Full wwPDB X-ray Structure Validation Report ⓘ

Aug 12, 2024 – 08:50 PM EDT

PDB ID	:	9CBS
Title	:	Crystal structure of Chaetomium thermophilum Gcn2 HisRS-like domain, catalytic domain
Authors	:	Zhang, J.
Deposited on	:	2024-06-20
Resolution	:	2.69 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:

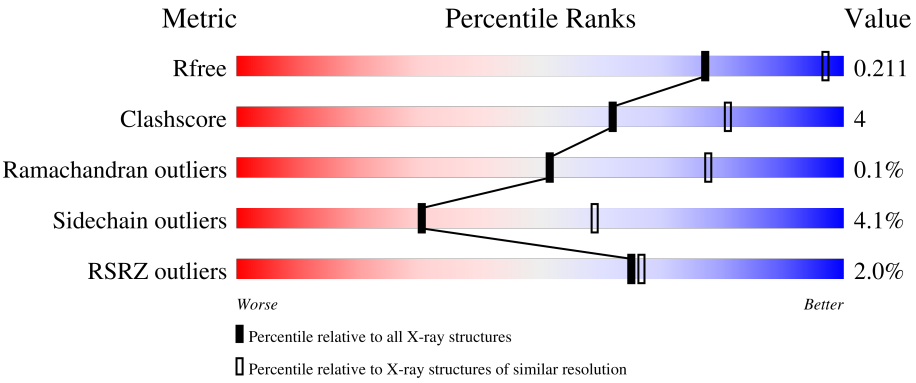
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.69 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	330	2% 88% 11% ••
1	B	330	2% 87% 11% ••
1	C	330	81% 14% •
1	D	330	4% 81% 13% • 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 10191 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 non-specific serine/threonine protein kinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	327	Total	C	N	O	S	Se	0	0	0
			2607	1667	455	476	3	6			
1	B	326	Total	C	N	O	S	Se	0	0	0
			2580	1648	450	473	3	6			
1	C	316	Total	C	N	O	S	Se	0	0	0
			2506	1604	440	453	3	6			
1	D	314	Total	C	N	O	S	Se	0	0	0
			2421	1552	424	436	3	6			

- Molecule 2 is water.

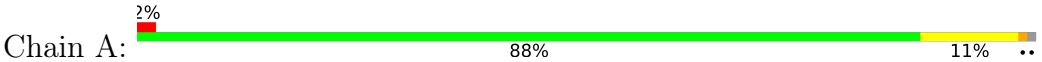
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	29	Total	O	0	0
			29	29		
2	B	21	Total	O	0	0
			21	21		
2	C	21	Total	O	0	0
			21	21		
2	D	6	Total	O	0	0
			6	6		

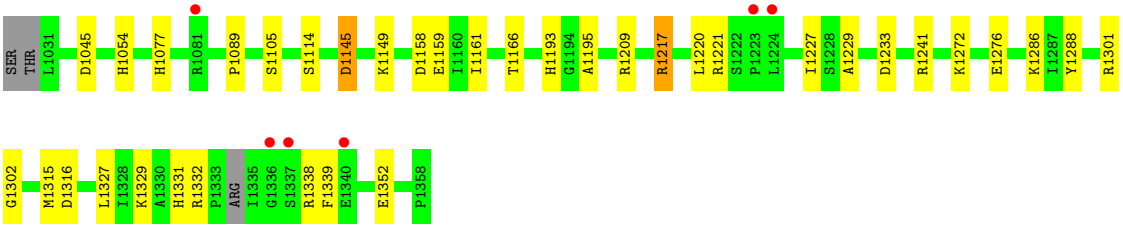
3 Residue-property plots (i)

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: non-specific serine/threonine protein kinase

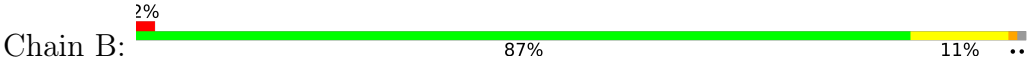
Chain A:

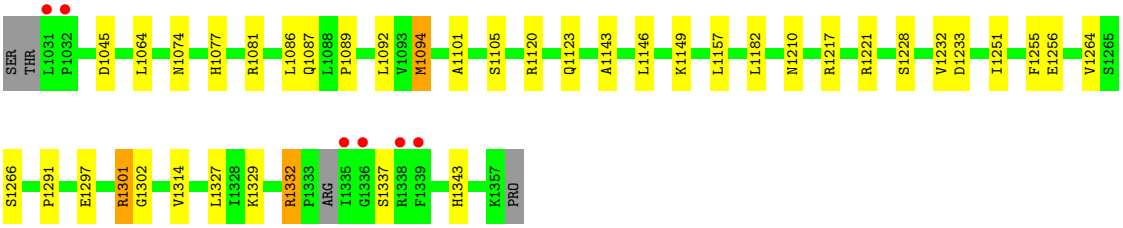




• Molecule 1: non-specific serine/threonine protein kinase

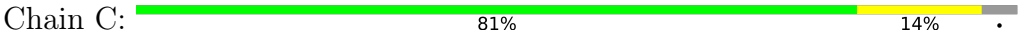
Chain B:

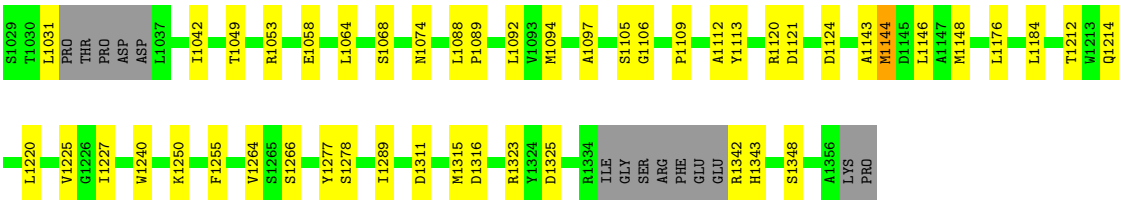




• Molecule 1: non-specific serine/threonine protein kinase

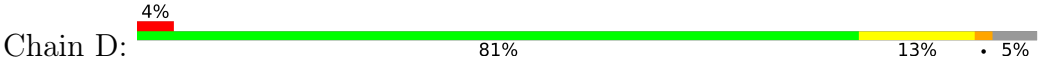
Chain C:

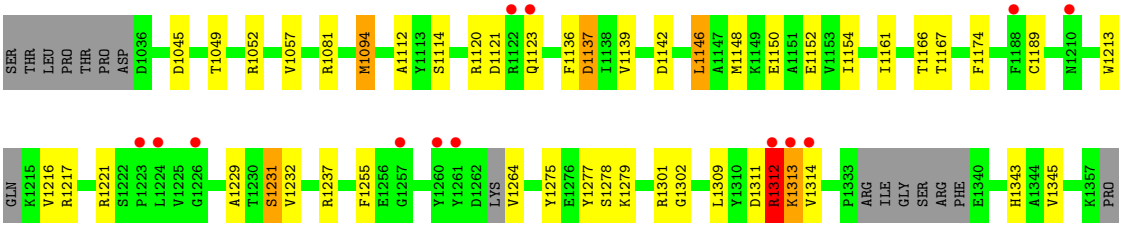




• Molecule 1: non-specific serine/threonine protein kinase

Chain D:




SER THR LEU PRO THR PRO ASP D1036 D1045 T1049 R1052 V1057 R1081 M1094 A1112 Y1113 S1114 R1120 D1121 R1122 Q1123 F1136 D1137 T1138 V1139 D1142 L1146 A1147 H1148 K1149 E1150 A1151 E1152 Y1153 I1154 I1161 T1166 T1167 F1174 F1188 C1189 N1210 W1213
GLN K1215 V1216 R1217 R1221 S1222 P1223 L1224 V1225 G1226 A1229 T1230 S1231 V1232 R1237 F1255 E1256 G1257 Y1260 Y1261 D1262 LYS V1264 Y1275 E1276 Y1277 S1278 K1279 R1301 G1302 L1309 Y1310 D1311 R1312 K1313 V1314 P1333 ARG ILE GLY SER ARG PHE E1340 H1343 A1344 V1345 K1357 PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	126.09Å 132.55Å 138.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.49 – 2.69 35.49 – 2.69	Depositor EDS
% Data completeness (in resolution range)	98.8 (35.49-2.69) 98.8 (35.49-2.69)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.96 (at 2.68Å)	Xtriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, R_{free}	0.177 , 0.210 0.180 , 0.211	Depositor DCC
R_{free} test set	1991 reflections (3.09%)	wwPDB-VP
Wilson B-factor (Å^2)	56.5	Xtriage
Anisotropy	0.538	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 33.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.019 for -h,l,k	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10191	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.78% 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	0.29	0/2662	0.54	0/3592
1	B	0.29	0/2634	0.54	0/3556
1	C	0.29	0/2556	0.55	0/3447
1	D	0.28	0/2470	0.57	0/3341
All	All	0.29	0/10322	0.55	0/13936

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	D	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	1311	ASP	Peptide
1	D	1312	ARG	Peptide

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	2607	0	2593	23	0
1	B	2580	0	2542	22	0
1	C	2506	0	2496	25	0
1	D	2421	0	2339	25	0
2	A	29	0	0	1	0
2	B	21	0	0	1	0
2	C	21	0	0	0	0
2	D	6	0	0	0	0
All	All	10191	0	9970	88	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 (88) 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:B:1210:ASN:HD21	1:B:1314:VAL:HG22	1.51	0.75
1:C:1109:PRO:HG2	1:C:1112:ALA:HB2	1.68	0.74
1:B:1217:ARG:NH2	1:B:1233:ASP:OD1	2.25	0.69
1:D:1150:GLU:O	1:D:1154:ILE:HD12	1.94	0.67
1:A:1209:ARG:NH1	1:A:1316:ASP:OD2	2.29	0.65
1:D:1313:LYS:HG2	1:D:1314:VAL:HG13	1.80	0.64
1:D:1161:ILE:O	1:D:1167:THR:OG1	2.17	0.63
1:C:1049:THR:O	1:C:1053:ARG:HG2	1.99	0.62
1:C:1184:LEU:HD23	1:C:1264:VAL:HG23	1.81	0.61
1:D:1148:MSE:HG3	1:D:1277:TYR:CE1	2.37	0.59
1:B:1143:ALA:HB2	1:B:1343:HIS:NE2	2.19	0.58
1:B:1143:ALA:HA	1:B:1146:LEU:HD21	1.85	0.58
1:B:1255:PHE:CE2	1:B:1264:VAL:HG21	2.39	0.58
1:D:1146:LEU:O	1:D:1150:GLU:HG3	2.03	0.58
1:D:1255:PHE:CE2	1:D:1264:VAL:HG21	2.39	0.57
1:B:1251:ILE:HD11	1:B:1291:PRO:HG3	1.86	0.57
1:B:1081:ARG:NH1	2:B:1401:HOH:O	2.38	0.56
1:B:1143:ALA:HB2	1:B:1343:HIS:CE1	2.41	0.56
1:C:1255:PHE:CE2	1:C:1264:VAL:HG21	2.40	0.56
1:C:1143:ALA:HB2	1:C:1343:HIS:CE1	2.41	0.55
1:D:1255:PHE:HE2	1:D:1264:VAL:HG21	1.72	0.54
1:A:1241:ARG:HH12	1:A:1315:MSE:HE1	1.74	0.53
1:A:1338:ARG:HH12	1:D:1081:ARG:HH11	1.55	0.52
1:C:1144:MSE:HE3	1:C:1144:MSE:HA	1.90	0.52
1:B:1228:SER:O	1:B:1232:VAL:HG23	2.10	0.52
1:A:1301:ARG:NH1	1:D:1052:ARG:O	2.43	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1338:ARG:NE	1:A:1339:PHE:O	2.39	0.51
1:A:1338:ARG:HH12	1:D:1081:ARG:NH1	2.09	0.50
1:C:1089:PRO:HD2	1:C:1092:LEU:HD12	1.93	0.50
1:C:1143:ALA:HA	1:C:1146:LEU:HD21	1.92	0.49
1:A:1327:LEU:HD22	1:A:1331:HIS:HD2	1.78	0.48
1:C:1184:LEU:HD23	1:C:1264:VAL:CG2	2.42	0.48
1:D:1137:ASP:HA	1:D:1345:VAL:O	2.14	0.48
1:C:1088:LEU:HD23	1:C:1097:ALA:HB1	1.95	0.48
1:C:1106:GLY:O	1:C:1342:ARG:NH2	2.48	0.47
1:C:1176:LEU:HB2	1:C:1289:ILE:HD13	1.95	0.47
1:A:1089:PRO:HB3	1:D:1049:THR:OG1	2.14	0.47
1:C:1255:PHE:CD2	1:C:1264:VAL:HG21	2.50	0.47
1:D:1301:ARG:HA	1:D:1302:GLY:HA2	1.54	0.47
1:D:1221:ARG:NH2	1:D:1229:ALA:HB2	2.30	0.46
1:D:1217:ARG:HG3	1:D:1232:VAL:HG11	1.97	0.46
1:B:1064:LEU:HD11	1:B:1074:ASN:HB3	1.97	0.46
1:A:1054:HIS:NE2	1:A:1159:GLU:HG3	2.31	0.45
1:C:1240:TRP:CE2	1:C:1250:LYS:HE2	2.52	0.45
1:B:1086:LEU:HD22	1:B:1101:ALA:HA	1.99	0.45
1:C:1058:GLU:HG3	1:C:1113:TYR:CE2	2.52	0.45
1:D:1094:MSE:H	1:D:1094:MSE:SE	2.49	0.45
1:A:1272:LYS:O	1:A:1276:GLU:HG2	2.17	0.44
1:A:1329:LYS:O	1:A:1332:ARG:HG2	2.17	0.44
1:A:1327:LEU:HD23	1:A:1327:LEU:HA	1.71	0.44
1:C:1148:MSE:HG3	1:C:1277:TYR:CE1	2.52	0.44
1:C:1323:ARG:HD3	1:C:1325:ASP:OD2	2.18	0.44
1:A:1301:ARG:HA	1:A:1302:GLY:HA2	1.55	0.44
1:D:1275:TYR:CZ	1:D:1279:LYS:HD3	2.52	0.44
1:B:1087:GLN:HG3	1:C:1042:ILE:HD11	2.00	0.43
1:D:1213:TRP:HA	1:D:1216:VAL:HB	2.00	0.43
1:D:1112:ALA:O	1:D:1136:PHE:O	2.36	0.43
1:D:1312:ARG:HB2	1:D:1313:LYS:H	1.66	0.43
1:D:1148:MSE:O	1:D:1152:GLU:HG3	2.17	0.43
1:A:1158:ASP:O	1:A:1161:ILE:HG12	2.18	0.43
1:A:1193:HIS:C	1:A:1195:ALA:H	2.22	0.42
1:B:1301:ARG:HA	1:B:1302:GLY:HA2	1.61	0.42
1:D:1139:VAL:HA	1:D:1343:HIS:O	2.19	0.42
1:C:1225:VAL:HG12	1:C:1227:ILE:HG13	1.99	0.42
1:C:1143:ALA:HB2	1:C:1343:HIS:NE2	2.33	0.42
1:D:1189:CYS:O	1:D:1231:SER:OG	2.28	0.42
1:A:1220:LEU:HD22	1:A:1227:ILE:HD12	2.01	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1327:LEU:HD12	1:B:1327:LEU:HA	1.77	0.42
1:B:1329:LYS:O	1:B:1332:ARG:HD3	2.18	0.42
1:B:1182:LEU:HD23	1:B:1297:GLU:HG3	2.00	0.42
1:B:1149:LYS:HA	1:B:1149:LYS:HD3	1.84	0.42
1:C:1311:ASP:HA	1:C:1315:MSE:HE3	2.01	0.42
1:A:1077:HIS:HD2	2:A:1408:HOH:O	2.02	0.42
1:B:1094:MSE:H	1:B:1094:MSE:SE	2.53	0.42
1:D:1121:ASP:N	1:D:1121:ASP:OD1	2.53	0.41
1:A:1217:ARG:NH1	1:A:1233:ASP:OD1	2.46	0.41
1:C:1220:LEU:HD22	1:C:1227:ILE:HD12	2.02	0.41
1:B:1221:ARG:HG2	1:B:1232:VAL:HG21	2.02	0.41
1:A:1145:ASP:O	1:A:1149:LYS:HG2	2.20	0.41
1:A:1221:ARG:HD2	1:A:1229:ALA:HA	2.01	0.41
1:A:1286:LYS:HD3	1:A:1288:TYR:CZ	2.56	0.41
1:C:1212:THR:OG1	1:C:1214:GLN:HG2	2.21	0.41
1:B:1089:PRO:HD2	1:B:1092:LEU:HD12	2.02	0.41
1:C:1031:LEU:HD23	1:C:1031:LEU:HA	1.97	0.40
1:C:1064:LEU:HD11	1:C:1074:ASN:HB3	2.03	0.40
1:A:1338:ARG:HH11	1:B:1123:GLN:HE22	1.68	0.40
1:A:1338:ARG:HH11	1:B:1123:GLN:NE2	2.20	0.40
1:D:1154:ILE:HG23	1:D:1174:PHE:CE1	2.57	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	323/330 (98%)	316 (98%)	7 (2%)	0	100	100
1	B	322/330 (98%)	316 (98%)	6 (2%)	0	100	100
1	C	310/330 (94%)	306 (99%)	4 (1%)	0	100	100
1	D	306/330 (93%)	296 (97%)	9 (3%)	1 (0%)	41	66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1261/1320 (96%)	1234 (98%)	26 (2%)	1 (0%)	51 78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	1137	ASP

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	280/277 (101%)	273 (98%)	7 (2%)	47 76
1	B	274/277 (99%)	263 (96%)	11 (4%)	31 60
1	C	267/277 (96%)	256 (96%)	11 (4%)	30 59
1	D	246/277 (89%)	231 (94%)	15 (6%)	18 41
All	All	1067/1108 (96%)	1023 (96%)	44 (4%)	30 59

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1045	ASP
1	A	1105	SER
1	A	1114	SER
1	A	1145	ASP
1	A	1166	THR
1	A	1217	ARG
1	A	1352	GLU
1	B	1045	ASP
1	B	1077	HIS
1	B	1094	MSE
1	B	1105	SER
1	B	1120	ARG
1	B	1157	LEU
1	B	1256	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1266	SER
1	B	1301	ARG
1	B	1332	ARG
1	B	1337	SER
1	C	1068	SER
1	C	1094	MSE
1	C	1105	SER
1	C	1120	ARG
1	C	1121	ASP
1	C	1124	ASP
1	C	1144	MSE
1	C	1266	SER
1	C	1278	SER
1	C	1316	ASP
1	C	1348	SER
1	D	1045	ASP
1	D	1057	VAL
1	D	1094	MSE
1	D	1114	SER
1	D	1120	ARG
1	D	1123	GLN
1	D	1142	ASP
1	D	1146	LEU
1	D	1166	THR
1	D	1231	SER
1	D	1237	ARG
1	D	1278	SER
1	D	1309	LEU
1	D	1312	ARG
1	D	1313	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1210	ASN
1	B	1236	GLN
1	C	1082	ASN
1	C	1210	ASN
1	C	1236	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, 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	321/330 (97%)	-0.21	6 (1%) 66 69	40, 54, 85, 131	0
1	B	320/330 (96%)	-0.16	6 (1%) 66 69	41, 63, 95, 128	0
1	C	310/330 (93%)	-0.28	0 100 100	41, 58, 90, 116	0
1	D	308/330 (93%)	0.02	13 (4%) 36 35	46, 78, 117, 146	0
All	All	1259/1320 (95%)	-0.16	25 (1%) 65 67	40, 61, 108, 146	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1336	GLY	4.1
1	D	1312	ARG	3.5
1	B	1339	PHE	3.3
1	D	1314	VAL	3.2
1	D	1261	TYR	3.1
1	B	1031	LEU	3.0
1	D	1224	LEU	2.8
1	D	1226	GLY	2.7
1	D	1257	GLY	2.7
1	A	1337	SER	2.7
1	B	1338	ARG	2.7
1	D	1313	LYS	2.6
1	B	1336	GLY	2.6
1	D	1188	PHE	2.6
1	B	1335	ILE	2.5
1	D	1223	PRO	2.4
1	D	1210	ASN	2.3
1	A	1081	ARG	2.2
1	B	1032	PRO	2.2
1	D	1123	GLN	2.2
1	D	1260	TYR	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1223	PRO	2.1
1	A	1340	GLU	2.1
1	A	1224	LEU	2.0
1	D	1122	ARG	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.