wwPDB X-ray Structure Validation Summary Report (i)

Dec 3, 2023 – 04:22 am GMT

PDB ID	:	2UZY
Title	:	Structure of the human receptor tyrosine kinase Met in complex with the Listeria monocytogenes invasion protein inlb: low resolution, Crystal form II
Authors	:	Niemann, H.H.; Jager, V.; Butler, P.J.G.; van den Heuvel, J.; Schmidt, S.; Ferraris, D.; Gherardi, E.; Heinz, D.W.
Deposited on	:	2007-05-02
Resolution	:	4.00 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 (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

PERCENTILES INFOmissingINFO

1 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 14329 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 INTERNALIN B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	286	Total	C	N	O	S	0	0	0
			2252	1435	378	437	2			
1	C	286	Total	C	N	O	S	0	0	0
			2252	1435	378	437	2			

- Molecule 2 is a protein called HEPATOCYTE GROWTH FACTOR RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	630	Total	C	N	O	S	0	0	0
			4959	3151	835	938	35			
2	D	618	Total	C	N	O	S	0	0	0
			4866	3097	818	918	33			

There are 4 discrepancies between the modelled and reference sequences:

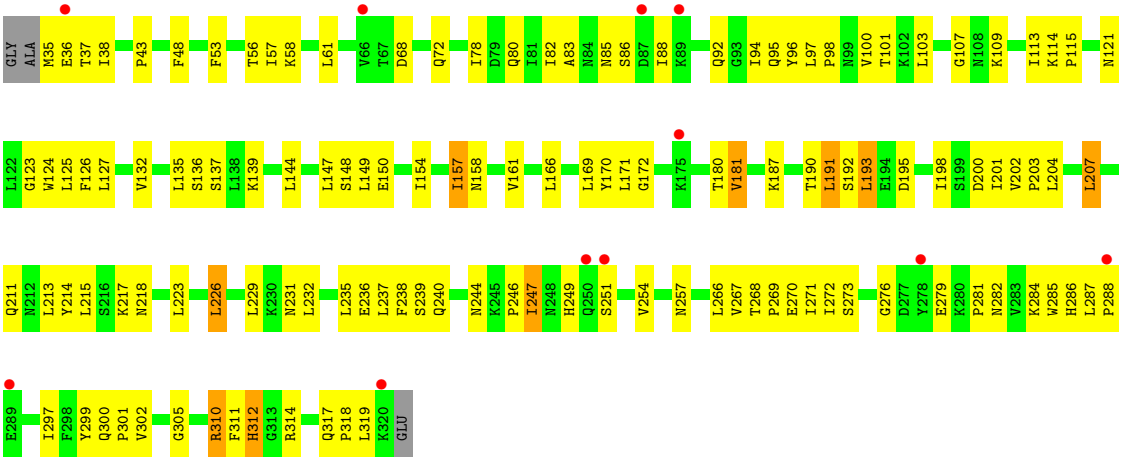
Chain	Residue	Modelled	Actual	Comment	Reference
B	41	CYS	TYR	conflict	UNP P08581
B	344	ALA	GLY	conflict	UNP P08581
D	41	CYS	TYR	conflict	UNP P08581
D	344	ALA	GLY	conflict	UNP P08581

2 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.

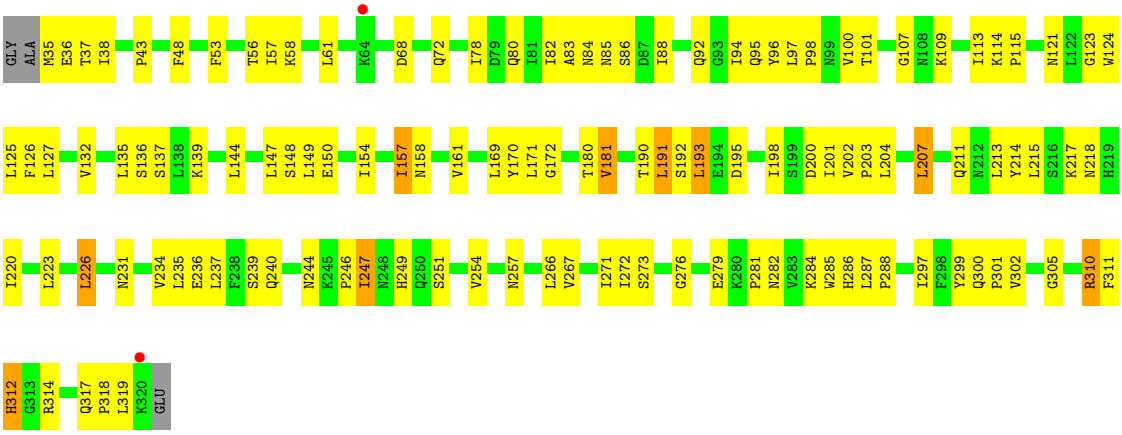
• Molecule 1: INTERNALIN B

Chain A:



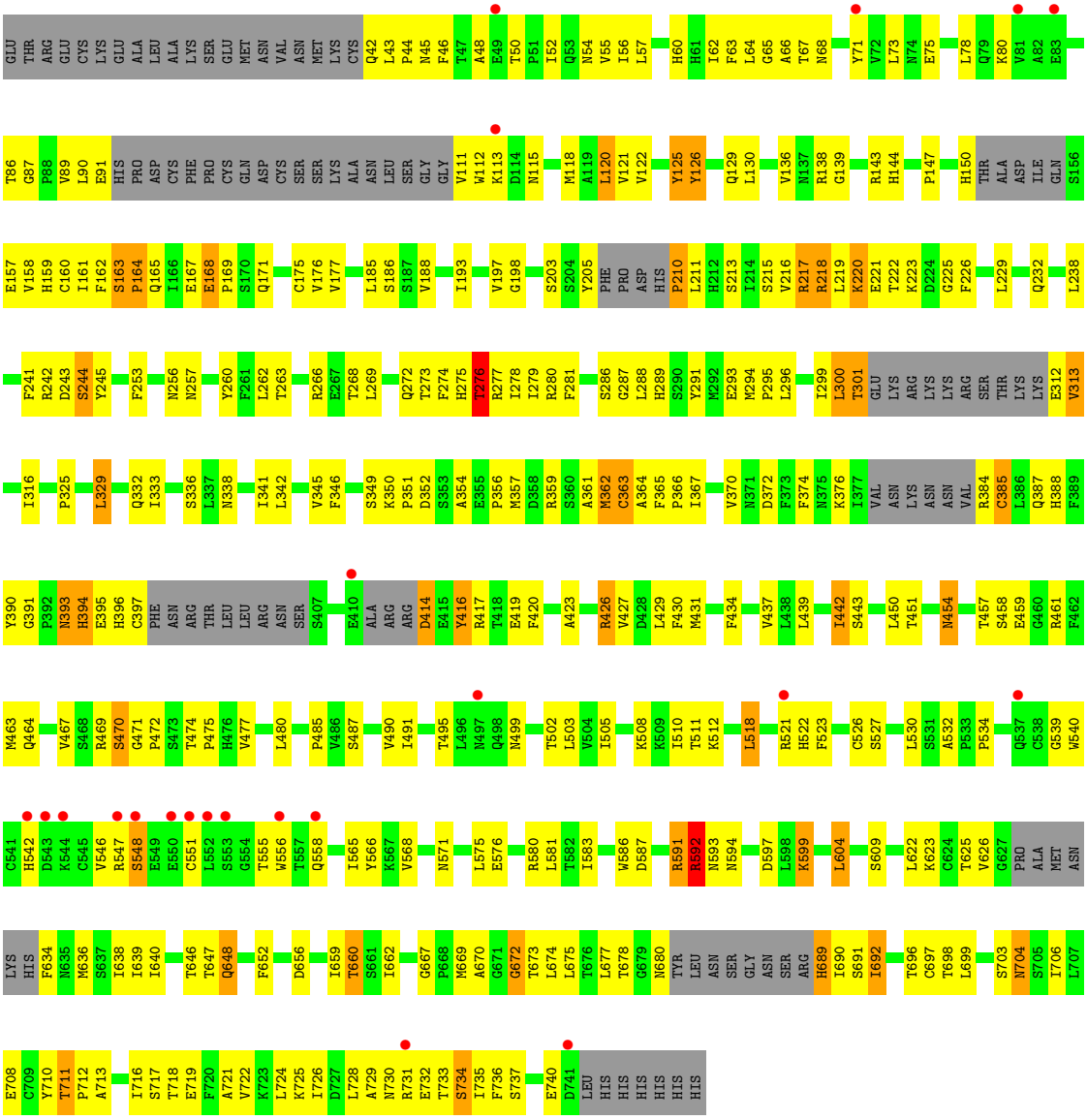
• Molecule 1: INTERNALIN B

Chain C:



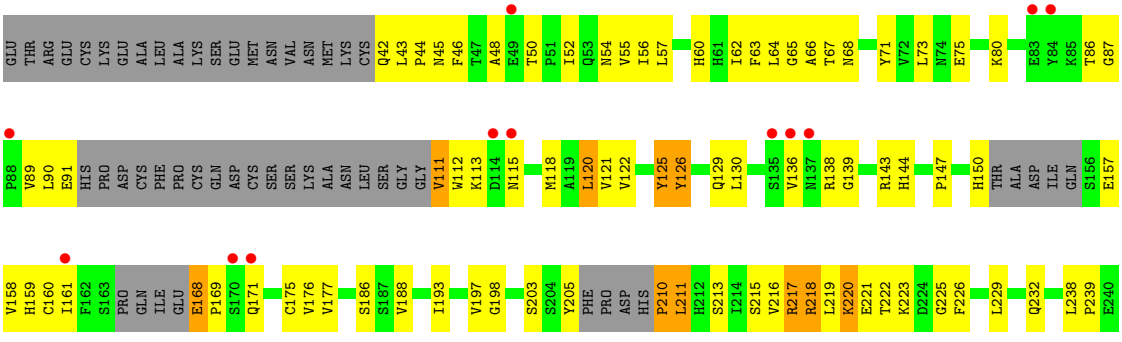
• Molecule 2: HEPATOCYTE GROWTH FACTOR RECEPTOR

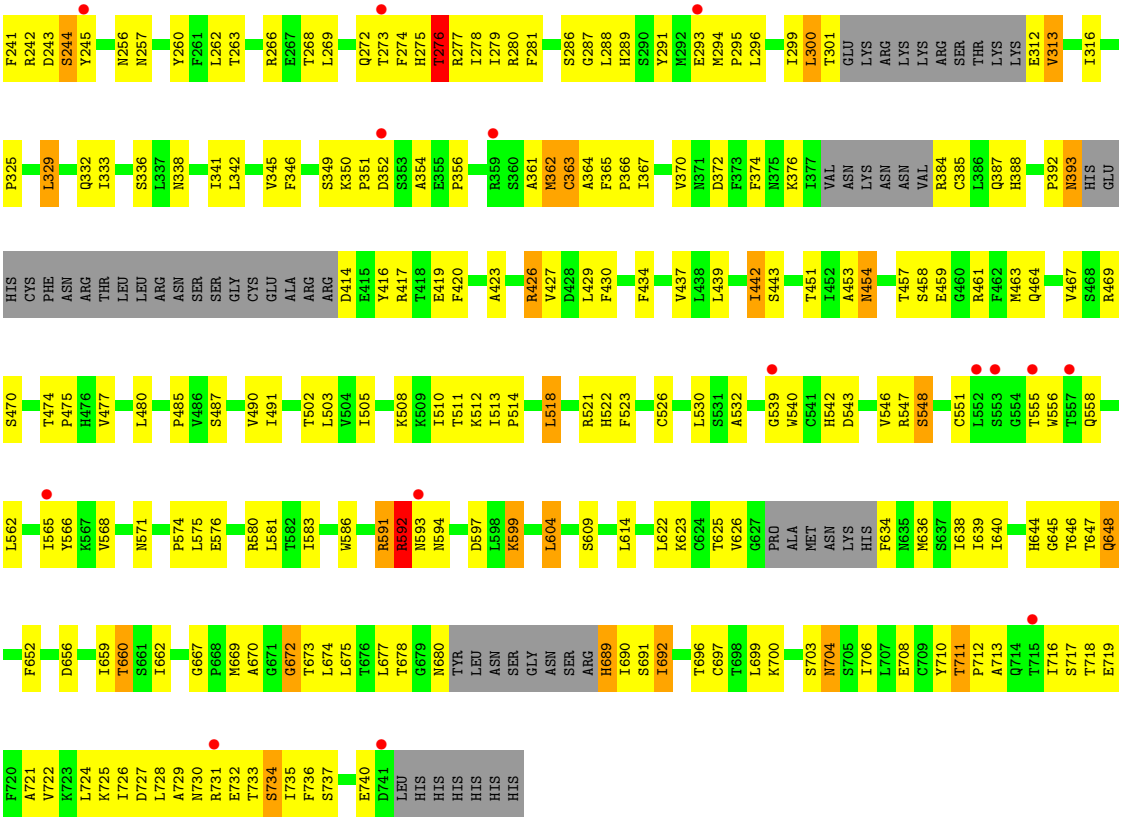
Chain B:



• Molecule 2: HEPATOCYTE GROWTH FACTOR RECEPTOR

Chain D:




F241 R242 D243 S244 Y245 N256 N257 Y260 F261 L262 T263 R266 E267 T268 L269 Q272 T273 F274 H275 T276 R277 I278 I279 R280 F281 S286 G287 L288 H289 S290 Y291 M292 E293 M294 P295 L296 I299 L300 T301 GLU LYS ARG LYS LYS ARG SER THR LYS LYS E312 V313 I316
P325 L329 Q332 I333 S336 L337 N338 I341 L342 V345 F346 S349 K350 P351 D352 S353 A354 E355 P356 H359 S360 A361 M362 C363 A364 F365 P366 I367 V370 N371 D372 F373 F374 M375 K376 I377 VAL ASN LYS ASN ASN VAL R384 C385 L386 Q387 H388 P392 N393 HIS GLU
HIS CYS PHE ASN ARG THR LEU LEU ARG ASN SER SER GLY CYS GLU ALA ARG ARG D414 E415 Y416 R417 T418 E419 F420 A423 R426 V427 D428 L429 F430 F434 V437 L438 L439 I442 S443 T451 I452 A453 N454 T457 S458 E459 G460 R461 F462 M463 Q464 V467 S468 R469
S470 T474 P475 H476 V477 L480 P485 V486 S487 V490 I491 T502 L503 V504 I505 K508 K509 I510 T511 K512 I513 P514 L518 R521 H522 F523 C526 L530 S531 A532 G539 W540 C541 H542 D543 V546 R547 S548 C551 L552 S553 G554 T555 W556 T557 Q558
L562 I565 Y566 K567 V568 N571 P574 L575 E576 R580 L581 T582 I583 W586 R591 R592 N593 N594 D597 L598 K599 L604 S609 L614 L622 K623 C624 T625 V626 G627 PRO ALA MET ASN LYS HIS F634 M635 M636 S637 I638 I639 I640 H644 G645 T646 T647 Q648
F652 D656 I659 T660 S661 I662 G667 F668 M669 A670 G671 G672 T673 L674 L675 T676 L677 T678 G679 N680 TYR LEU ASN SER GLY ASN SER ARG H689 I690 S691 I692 T696 C697 T698 L699 K700 S703 N704 S705 I706 L707 E708 C709 Y710 T711 P712 A713 Q714 T715 I716 S717 T718 E719
F720 A721 V722 K723 L724 K725 I726 D727 L728 A729 N730 R731 E732 T733 S734 I735 F736 S737 E740 D741 LEU HIS HIS HIS HIS HIS HIS

3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	139.00Å 144.98Å 150.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 4.00 48.16 – 4.00	Depositor EDS
% Data completeness (in resolution range)	97.5 (15.00-4.00) 99.6 (48.16-4.00)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.14 (at 4.00Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.251 , 0.301 0.257 , 0.294	Depositor DCC
R_{free} test set	1302 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	119.6	Xtriage
Anisotropy	0.833	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 92.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.21$	Xtriage
Estimated twinning fraction	0.088 for -h,l,k	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	14329	wwPDB-VP
Average B, all atoms (Å^2)	136.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.29% 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.

4 Model quality

4.1 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	0.23	0/2288	0.42	0/3105
1	C	0.24	0/2288	0.42	0/3105
2	B	0.33	1/5068 (0.0%)	0.47	1/6868 (0.0%)
2	D	0.33	2/4972 (0.0%)	0.48	0/6738
All	All	0.30	3/14616 (0.0%)	0.46	1/19816 (0.0%)

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
2	B	0	1
2	D	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	113	LYS	C-N	12.95	1.63	1.34
2	D	113	LYS	C-N	9.64	1.56	1.34
2	D	111	VAL	CB-CG2	-5.24	1.41	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	414	ASP	N-CA-C	-6.07	94.61	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	300	LEU	Peptide
2	D	300	LEU	Peptide

4.2 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	2252	0	2323	125	0
1	C	2252	0	2323	111	0
2	B	4959	0	4832	301	0
2	D	4866	0	4760	270	0
All	All	14329	0	14238	781	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 27.

The worst 5 of 781 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:310:ARG:HH11	1:A:310:ARG:HG3	1.32	0.94
1:C:310:ARG:HG3	1:C:310:ARG:HH11	1.33	0.92
1:A:193:LEU:HD22	1:A:198:ILE:HD11	1.53	0.90
1:C:193:LEU:HD22	1:C:198:ILE:HD11	1.52	0.90
2:B:521:ARG:HE	2:B:546:VAL:HG23	1.37	0.90

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	284/289 (98%)	226 (80%)	54 (19%)	4 (1%)	11	46
1	C	284/289 (98%)	227 (80%)	53 (19%)	4 (1%)	11	46
2	B	610/727 (84%)	494 (81%)	101 (17%)	15 (2%)	5	35
2	D	598/727 (82%)	489 (82%)	96 (16%)	13 (2%)	6	37
All	All	1776/2032 (87%)	1436 (81%)	304 (17%)	36 (2%)	7	40

5 of 36 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	193	LEU
2	B	163	SER
2	B	164	PRO
2	B	211	LEU
2	B	276	THR

4.3.2 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	264/265 (100%)	251 (95%)	13 (5%)	25	52
1	C	264/265 (100%)	251 (95%)	13 (5%)	25	52
2	B	566/654 (86%)	524 (93%)	42 (7%)	13	41
2	D	555/654 (85%)	514 (93%)	41 (7%)	13	41
All	All	1649/1838 (90%)	1540 (93%)	109 (7%)	16	45

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	135	LEU
2	D	150	HIS
2	D	592	ARG
1	C	180	THR
1	C	302	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	76	ASN
2	D	42	GLN
2	D	571	ASN
1	C	80	GLN
1	C	211	GLN

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

There are no ligands in this entry.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	113:LYS	C	114:ASP	N	1.63

5 Fit of model and data

5.1 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	286/289 (98%)	0.11	11 (3%)	40	32	89, 125, 174, 201	0
1	C	286/289 (98%)	0.02	2 (0%)	87	82	76, 112, 158, 191	0
2	B	630/727 (86%)	0.27	22 (3%)	44	35	87, 139, 208, 245	0
2	D	618/727 (85%)	0.34	27 (4%)	34	28	87, 137, 191, 236	0
All	All	1820/2032 (89%)	0.23	62 (3%)	45	36	76, 132, 192, 245	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	136	VAL	5.9
2	B	552	LEU	5.1
2	D	555	THR	4.9
2	D	135	SER	4.3
2	B	544	LYS	4.0

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

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

5.3 Carbohydrates

There are no monosaccharides in this entry.

5.4 Ligands

There are no ligands in this entry.

5.5 Other polymers (i)

There are no such residues in this entry.