WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Feb 15, 2024 – 09:10 AM EST

PDB ID	:	3PNO
Title	:	Crystal Structure of E.coli Dha kinase DhaK (H56N)
Authors	:	Shi, R.; McDonald, L.; Matte, A.; Cygler, M.; Ekiel, I.; Montreal-Kingston Bacterial Structural Genomics Initiative (BSGI)
Deposited on	:	2010-11-19
Resolution	:	1.97 Å(reported)

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

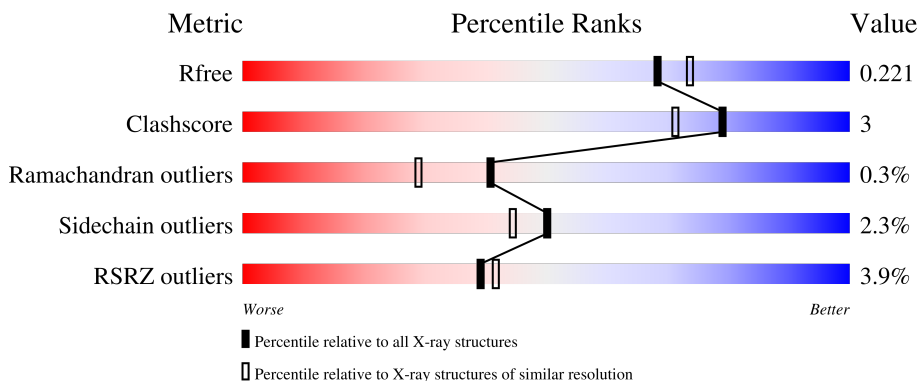|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.97 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	Percentile Ranks	Value
Rfree		0.221
Clashscore		3
Ramachandran outliers		0.3%
Sidechain outliers		2.3%
RSRZ outliers		3.9%

Worse ← → *Better*

▌ Percentile relative to all X-ray structures

▊ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	357	3% / 87% / 7% / 6%
1	B	357	4% / 85% / 10% / 6%
1	C	357	4% / 89% / 6% / 5%
1	D	357	4% / 89% / 6% · 5%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 11284 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 PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit dhaK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	336	Total	C	N	O	S	0	2	0
			2547	1595	436	503	13			
1	B	336	Total	C	N	O	S	0	5	0
			2559	1605	436	505	13			
1	C	339	Total	C	N	O	S	0	6	0
			2589	1624	441	511	13			
1	D	339	Total	C	N	O	S	0	2	0
			2571	1611	439	508	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP P76015
A	1	SER	-	expression tag	UNP P76015
A	56	ASN	HIS	engineered mutation	UNP P76015
B	0	GLY	-	expression tag	UNP P76015
B	1	SER	-	expression tag	UNP P76015
B	56	ASN	HIS	engineered mutation	UNP P76015
C	0	GLY	-	expression tag	UNP P76015
C	1	SER	-	expression tag	UNP P76015
C	56	ASN	HIS	engineered mutation	UNP P76015
D	0	GLY	-	expression tag	UNP P76015
D	1	SER	-	expression tag	UNP P76015
D	56	ASN	HIS	engineered mutation	UNP P76015

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	293	Total	O	0	0
			293	293		
2	B	206	Total	O	0	0
			206	206		

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	274	Total 274	O 274	0	0
2	D	245	Total 245	O 245	0	0

3 Residue-property plots ⓘ

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.
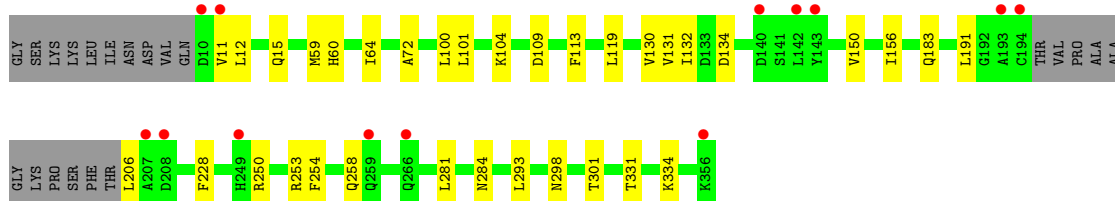
- Molecule 1: PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit dhaK
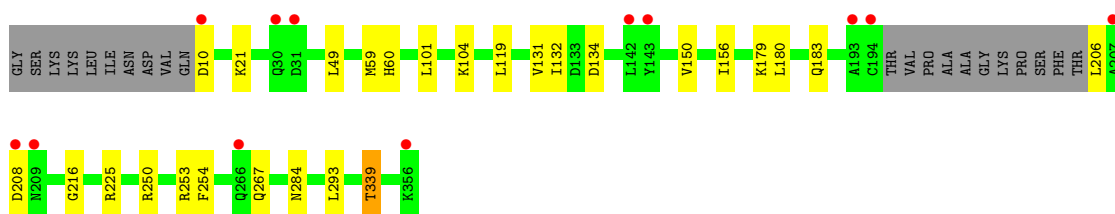
- Molecule 1: PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit dhaK

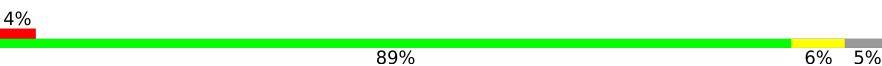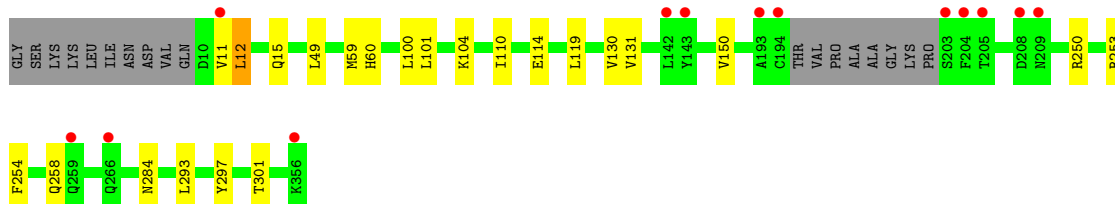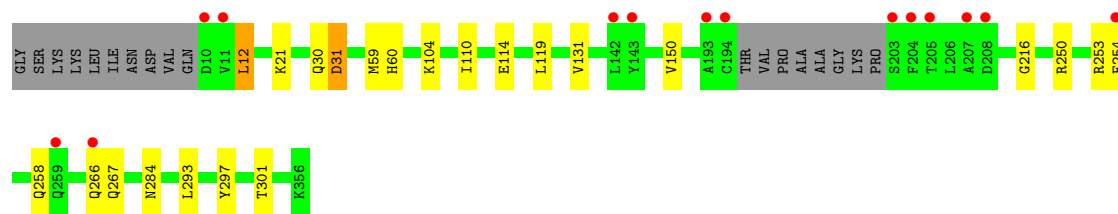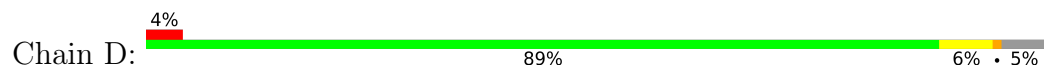- Molecule 1: PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit dhaK

- Molecule 1: PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit dhaK

Chain D:

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	59.69Å 82.56Å 92.86Å 77.91° 78.14° 71.07°	Depositor
Resolution (Å)	41.52 – 1.97 41.52 – 1.97	Depositor EDS
% Data completeness (in resolution range)	97.8 (41.52-1.97) 97.9 (41.52-1.97)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.62 (at 1.97Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.187 , 0.220 0.188 , 0.221	Depositor DCC
R_{free} test set	5617 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	15.8	Xtriage
Anisotropy	0.565	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 42.2	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11284	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.85% 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	0.47	0/2599	0.59	0/3530
1	B	0.43	0/2620	0.56	0/3559
1	C	0.46	0/2655	0.59	0/3607
1	D	0.45	0/2624	0.59	0/3564
All	All	0.46	0/10498	0.58	0/14260

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 (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	2547	0	2492	16	0
1	B	2559	0	2515	18	0
1	C	2589	0	2539	13	0
1	D	2571	0	2513	11	0
2	A	293	0	0	6	0
2	B	206	0	0	5	0
2	C	274	0	0	4	0
2	D	245	0	0	4	0
All	All	11284	0	10059	56	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 (56) 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:C:60:HIS:NE2	1:C:104:LYS:NZ	2.27	0.83
1:D:131:VAL:HG23	2:D:464:HOH:O	1.79	0.82
1:C:301[B]:THR:HG21	1:D:21:LYS:HD3	1.62	0.81
1:A:60:HIS:NE2	1:A:104:LYS:NZ	2.31	0.79
1:D:60:HIS:NE2	1:D:104:LYS:NZ	2.36	0.74
1:B:60:HIS:NE2	1:B:104:LYS:NZ	2.36	0.74
1:A:21:LYS:HD3	1:B:301[B]:THR:HG21	1.70	0.73
1:D:12:LEU:HD11	2:D:415:HOH:O	1.89	0.72
1:B:101[B]:LEU:HD23	1:B:113:PHE:HD2	1.56	0.70
1:D:31:ASP:HB3	2:D:729:HOH:O	1.91	0.70
1:C:131:VAL:HG23	2:C:397:HOH:O	1.93	0.67
1:B:131:VAL:HG23	2:B:516:HOH:O	1.93	0.67
1:A:49:LEU:HD22	1:A:101:LEU:HD22	1.80	0.62
1:C:254[A]:PHE:CD1	2:C:397:HOH:O	2.50	0.62
1:C:297:TYR:O	1:C:301[A]:THR:HG23	2.01	0.60
1:B:298:ASN:O	1:B:301[B]:THR:HG22	2.02	0.59
1:A:179:LYS:NZ	1:A:183:GLN:HE22	2.01	0.58
1:A:206:LEU:N	2:A:1010:HOH:O	2.38	0.57
1:A:131:VAL:HG23	2:A:461:HOH:O	2.06	0.56
1:B:254[A]:PHE:CD1	2:B:516:HOH:O	2.53	0.55
1:A:180:LEU:HD21	1:A:339:THR:HG21	1.90	0.53
1:B:100:LEU:HD11	1:B:130:VAL:HG23	1.92	0.51
1:D:284:ASN:HB2	1:D:293:LEU:HD11	1.92	0.50
1:C:254[A]:PHE:HD1	2:C:397:HOH:O	1.93	0.50
1:B:284:ASN:HB2	1:B:293:LEU:HD11	1.93	0.49
1:A:179:LYS:HZ2	1:A:183:GLN:HE22	1.59	0.48
1:A:254[A]:PHE:HE2	1:A:267:GLN:HE22	1.62	0.48
1:A:284:ASN:HB2	1:A:293:LEU:HD11	1.95	0.47
1:A:254[A]:PHE:HD1	2:A:461:HOH:O	1.97	0.47
1:C:12:LEU:HD21	2:C:906:HOH:O	2.13	0.47
1:D:254[A]:PHE:HE2	1:D:267:GLN:HE22	1.63	0.47
1:A:59:MET:HA	1:A:60:HIS:HA	1.76	0.47
1:B:254[A]:PHE:HD1	2:B:516:HOH:O	1.95	0.46
1:D:297:TYR:O	1:D:301:THR:HG23	2.15	0.46
1:A:225:ARG:CZ	2:A:1010:HOH:O	2.63	0.46
1:C:59:MET:HA	1:C:60:HIS:HA	1.78	0.45
1:B:183:GLN:HB3	1:B:334:LYS:HB3	1.98	0.45
1:B:134:ASP:HB3	2:B:358:HOH:O	2.15	0.45
1:C:284:ASN:HB2	1:C:293:LEU:HD11	1.99	0.44
1:B:132:ILE:HD12	1:B:156:ILE:HD12	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:59:MET:HA	1:D:60:HIS:HA	1.75	0.43
1:B:191:LEU:HD21	1:B:228:PHE:CD1	2.53	0.43
1:D:110:ILE:O	1:D:114:GLU:HG3	2.19	0.43
1:B:11:VAL:O	1:B:15:GLN:HG3	2.18	0.43
1:A:134:ASP:HB3	2:A:380:HOH:O	2.19	0.43
1:C:11:VAL:O	1:C:15:GLN:HG3	2.19	0.43
1:B:64:ILE:HD11	1:B:72:ALA:HB3	2.00	0.42
1:B:281:LEU:HB3	1:B:331:THR:HB	2.00	0.42
1:B:59:MET:HA	1:B:60:HIS:HA	1.73	0.42
1:A:132:ILE:HD12	1:A:156:ILE:HD12	2.02	0.42
1:D:216:GLY:HA2	2:D:384:HOH:O	2.20	0.42
1:B:206:LEU:N	2:B:501:HOH:O	2.52	0.41
1:C:110:ILE:O	1:C:114:GLU:HG3	2.21	0.41
1:A:216:GLY:HA2	2:A:480:HOH:O	2.19	0.41
1:C:100:LEU:HD11	1:C:130:VAL:HG23	2.03	0.40
1:C:49:LEU:HD22	1:C:101:LEU:HD22	2.02	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	334/357 (94%)	326 (98%)	7 (2%)	1 (0%)	41	29
1	B	337/357 (94%)	329 (98%)	7 (2%)	1 (0%)	41	29
1	C	341/357 (96%)	331 (97%)	9 (3%)	1 (0%)	41	29
1	D	337/357 (94%)	327 (97%)	9 (3%)	1 (0%)	41	29
All	All	1349/1428 (94%)	1313 (97%)	32 (2%)	4 (0%)	41	29

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	150	VAL
1	B	150	VAL
1	C	150	VAL
1	D	150	VAL

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	273/288 (95%)	267 (98%)	6 (2%)	52	46
1	B	276/288 (96%)	270 (98%)	6 (2%)	52	46
1	C	280/288 (97%)	275 (98%)	5 (2%)	59	51
1	D	276/288 (96%)	268 (97%)	8 (3%)	42	31
All	All	1105/1152 (96%)	1080 (98%)	25 (2%)	50	44

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	ASP
1	A	119	LEU
1	A	208	ASP
1	A	250	ARG
1	A	253	ARG
1	A	339	THR
1	B	12	LEU
1	B	109	ASP
1	B	119	LEU
1	B	250	ARG
1	B	253	ARG
1	B	258	GLN
1	C	12	LEU
1	C	119	LEU
1	C	250	ARG
1	C	253	ARG
1	C	258	GLN
1	D	12	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	30	GLN
1	D	31	ASP
1	D	119	LEU
1	D	250	ARG
1	D	253	ARG
1	D	258	GLN
1	D	266	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	183	GLN
1	A	249	HIS
1	B	218	HIS
1	B	305	GLN
1	C	183	GLN
1	C	263	GLN
1	C	266	GLN
1	D	249	HIS
1	D	263	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 ⓘ

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, 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	336/357 (94%)	-0.03	12 (3%)	42	45	7, 14, 29, 41	0
1	B	336/357 (94%)	0.07	13 (3%)	39	42	9, 18, 33, 43	0
1	C	339/357 (94%)	0.04	13 (3%)	40	43	7, 15, 31, 47	0
1	D	339/357 (94%)	0.17	14 (4%)	37	39	8, 17, 33, 47	0
All	All	1350/1428 (94%)	0.06	52 (3%)	39	42	7, 16, 32, 47	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	193	ALA	5.9
1	C	142	LEU	5.7
1	C	205	THR	5.4
1	D	208	ASP	5.3
1	D	142	LEU	5.2
1	B	142	LEU	5.0
1	D	205	THR	4.9
1	C	193	ALA	4.6
1	C	204	PHE	4.5
1	C	11	VAL	4.4
1	D	11	VAL	4.2
1	D	194	CYS	4.1
1	D	207	ALA	4.1
1	C	208	ASP	3.8
1	A	208	ASP	3.8
1	A	194	CYS	3.7
1	D	204	PHE	3.7
1	C	194	CYS	3.7
1	A	207	ALA	3.6
1	D	193	ALA	3.6
1	B	10	ASP	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	142	LEU	3.4
1	B	208	ASP	3.4
1	D	203	SER	3.1
1	B	143	TYR	3.1
1	A	30	GLN	3.1
1	D	259	GLN	3.0
1	B	266	GLN	2.9
1	C	203	SER	2.9
1	C	209	ASN	2.9
1	B	249	HIS	2.8
1	B	193	ALA	2.8
1	D	10	ASP	2.8
1	A	10	ASP	2.7
1	A	356	LYS	2.7
1	B	356	LYS	2.7
1	B	259	GLN	2.6
1	D	266	GLN	2.6
1	A	31	ASP	2.5
1	D	254[A]	PHE	2.5
1	B	194	CYS	2.5
1	A	209	ASN	2.4
1	C	356	LYS	2.3
1	A	266	GLN	2.3
1	A	143	TYR	2.2
1	C	266	GLN	2.2
1	C	259	GLN	2.1
1	B	207	ALA	2.1
1	D	143	TYR	2.1
1	B	140	ASP	2.0
1	C	143	TYR	2.0
1	B	11	VAL	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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.