



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 10:21 pm BST

PDB ID : 6E2Q
Title : Structure of human JAK2 FERM/SH2 in complex with Erythropoietin Receptor
Authors : Ferrao, R.; Lupardus, P.J.
Deposited on : 2018-07-11
Resolution : 2.65 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

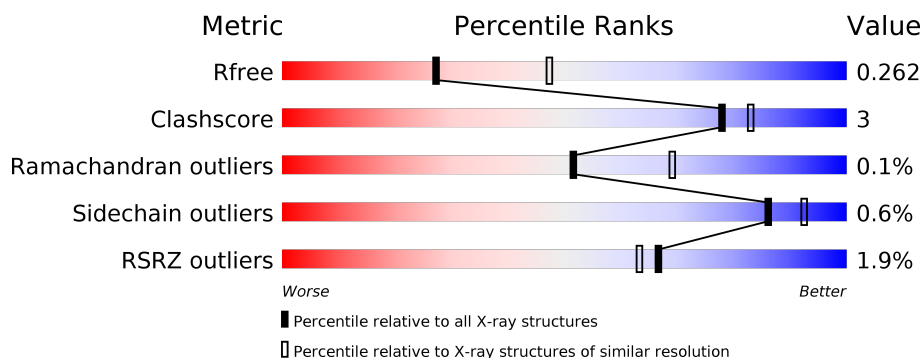
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.65 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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 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 $\leq 5\%$. 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	483	<div> <div>90%</div> <div>8%</div> <div>•</div> </div>
1	B	483	<div> <div>%</div> <div>90%</div> <div>7%</div> <div>•</div> </div>
1	C	483	<div> <div>%</div> <div>86%</div> <div>10%</div> <div>•</div> </div>
1	D	483	<div> <div>2%</div> <div>81%</div> <div>8%</div> <div>11%</div> </div>
2	M	83	<div> <div>4%</div> <div>45%</div> <div>7%</div> <div>•</div> <div>47%</div> </div>
2	N	83	<div> <div>5%</div> <div>64%</div> <div>•</div> <div>•</div> <div>31%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	O	83	<div><div><div>5%</div><div>59%</div><div>37%</div></div></div>
2	P	83	<div><div><div>5%</div><div>54%</div><div>43%</div></div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16599 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 Tyrosine-protein kinase JAK2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	473	Total	C	N	O	S	0	0	0
			3840	2457	660	702	21			
1	B	470	Total	C	N	O	S	0	0	0
			3809	2440	652	695	22			
1	C	463	Total	C	N	O	S	0	0	0
			3705	2378	630	676	21			
1	D	431	Total	C	N	O	S	0	0	0
			3493	2248	595	632	18			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	GLY	-	expression tag	UNP O60674
A	35	SER	-	expression tag	UNP O60674
A	515	GLY	-	expression tag	UNP O60674
A	516	SER	-	expression tag	UNP O60674
B	34	GLY	-	expression tag	UNP O60674
B	35	SER	-	expression tag	UNP O60674
B	515	GLY	-	expression tag	UNP O60674
B	516	SER	-	expression tag	UNP O60674
C	34	GLY	-	expression tag	UNP O60674
C	35	SER	-	expression tag	UNP O60674
C	515	GLY	-	expression tag	UNP O60674
C	516	SER	-	expression tag	UNP O60674
D	34	GLY	-	expression tag	UNP O60674
D	35	SER	-	expression tag	UNP O60674
D	515	GLY	-	expression tag	UNP O60674
D	516	SER	-	expression tag	UNP O60674

- Molecule 2 is a protein called Erythropoietin receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	M	44	Total	C	N	O	S	0	0	0
			362	239	57	64	2			
2	N	57	Total	C	N	O	S	0	0	0
			454	296	70	86	2			
2	O	52	Total	C	N	O	S	0	0	0
			417	273	64	79	1			
2	P	47	Total	C	N	O	S	0	0	0
			374	247	59	66	2			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	259	GLY	-	expression tag	UNP P19235
M	260	SER	-	expression tag	UNP P19235
M	261	GLY	-	expression tag	UNP P19235
M	262	SER	-	expression tag	UNP P19235
M	263	GLY	-	expression tag	UNP P19235
M	264	SER	-	expression tag	UNP P19235
M	265	GLY	-	expression tag	UNP P19235
M	266	SER	-	expression tag	UNP P19235
M	267	GLY	-	expression tag	UNP P19235
M	268	SER	-	expression tag	UNP P19235
M	269	GLY	-	expression tag	UNP P19235
M	270	SER	-	expression tag	UNP P19235
M	271	GLY	-	expression tag	UNP P19235
M	272	SER	-	expression tag	UNP P19235
M	339	GLY	-	expression tag	UNP P19235
M	340	ASN	-	expression tag	UNP P19235
M	341	SER	-	expression tag	UNP P19235
N	259	GLY	-	expression tag	UNP P19235
N	260	SER	-	expression tag	UNP P19235
N	261	GLY	-	expression tag	UNP P19235
N	262	SER	-	expression tag	UNP P19235
N	263	GLY	-	expression tag	UNP P19235
N	264	SER	-	expression tag	UNP P19235
N	265	GLY	-	expression tag	UNP P19235
N	266	SER	-	expression tag	UNP P19235
N	267	GLY	-	expression tag	UNP P19235
N	268	SER	-	expression tag	UNP P19235
N	269	GLY	-	expression tag	UNP P19235
N	270	SER	-	expression tag	UNP P19235
N	271	GLY	-	expression tag	UNP P19235
N	272	SER	-	expression tag	UNP P19235

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
N	339	GLY	-	expression tag	UNP P19235
N	340	ASN	-	expression tag	UNP P19235
N	341	SER	-	expression tag	UNP P19235
O	259	GLY	-	expression tag	UNP P19235
O	260	SER	-	expression tag	UNP P19235
O	261	GLY	-	expression tag	UNP P19235
O	262	SER	-	expression tag	UNP P19235
O	263	GLY	-	expression tag	UNP P19235
O	264	SER	-	expression tag	UNP P19235
O	265	GLY	-	expression tag	UNP P19235
O	266	SER	-	expression tag	UNP P19235
O	267	GLY	-	expression tag	UNP P19235
O	268	SER	-	expression tag	UNP P19235
O	269	GLY	-	expression tag	UNP P19235
O	270	SER	-	expression tag	UNP P19235
O	271	GLY	-	expression tag	UNP P19235
O	272	SER	-	expression tag	UNP P19235
O	339	GLY	-	expression tag	UNP P19235
O	340	ASN	-	expression tag	UNP P19235
O	341	SER	-	expression tag	UNP P19235
P	259	GLY	-	expression tag	UNP P19235
P	260	SER	-	expression tag	UNP P19235
P	261	GLY	-	expression tag	UNP P19235
P	262	SER	-	expression tag	UNP P19235
P	263	GLY	-	expression tag	UNP P19235
P	264	SER	-	expression tag	UNP P19235
P	265	GLY	-	expression tag	UNP P19235
P	266	SER	-	expression tag	UNP P19235
P	267	GLY	-	expression tag	UNP P19235
P	268	SER	-	expression tag	UNP P19235
P	269	GLY	-	expression tag	UNP P19235
P	270	SER	-	expression tag	UNP P19235
P	271	GLY	-	expression tag	UNP P19235
P	272	SER	-	expression tag	UNP P19235
P	339	GLY	-	expression tag	UNP P19235
P	340	ASN	-	expression tag	UNP P19235
P	341	SER	-	expression tag	UNP P19235

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	55	Total O 55 55	0	0

Continued on next page...

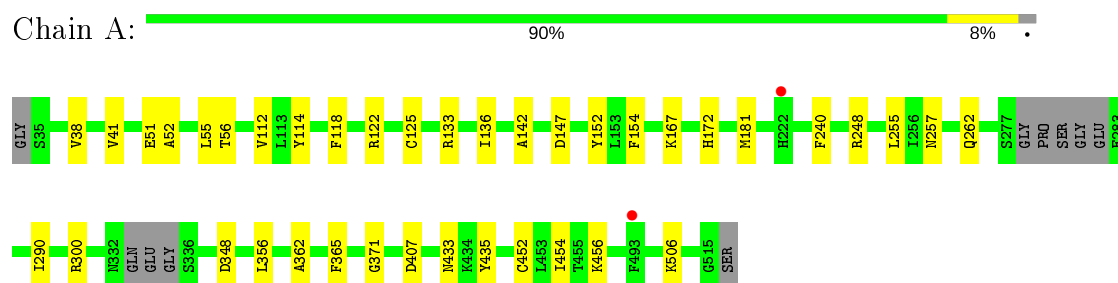
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	38	Total 38	O 38	0	0
3	C	31	Total 31	O 31	0	0
3	D	11	Total 11	O 11	0	0
3	M	3	Total 3	O 3	0	0
3	N	4	Total 4	O 4	0	0
3	O	2	Total 2	O 2	0	0
3	P	1	Total 1	O 1	0	0

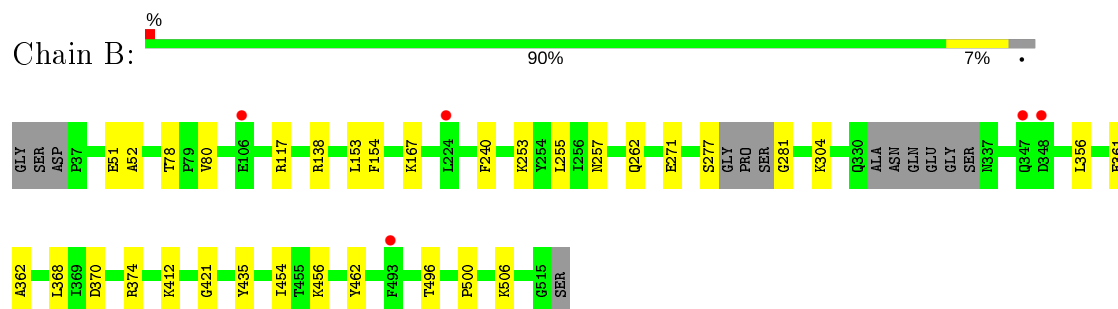
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

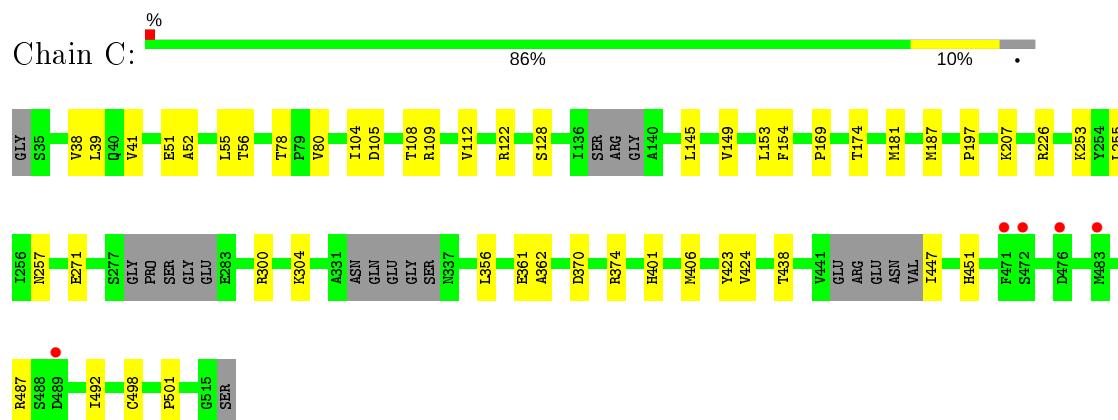
- Molecule 1: Tyrosine-protein kinase JAK2



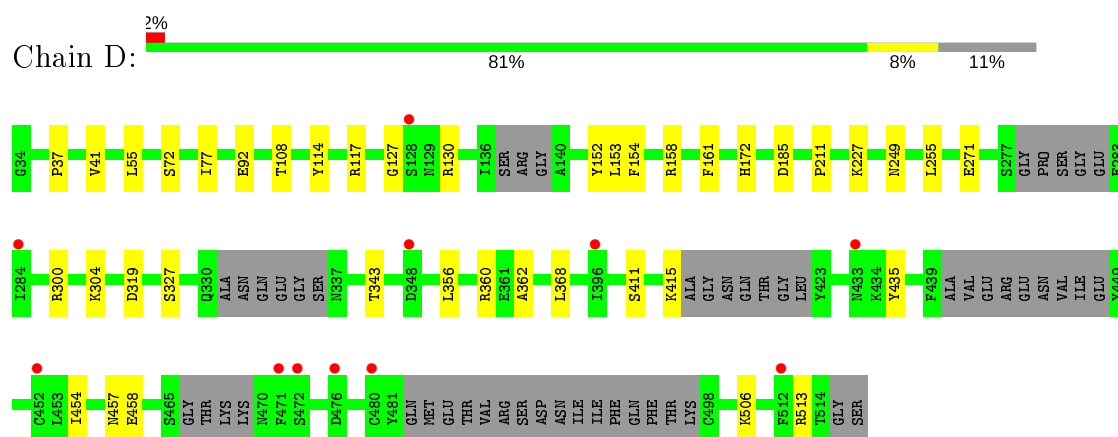
- Molecule 1: Tyrosine-protein kinase JAK2



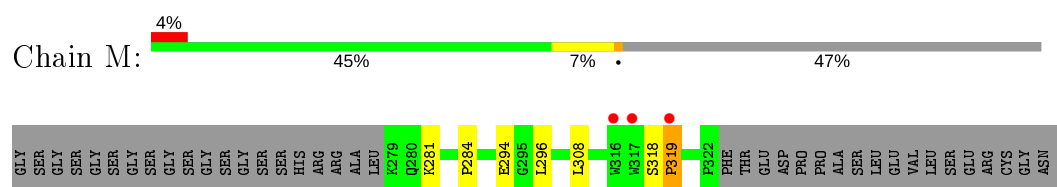
- Molecule 1: Tyrosine-protein kinase JAK2



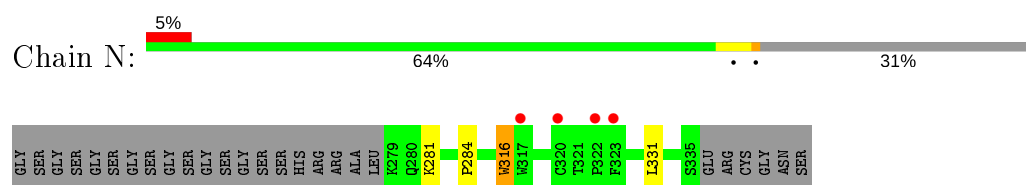
- Molecule 1: Tyrosine-protein kinase JAK2



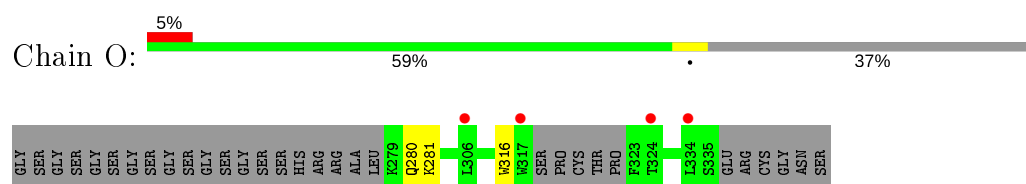
- Molecule 2: Erythropoietin receptor



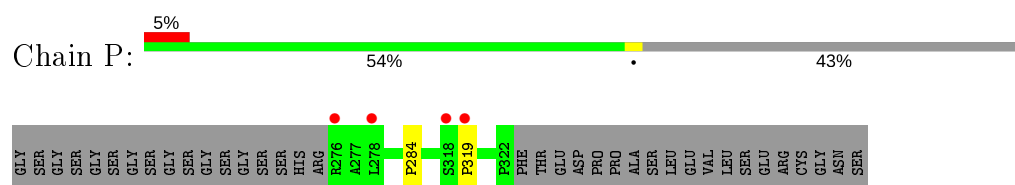
- Molecule 2: Erythropoietin receptor



- Molecule 2: Erythropoietin receptor



- Molecule 2: Erythropoietin receptor



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	178.49Å 114.88Å 179.82Å 90.00° 93.21° 90.00°	Depositor
Resolution (Å)	48.44 – 2.65 48.44 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.44-2.65) 92.1 (48.44-2.65)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.95 (at 2.65Å)	Xtriage
Refinement program	PHENIX (1.13 _2998: ???)	Depositor
R, R_{free}	0.225 , 0.263 0.225 , 0.262	Depositor DCC
R_{free} test set	1997 reflections (1.90%)	wwPDB-VP
Wilson B-factor (Å ²)	54.0	Xtriage
Anisotropy	0.519	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 45.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	16599	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.33% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.24	0/3933	0.39	0/5319
1	B	0.24	0/3902	0.39	0/5277
1	C	0.24	0/3795	0.39	0/5141
1	D	0.23	0/3578	0.38	0/4839
2	M	0.24	0/379	0.38	0/520
2	N	0.24	0/473	0.39	0/651
2	O	0.23	0/433	0.37	0/594
2	P	0.24	0/391	0.40	0/538
All	All	0.24	0/16884	0.39	0/22879

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	3840	0	3769	22	0
1	B	3809	0	3736	19	0
1	C	3705	0	3582	27	0
1	D	3493	0	3403	24	0
2	M	362	0	327	5	0
2	N	454	0	412	4	0
2	O	417	0	369	1	0
2	P	374	0	329	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	55	0	0	0	0
3	B	38	0	0	1	0
3	C	31	0	0	1	0
3	D	11	0	0	0	0
3	M	3	0	0	0	0
3	N	4	0	0	0	0
3	O	2	0	0	0	0
3	P	1	0	0	0	0
All	All	16599	0	15927	93	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.

The worst 5 of 93 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:300:ARG:NH2	2:M:281:LYS:O	2.25	0.69
1:A:122:ARG:NH1	1:A:125:CYS:SG	2.74	0.60
1:A:248:ARG:NH2	2:N:316:TRP:O	2.35	0.60
1:B:78:THR:HG22	1:B:80:VAL:H	1.68	0.59
1:C:300:ARG:NH2	2:N:281:LYS:O	2.36	0.57

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	467/483 (97%)	455 (97%)	12 (3%)	0	100	100
1	B	464/483 (96%)	446 (96%)	18 (4%)	0	100	100
1	C	453/483 (94%)	440 (97%)	13 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	415/483 (86%)	399 (96%)	16 (4%)	0	100	100
2	M	42/83 (51%)	39 (93%)	2 (5%)	1 (2%)	6	7
2	N	55/83 (66%)	51 (93%)	4 (7%)	0	100	100
2	O	48/83 (58%)	46 (96%)	2 (4%)	0	100	100
2	P	45/83 (54%)	41 (91%)	4 (9%)	0	100	100
All	All	1989/2264 (88%)	1917 (96%)	71 (4%)	1 (0%)	51	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	M	319	PRO

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	419/432 (97%)	416 (99%)	3 (1%)	84	91
1	B	415/432 (96%)	414 (100%)	1 (0%)	93	97
1	C	395/432 (91%)	392 (99%)	3 (1%)	81	89
1	D	377/432 (87%)	377 (100%)	0	100	100
2	M	39/69 (56%)	38 (97%)	1 (3%)	46	64
2	N	50/69 (72%)	49 (98%)	1 (2%)	55	73
2	O	44/69 (64%)	42 (96%)	2 (4%)	27	42
2	P	38/69 (55%)	38 (100%)	0	100	100
All	All	1777/2004 (89%)	1766 (99%)	11 (1%)	86	92

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	109	ARG
1	C	112	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	N	316	TRP
1	B	240	PHE
2	M	296	LEU

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

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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	473/483 (97%)	-0.09	2 (0%) 92 93	34, 57, 96, 118	0
1	B	470/483 (97%)	-0.06	5 (1%) 80 79	39, 64, 102, 125	0
1	C	463/483 (95%)	-0.02	5 (1%) 80 79	38, 68, 118, 159	0
1	D	431/483 (89%)	0.05	11 (2%) 56 52	45, 78, 121, 135	0
2	M	44/83 (53%)	0.17	3 (6%) 17 14	48, 64, 100, 129	0
2	N	57/83 (68%)	0.41	4 (7%) 16 13	53, 78, 155, 168	0
2	O	52/83 (62%)	0.33	4 (7%) 13 10	64, 81, 115, 140	0
2	P	47/83 (56%)	0.27	4 (8%) 10 8	53, 73, 127, 135	0
All	All	2037/2264 (89%)	0.00	38 (1%) 66 63	34, 68, 113, 168	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	P	278	LEU	5.0
1	D	128	SER	4.5
2	N	322	PRO	3.9
1	C	471	PHE	3.9
2	O	317	TRP	3.8

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 carbohydrates 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.