wwPDB X-ray Structure Validation Summary Report (i)

Oct 3, 2021 – 10:11 AM EDT

PDB ID	:	3II6
Title	:	Structure of human Xrcc4 in complex with the tandem BRCT domains of DNA LigaseIV.
Authors	:	Meesala, S.; Junop, M.
Deposited on	:	2009-07-31
Resolution	:	2.40 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 following versions of software and data (see references (1)) were used in the production of this report:

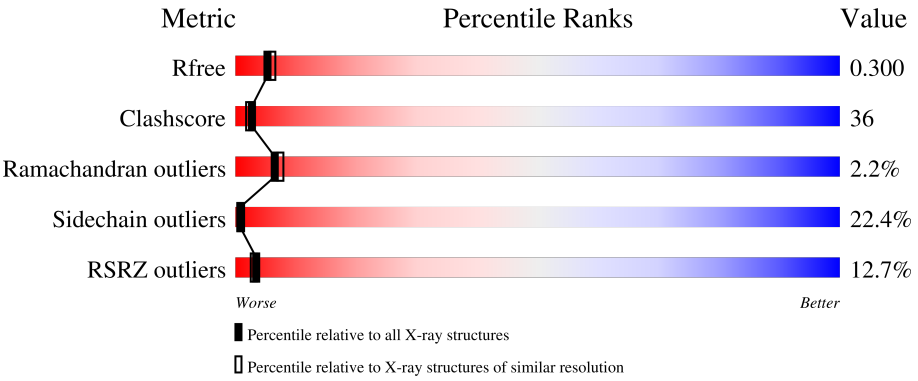
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.40 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.



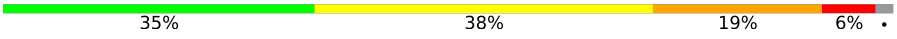
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	203	6% 39% 39% 15% 5%
1	B	203	34% 31% 41% 21%
1	C	203	4% 39% 37% 18%
1	D	203	34% 32% 40% 18% 6%
2	X	263	% 29% 33% 25% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	Y	263	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10925 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 DNA repair protein XRCC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	201	Total	C	N	O	S	0	0	0
			1631	1031	278	315	7			
1	B	196	Total	C	N	O	S	0	0	0
			1606	1016	273	310	7			
1	C	201	Total	C	N	O	S	0	0	0
			1631	1031	278	315	7			
1	D	195	Total	C	N	O	S	0	0	0
			1598	1012	272	307	7			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	60	GLU	ALA	engineered mutation	UNP Q13426
A	134	THR	ILE	engineered mutation	UNP Q13426
B	60	GLU	ALA	engineered mutation	UNP Q13426
B	134	THR	ILE	engineered mutation	UNP Q13426
C	60	GLU	ALA	engineered mutation	UNP Q13426
C	134	THR	ILE	engineered mutation	UNP Q13426
D	60	GLU	ALA	engineered mutation	UNP Q13426
D	134	THR	ILE	engineered mutation	UNP Q13426

- Molecule 2 is a protein called DNA ligase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	X	256	Total	C	N	O	S	0	0	0
			2081	1326	351	391	13			
2	Y	258	Total	C	N	O	S	0	0	0
			2095	1333	353	396	13			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	649	GLY	-	expression tag	UNP P49917
X	650	ALA	-	expression tag	UNP P49917
X	651	MET	-	expression tag	UNP P49917
X	652	GLY	-	expression tag	UNP P49917
X	653	SER	-	expression tag	UNP P49917
Y	649	GLY	-	expression tag	UNP P49917
Y	650	ALA	-	expression tag	UNP P49917
Y	651	MET	-	expression tag	UNP P49917
Y	652	GLY	-	expression tag	UNP P49917
Y	653	SER	-	expression tag	UNP P49917

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	Y	2	Total Cl 2 2	0	0

- Molecule 4 is water.

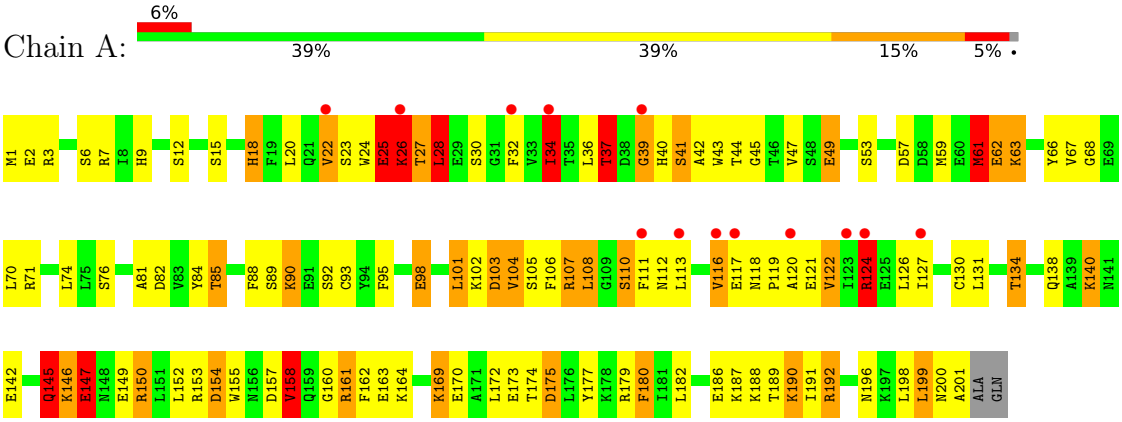
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	44	Total O 44 44	0	0
4	B	25	Total O 25 25	0	0
4	C	41	Total O 41 41	0	0
4	D	29	Total O 29 29	0	0
4	X	83	Total O 83 83	0	0
4	Y	59	Total O 59 59	0	0

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.

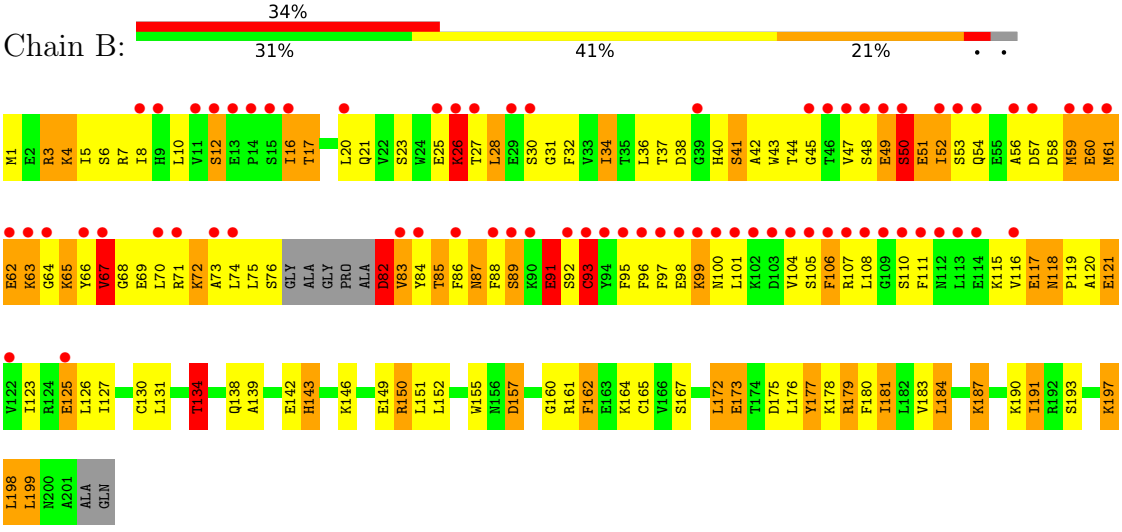
• Molecule 1: DNA repair protein XRCC4

Chain A:



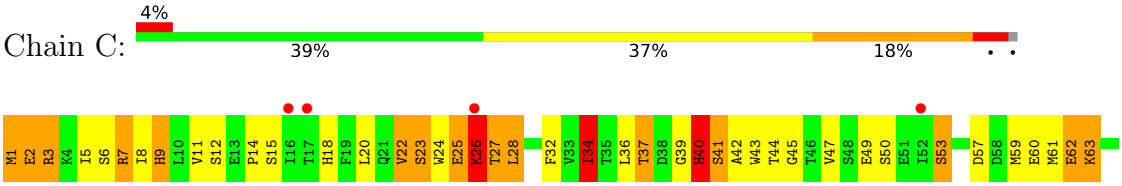
• Molecule 1: DNA repair protein XRCC4

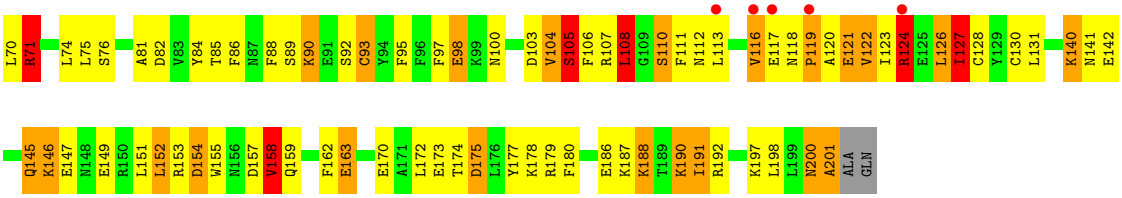
Chain B:



• Molecule 1: DNA repair protein XRCC4

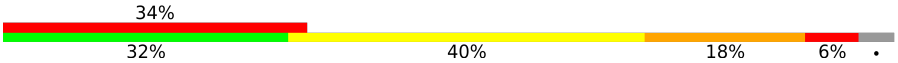
Chain C:

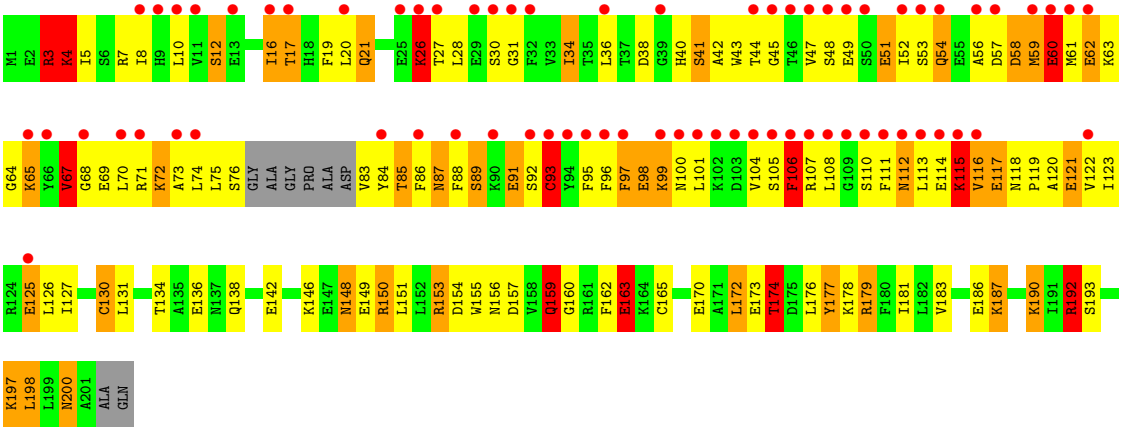




• Molecule 1: DNA repair protein XRCC4

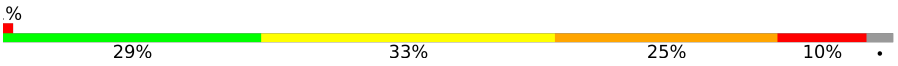
Chain D:

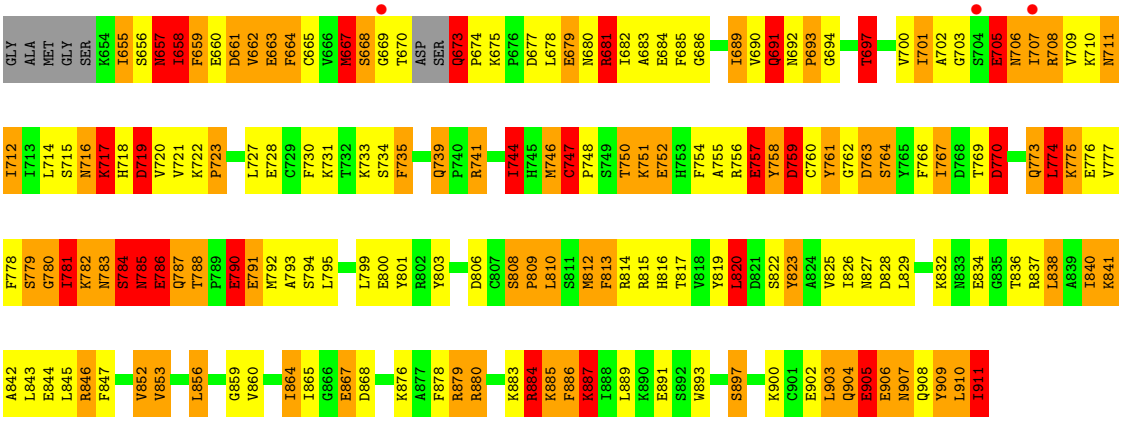




• Molecule 2: DNA ligase 4

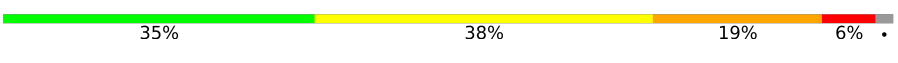
Chain X:

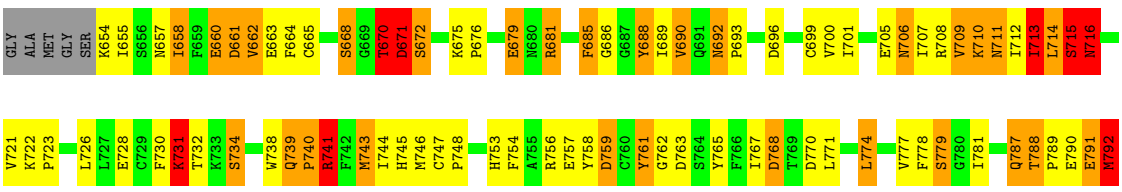


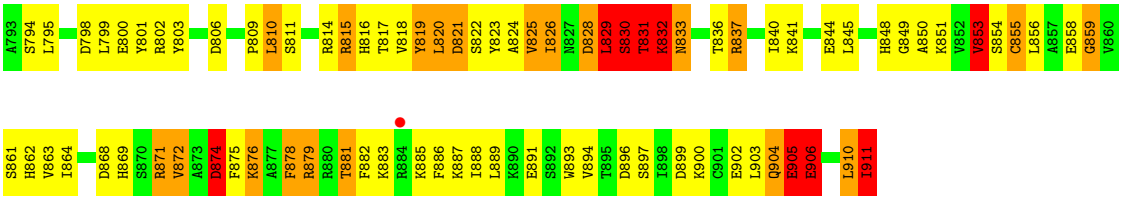


• Molecule 2: DNA ligase 4

Chain Y:






A793 S794 L795 D798 L799 E800 Y801 R802 Y803 D806 P809 L810 S811 R814 R815 H816 T817 V818 Y819 L820 D821 S822 Y823 A824 V825 I826 N827 D828 L829 S830 T831 K832 N833 T836 R837 I840 K841 E844 L845 H848 G849 A850 K851 V852 V853 S854 C855 L856 A857 E858 G859 V860
S861 H862 V863 I864 D868 H869 S870 R871 V872 A873 D874 F875 K876 A877 F878 R879 R880 T881 F882 K883 R884 K885 F886 K887 I888 L889 K890 E891 S892 W893 V894 T895 D896 S897 I898 D899 K900 C901 E902 L903 Q904 E905 E906 L910 I911

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	67.24Å 85.98Å 111.61Å 67.34° 82.86° 74.52°	Depositor
Resolution (Å)	20.00 – 2.40 45.26 – 2.28	Depositor EDS
% Data completeness (in resolution range)	86.9 (20.00-2.40) 95.4 (45.26-2.28)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.31 (at 2.29Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.240 , 0.280 0.254 , 0.300	Depositor DCC
R_{free} test set	3599 reflections (3.72%)	wwPDB-VP
Wilson B-factor (Å^2)	48.0	Xtriage
Anisotropy	0.118	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 52.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,-k+l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	10925	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.59% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: CL

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	1.93	44/1660 (2.7%)	1.68	38/2232 (1.7%)
1	B	1.81	36/1633 (2.2%)	1.41	23/2193 (1.0%)
1	C	1.91	44/1660 (2.7%)	1.63	39/2232 (1.7%)
1	D	1.84	29/1625 (1.8%)	1.54	27/2182 (1.2%)
2	X	2.20	84/2129 (3.9%)	1.95	73/2873 (2.5%)
2	Y	2.01	55/2144 (2.6%)	1.78	54/2895 (1.9%)
All	All	1.97	292/10851 (2.7%)	1.69	254/14607 (1.7%)

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	3
1	B	0	2
1	C	0	3
1	D	0	1
2	X	1	14
2	Y	0	12
All	All	1	35

The worst 5 of 292 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	91	GLU	CD-OE1	20.03	1.47	1.25
1	B	91	GLU	CD-OE1	15.02	1.42	1.25
2	X	909	TYR	CD1-CE1	13.72	1.59	1.39
1	A	179	ARG	CZ-NH1	-13.64	1.15	1.33
1	D	98	GLU	CD-OE1	13.47	1.40	1.25

The worst 5 of 254 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	179	ARG	NE-CZ-NH2	17.97	129.28	120.30
1	D	179	ARG	NE-CZ-NH1	-16.36	112.12	120.30
1	C	147	GLU	OE1-CD-OE2	-13.46	107.15	123.30
1	A	154	ASP	CB-CG-OD1	13.08	130.07	118.30
2	X	837	ARG	NE-CZ-NH2	-13.07	113.77	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	X	911	ILE	CA

5 of 35 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	25	GLU	Peptide
1	A	26	LYS	Peptide
1	A	39	GLY	Peptide
1	B	82	ASP	Peptide
1	B	83	VAL	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	1631	0	1617	89	0
1	B	1606	0	1593	147	0
1	C	1631	0	1617	94	0
1	D	1598	0	1589	140	0
2	X	2081	0	2036	192	0
2	Y	2095	0	2046	146	0
3	Y	2	0	0	0	0
4	A	44	0	0	3	0
4	B	25	0	0	2	0
4	C	41	0	0	4	0
4	D	29	0	0	5	0
4	X	83	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	Y	59	0	0	3	0
All	All	10925	0	10498	756	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 36.

The worst 5 of 756 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)
2:X:846:ARG:CD	2:X:846:ARG:CG	1.75	1.61
1:B:82:ASP:N	1:B:83:VAL:HG23	1.09	1.36
1:A:127:ILE:HD11	1:B:127:ILE:CD1	1.65	1.27
1:B:82:ASP:N	1:B:83:VAL:CG2	1.95	1.26
1:C:40:HIS:CD2	1:D:120:ALA:HB2	1.73	1.22

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	199/203 (98%)	170 (85%)	26 (13%)	3 (2%)	10 14
1	B	192/203 (95%)	168 (88%)	19 (10%)	5 (3%)	5 5
1	C	199/203 (98%)	172 (86%)	22 (11%)	5 (2%)	5 6
1	D	191/203 (94%)	167 (87%)	19 (10%)	5 (3%)	5 5
2	X	252/263 (96%)	217 (86%)	32 (13%)	3 (1%)	13 19
2	Y	256/263 (97%)	221 (86%)	27 (10%)	8 (3%)	4 3
All	All	1289/1338 (96%)	1115 (86%)	145 (11%)	29 (2%)	6 7

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	50	SER
2	X	747	CYS
2	X	786	GLU
2	Y	829	LEU
1	A	34	ILE

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	181/182 (100%)	142 (78%)	39 (22%)	1	1
1	B	180/182 (99%)	134 (74%)	46 (26%)	0	0
1	C	181/182 (100%)	144 (80%)	37 (20%)	1	1
1	D	179/182 (98%)	135 (75%)	44 (25%)	0	0
2	X	232/236 (98%)	175 (75%)	57 (25%)	0	0
2	Y	234/236 (99%)	191 (82%)	43 (18%)	1	2
All	All	1187/1200 (99%)	921 (78%)	266 (22%)	1	1

5 of 266 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	Y	711	ASN
2	Y	774	LEU
2	Y	878	PHE
1	C	49	GLU
1	C	28	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	Y	745	HIS
2	Y	787	GLN
2	Y	908	GLN
1	D	100	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	87	ASN

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

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	201/203 (99%)	0.52	13 (6%) 18 17	35, 71, 104, 125	0
1	B	196/203 (96%)	1.75	70 (35%) 0 0	33, 99, 134, 147	0
1	C	201/203 (99%)	0.34	9 (4%) 33 31	32, 72, 103, 132	0
1	D	195/203 (96%)	1.76	70 (35%) 0 0	29, 94, 134, 148	0
2	X	256/263 (97%)	-0.01	3 (1%) 79 77	31, 50, 83, 96	0
2	Y	258/263 (98%)	-0.01	1 (0%) 92 91	33, 49, 79, 112	0
All	All	1307/1338 (97%)	0.65	166 (12%) 3 3	29, 63, 125, 148	0

The worst 5 of 166 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	56	ALA	11.7
1	B	56	ALA	11.4
1	D	103	ASP	11.1
1	D	106	PHE	9.2
1	D	74	LEU	8.7

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 monosaccharides in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	CL	Y	302	1/1	0.87	0.26	65,65,65,65	0
3	CL	Y	301	1/1	0.96	0.07	65,65,65,65	0

6.5 Other polymers

There are no such residues in this entry.