wwPDB X-ray Structure Validation Summary Report (i)

Mar 1, 2026 – 08:50 PM UTC

PDB ID	:	5I8I / pdb_00005i8i
Title	:	Crystal Structure of the K. lactis Urea Amidolyase
Authors	:	Zhao, J.; Xiang, S.
Deposited on	:	2016-02-19
Resolution	:	6.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i)) were used in the production of this report:

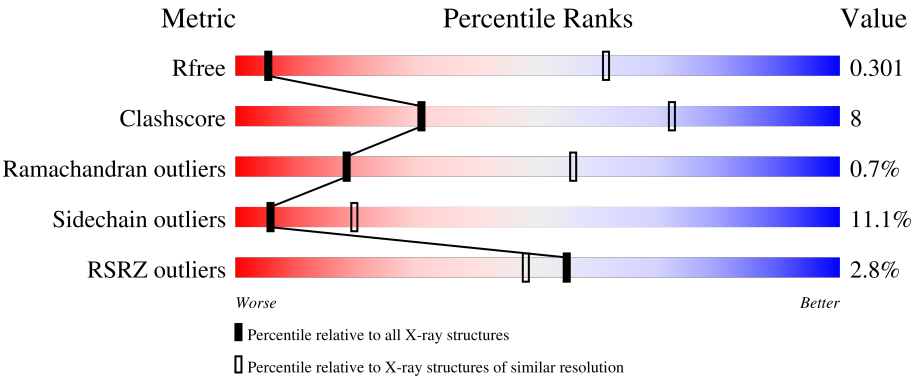
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 6.50 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	1162 (9.00-4.00)
Clashscore	190562	1020 (8.98-4.02)
Ramachandran outliers	187476	1050 (9.00-4.00)
Sidechain outliers	187428	1014 (9.00-4.00)
RSRZ outliers	180081	1155 (9.00-4.00)

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	1829	4% 69% 18% · 9%
1	B	1829	2% 66% 19% 5% · 9%
1	C	1829	2% 69% 18% · 9%
1	D	1829	2% 67% 19% · 9%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 51722 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 Urea Amidolyase.

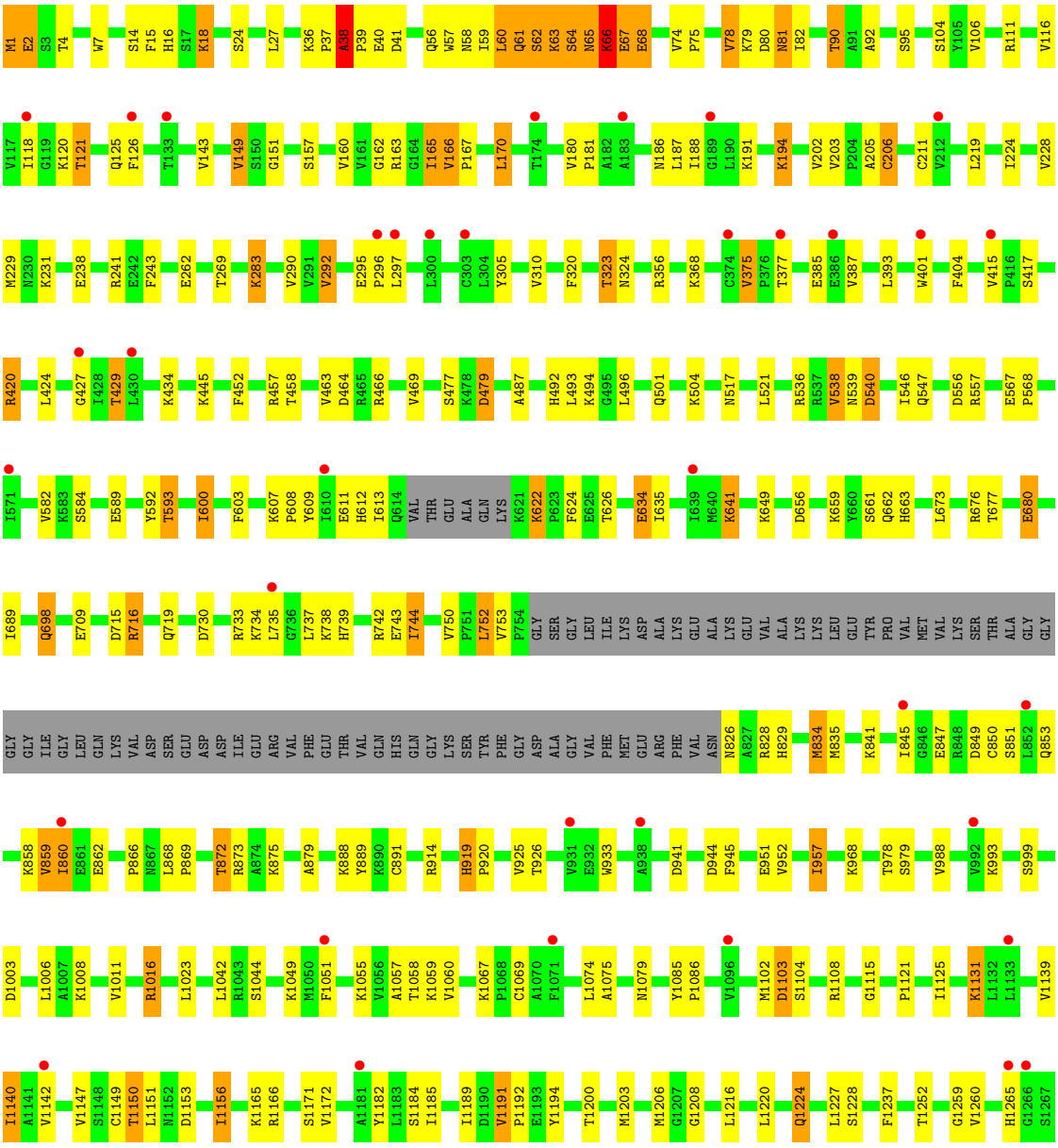
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1660	Total	C	N	O	S	0	0	0
			12939	8253	2183	2456	47			
1	B	1658	Total	C	N	O	S	0	0	0
			12922	8243	2181	2452	46			
1	C	1660	Total	C	N	O	S	0	0	0
			12939	8253	2183	2456	47			
1	D	1658	Total	C	N	O	S	0	0	0
			12922	8243	2181	2452	46			

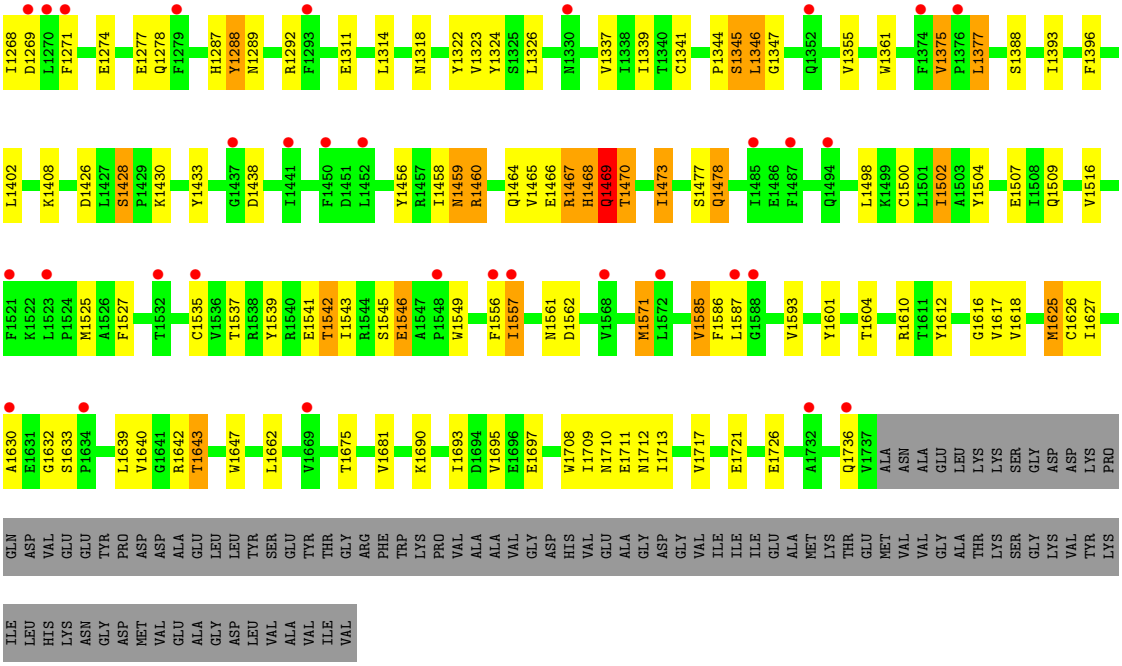
3 Residue-property plots (i)

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: Urea Amidolyase

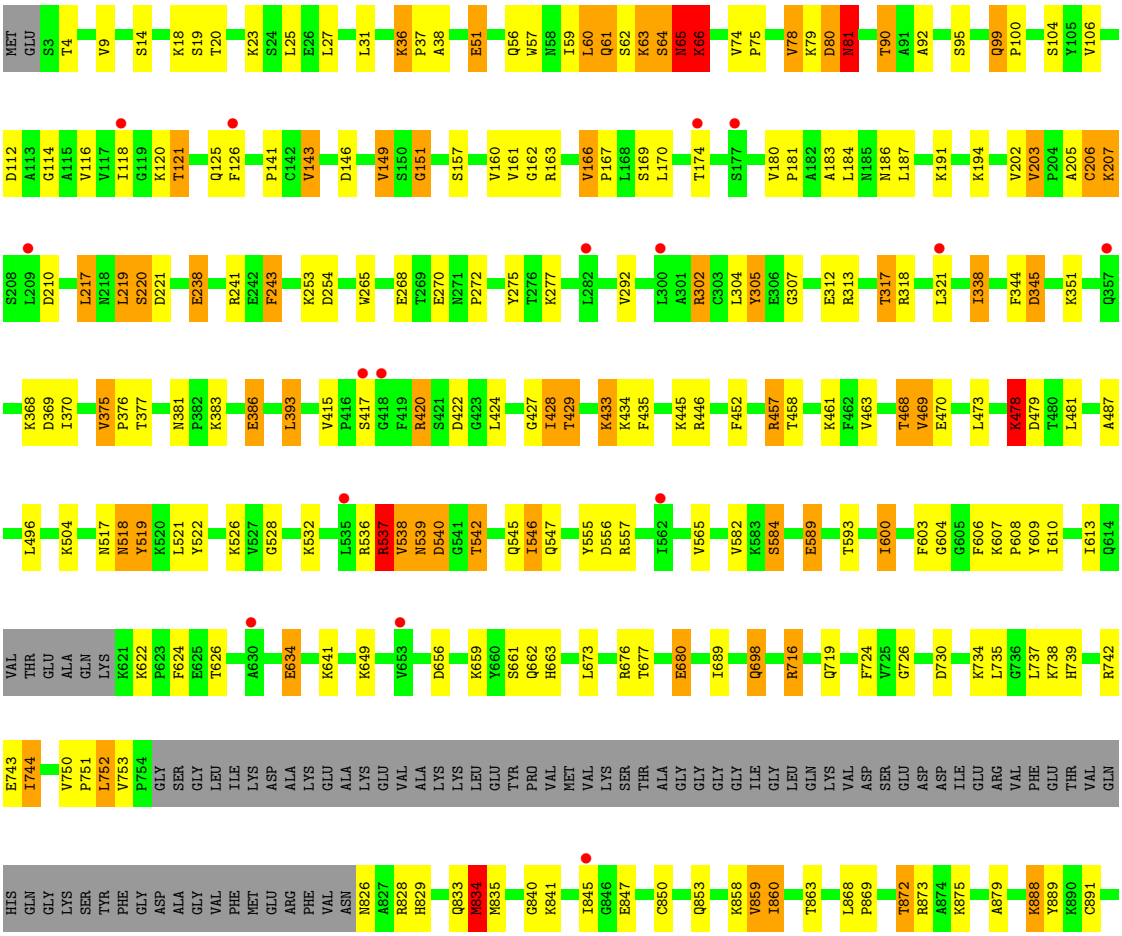
Chain A:

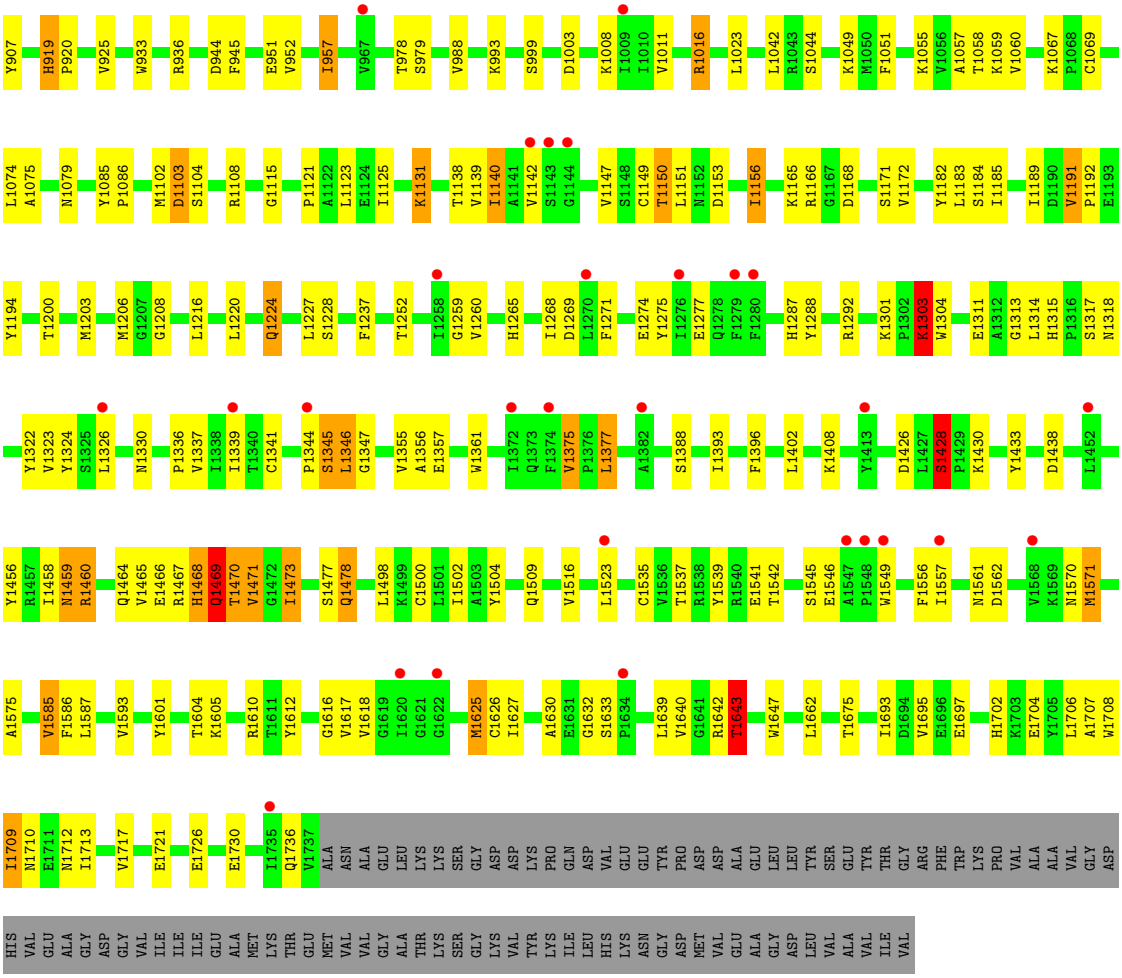




• Molecule 1: Urea Amidolyase

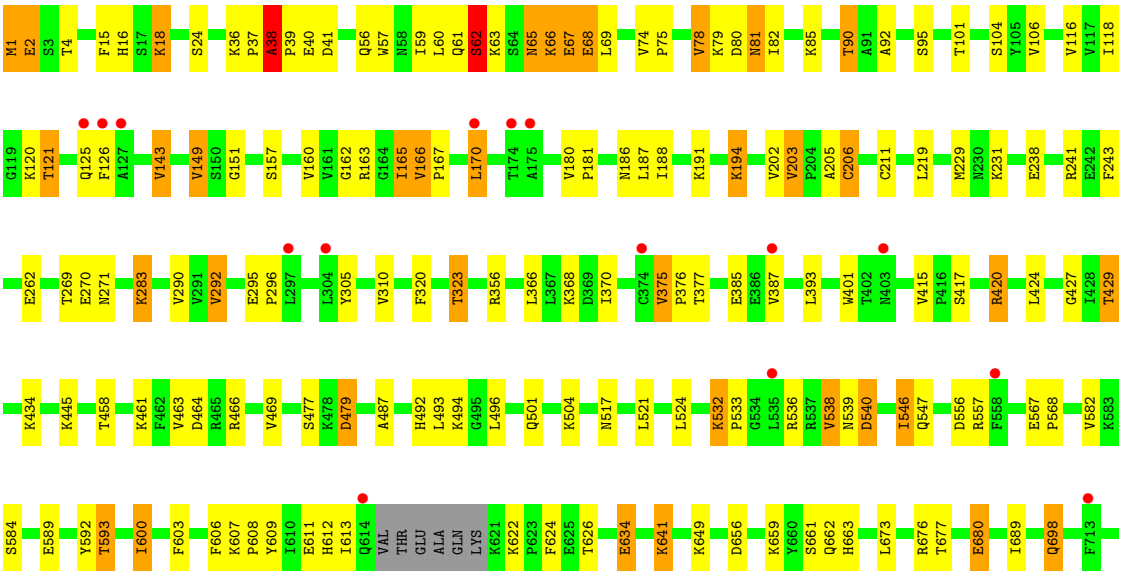
Chain B:

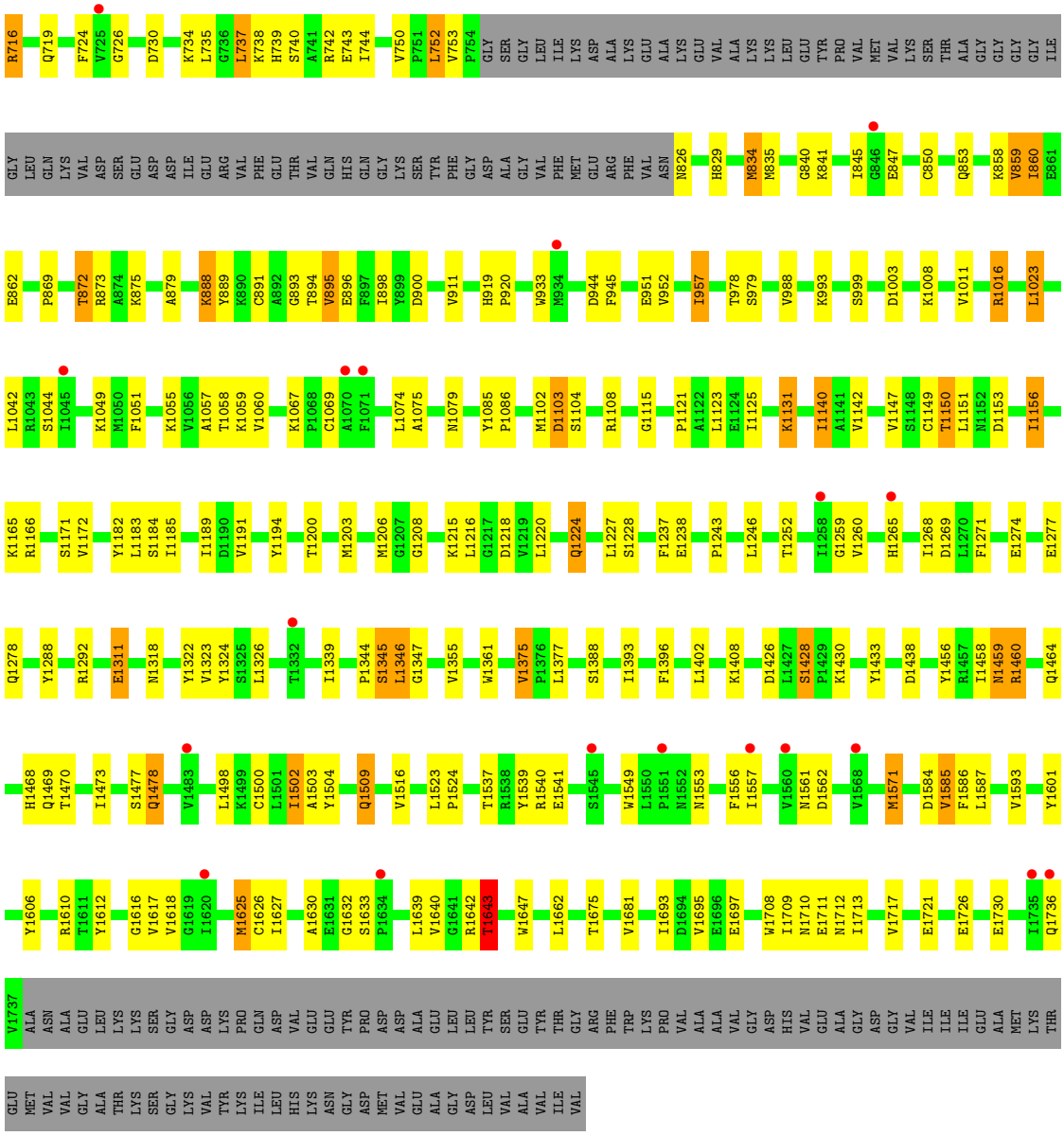




• Molecule 1: Urea Amidolyase

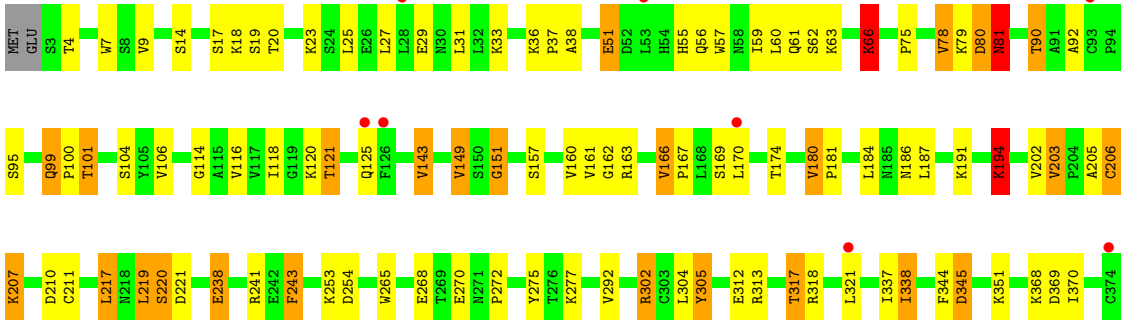
Chain C:

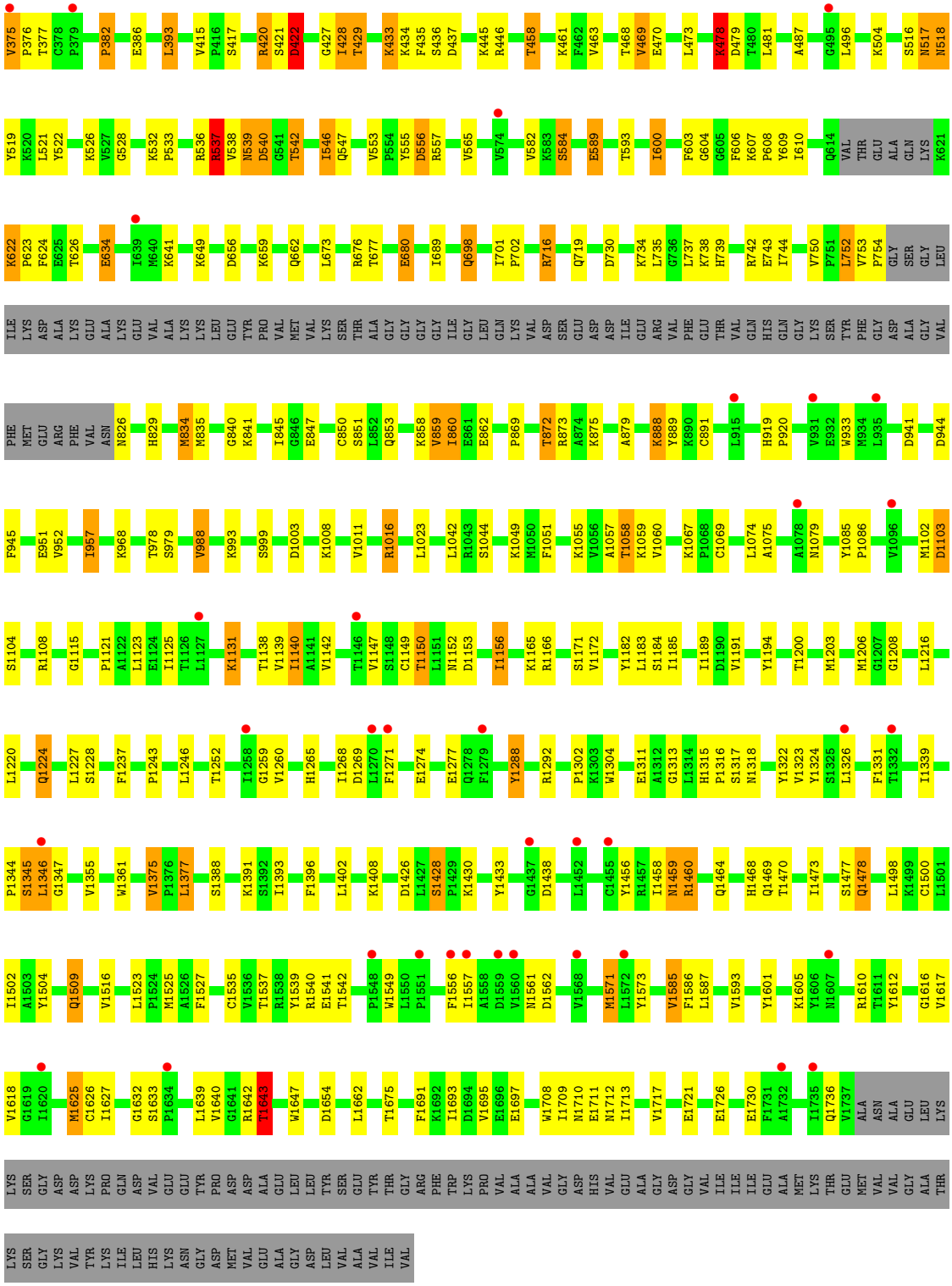




• Molecule 1: Urea Amidolyase

Chain D:




V375 P376 T377 C378 P379 P382 E386 L393 V415 P416 S417 R420 S421 D422 G427 I428 T429 K433 K434 F435 S436 D437 K445 R446 T468 K461 F462 V463 T468 V469 E470 L473 K478 D479 T480 L481 A487 G495 L496 K504 S516 N517 N518
Y519 K520 L521 Y522 K526 V527 G528 K532 P533 R536 R537 V538 N539 D540 G541 T542 I546 Q547 V553 P554 Y555 D556 R557 V565 V574 V582 K583 S584 E589 T593 I600 F603 G604 G605 F606 K607 P608 Y609 I610 Q614 VAL THR GLU ALA GLN LYS K621
K622 P623 F624 E625 T626 E634 I639 M640 K641 K649 D656 K659 Q662 L673 R676 T677 E680 I689 Q698 I701 P702 R716 Q719 D730 K734 L735 G736 L737 K738 H739 R742 E743 I744 V750 P751 L752 V753 P754 GLY SER GLY LEU
ILE LYS ASP ALA LYS GLU ALA LYS GLU VAL ALA LYS LYS LEU GLU TYR PRO VAL MET VAL LYS SER THR ALA GLY GLY GLY ILE GLY LEU GLY LYS VAL ASP SER GLU ASP ASP ILE GLU ARG VAL PHE GLU THR VAL GLN HIS GLN GLY LYS SER TYR PHE GLY ASP ALA GLY VAL
PHE MET GLU ARG PHE VAL ASN N826 H829 M834 M835 G840 K841 I845 G846 E847 C850 S851 L852 Q853 K858 V859 I860 E861 E862 P869 T872 R873 A874 K875 A879 K888 Y889 K890 C891 L915 H919 P920 V931 E932 W933 M934 L935 D941 D944
F945 E951 V952 I957 K968 T978 S979 V988 K993 S999 D1003 K1008 V1011 R1016 L1023 L1042 R1043 S1044 K1049 M1050 F1051 K1055 V1056 A1057 T1058 K1059 V1060 K1067 P1068 C1069 L1074 A1075 A1078 N1079 Y1085 P1086 V1096 M1102 D1103
S1104 R1108 G1115 P1121 A1122 L1123 E1124 I1125 T1126 L1127 K1131 T1138 V1139 I1140 A1141 V1142 T1146 V1147 S1148 C1149 T1150 L1151 N1152 D1153 I1156 K1165 R1166 S1171 V1172 Y1182 L1183 S1184 I1185 I1189 D1190 V1191 Y1194 T1200 M1203 M1206 G1207 G1208 L1216
L1220 Q1224 L1227 S1228 F1237 P1243 L1246 T1252 I1258 G1259 V1260 H1265 I1268 D1269 L1270 F1271 E1274 E1277 Q1278 F1279 Y1288 R1292 P1302 K1303 W1304 E1311 A1312 G1313 L1314 H1315 P1316 S1317 N1318 Y1322 V1323 Y1324 S1325 L1326 F1331 T1332 I1339
P1344 S1345 L1346 G1347 V1355 W1361 V1375 P1376 L1377 S1388 K1391 S1392 I1393 F1396 L1402 K1408 D1426 L1427 S1428 P1429 K1430 Y1433 G1437 D1438 L1452 C1455 Y1456 R1457 I1458 N1459 R1460 Q1464 H1468 Q1469 T1470 I1473 S1477 Q1478 L1498 K1499 C1500 L1501
I1502 A1503 Y1504 Q1509 V1516 L1523 P1524 M1525 A1526 F1527 C1535 V1536 T1537 R1538 Y1539 R1540 E1541 T1542 P1548 W1549 L1550 P1551 F1556 I1557 A1558 D1559 Y1560 N1561 D1562 V1568 M1571 L1572 Y1573 V1585 F1586 L1587 V1593 Y1601 K1605 Y1606 N1607 R1610 T1611 Y1612 G1616 V1617
V1618 G1619 I1620 M1625 C1626 I1627 G1632 S1633 P1634 L1639 V1640 G1641 R1642 T1643 W1647 D1654 L1662 T1675 F1691 K1692 I1693 D1694 V1695 E1696 E1697 W1708 I1709 N1710 E1711 N1712 I1713 V1717 E1721 E1726 E1730 F1731 A1732 I1735 Q1736 V1737 ALA ASN ALA GLU LEU LYS
LYS SER GLY ASP ASP LYS PRO GLN ASP VAL GLU GLU TYR PRO ASP ASP ASP ALA GLU LEU LEU TYR SER GLU TYR THR GLY ARG PHE TRP LYS PRO VAL ALA ALA VAL GLY ASP HIS VAL GLU ALA GLY ASP GLY VAL ILE ILE ILE GLU ALA MET LYS THR GLU MET VAL VAL GLY ALA THR
LYS SER GLY LYS VAL TYR LYS ILE LEU HIS LYS ASN GLY ASP MET VAL GLU ALA GLY ASP LEU VAL ALA VAL ILE VAL

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	105.74Å 181.94Å 549.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.98 – 6.50 29.98 – 6.50	Depositor EDS
% Data completeness (in resolution range)	95.8 (29.98-6.50) 94.7 (29.98-6.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	2.47 (at 6.58Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.278 , 0.302 0.274 , 0.301	Depositor DCC
R_{free} test set	1017 reflections (2.08%)	wwPDB-VP
Wilson B-factor (Å^2)	216.1	Xtriage
Anisotropy	0.897	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 331.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.78	EDS
Total number of atoms	51722	wwPDB-VP
Average B, all atoms (Å^2)	308.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% 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	0.53	8/13231 (0.1%)	1.02	44/17964 (0.2%)
1	B	0.66	10/13214 (0.1%)	1.19	100/17942 (0.6%)
1	C	0.51	0/13231	1.01	37/17964 (0.2%)
1	D	0.63	4/13214 (0.0%)	1.16	68/17942 (0.4%)
All	All	0.59	22/52890 (0.0%)	1.10	249/71812 (0.3%)

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	A	0	2
1	B	0	2
1	C	0	3
1	D	0	6
All	All	0	13

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	63	LYS	CA-C	8.10	1.63	1.52
1	B	63	LYS	CA-C	7.89	1.63	1.52
1	B	61	GLN	CA-C	-7.22	1.42	1.52
1	A	64	SER	N-CA	7.16	1.54	1.46
1	B	64	SER	N-CA	7.12	1.54	1.46

The worst 5 of 249 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	519	TYR	N-CA-C	10.51	126.42	110.14
1	B	305	TYR	N-CA-C	10.26	122.22	111.14
1	D	305	TYR	N-CA-C	9.69	121.60	111.14

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1303	LYS	N-CA-C	-9.66	93.52	109.07
1	D	556	ASP	N-CA-CB	-9.62	94.41	110.39

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1469	GLN	Mainchain
1	A	539	ASN	Peptide
1	B	1469	GLN	Mainchain
1	B	539	ASN	Peptide
1	C	539	ASN	Peptide

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	12939	0	12896	234	0
1	B	12922	0	12878	228	1
1	C	12939	0	12896	214	1
1	D	12922	0	12878	226	0
All	All	51722	0	51548	853	1

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 8.

The worst 5 of 853 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:59:ILE:O	1:B:62:SER:OG	1.78	1.01
1:B:79:LYS:HA	1:B:121:THR:HG22	1.45	0.97
1:C:1509:GLN:NE2	1:D:101:THR:O	1.97	0.97
1:D:79:LYS:HA	1:D:121:THR:HG22	1.45	0.95
1:A:493:LEU:HB2	1:A:496:LEU:HD12	1.47	0.94

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1288:TYR:OH	1:C:1541:GLU:OE2[2_455]	1.97	0.23

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	1654/1829 (90%)	1540 (93%)	101 (6%)	13 (1%)	16	54
1	B	1652/1829 (90%)	1544 (94%)	97 (6%)	11 (1%)	18	56
1	C	1654/1829 (90%)	1539 (93%)	102 (6%)	13 (1%)	16	54
1	D	1652/1829 (90%)	1544 (94%)	97 (6%)	11 (1%)	18	56
All	All	6612/7316 (90%)	6167 (93%)	397 (6%)	48 (1%)	18	56

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	38	ALA
1	A	1166	ARG
1	A	1711	GLU
1	B	1166	ARG
1	C	38	ALA

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	1413/1547 (91%)	1263 (89%)	150 (11%)	6	21
1	B	1411/1547 (91%)	1248 (88%)	163 (12%)	5	18
1	C	1413/1547 (91%)	1262 (89%)	151 (11%)	6	21
1	D	1411/1547 (91%)	1250 (89%)	161 (11%)	5	18
All	All	5648/6188 (91%)	5023 (89%)	625 (11%)	6	20

5 of 625 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1736	GLN
1	D	1055	LYS
1	D	116	VAL
1	C	1726	GLU
1	D	526	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1509	GLN
1	D	594	GLN
1	C	1659	HIS
1	D	230	ASN
1	D	1315	HIS

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 oligosaccharides 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	1660/1829 (90%)	0.14	68 (4%)	41	41	236, 291, 352, 387	0
1	B	1658/1829 (90%)	-0.00	44 (2%)	56	48	245, 295, 361, 403	0
1	C	1660/1829 (90%)	-0.02	34 (2%)	65	56	269, 323, 368, 413	0
1	D	1658/1829 (90%)	0.08	43 (2%)	57	49	254, 306, 345, 377	0
All	All	6636/7316 (90%)	0.05	189 (2%)	55	48	236, 307, 356, 413	0

The worst 5 of 189 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1568	VAL	10.5
1	D	1568	VAL	7.6
1	A	1557	ILE	6.9
1	B	1548	PRO	6.8
1	A	1587	LEU	5.8

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.