214	PRO	3.9
1	D	214	PRO	3.8
1	C	151	THR	3.7
1	C	131	GLU	3.5
1	C	243	ARG	3.5
1	C	140	GLN	3.4
1	C	165	SER	3.4
1	A	203[A]	ARG	3.3
1	D	149	SER	3.3
1	D	151	THR	3.2
1	C	190	ASN	3.2
1	D	132	SER	3.1
1	C	143	THR	3.1
1	D	211	ASN	3.0
1	D	204	THR	3.0
1	D	-1	ASN	3.0
1	C	181	ILE	3.0
1	D	205	ARG	3.0
1	C	144	CYS	3.0
1	C	139	ASN	3.0
1	A	-1	ASN	2.9
1	C	184	GLN	2.8
1	C	157	PHE	2.8
1	C	159	ASP	2.8
1	C	160	VAL	2.8
1	C	153	GLN	2.7
1	C	245	GLU	2.7
1	C	194	VAL	2.6
1	C	196	TRP	2.6
1	B	218	SER	2.6
1	C	189	PRO	2.6
1	A	138	VAL	2.6
1	C	204	THR	2.5
1	B	131	GLU	2.5
1	C	176	GLU	2.5
1	B	219	GLN	2.4
1	C	177	GLU	2.4
1	C	169	GLN	2.4
1	B	130	GLU	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	91	ASP	2.4
1	D	212	ASN	2.3
1	A	32	ILE	2.3
1	D	131	GLU	2.3
1	D	150	SER	2.2
1	C	166	ARG	2.2
1	C	91	ASP	2.2
1	D	143	THR	2.2
1	D	147	TYR	2.2
1	C	246	ASP	2.2
1	C	152	PRO	2.2
1	A	33	LEU	2.1
1	B	213	LEU	2.1
1	C	280	ILE	2.1
1	D	243	ARG	2.1
1	C	218	SER	2.1
1	C	164	ALA	2.1
1	A	202	ILE	2.1
1	D	228	GLN	2.0
1	C	191	GLN	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 carbohydrates 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.