Full wwPDB X-ray Structure Validation Report ⓘ

Dec 17, 2023 – 12:13 am GMT

PDB ID	:	4C0D
Title	:	Structure of the NOT module of the human CCR4-NOT complex (CNOT1-CNOT2-CNOT3)
Authors	:	Raisch, T.; Jonas, S.; Boland, A.; Chen, Y.; Izaurralde, E.; Weichenrieder, O.
Deposited on	:	2013-08-01
Resolution	:	3.20 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:

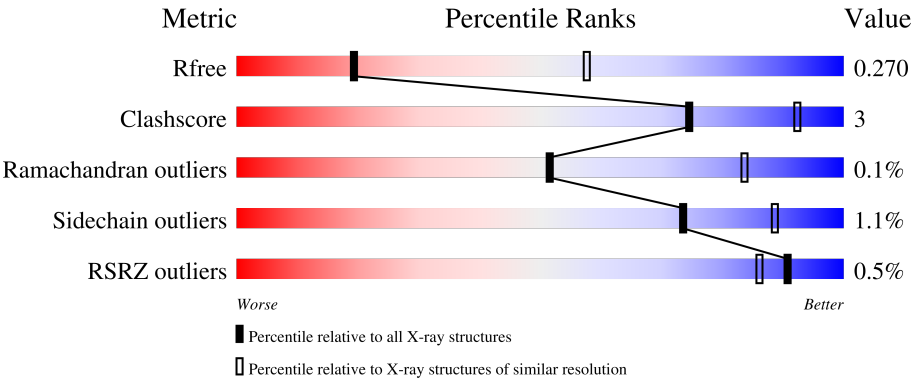
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	812	58% 5% 37%
2	B	201	83% 12% 5%
3	C	166	% 76% 10% 14%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6864 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 CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	512	Total	C	N	O	S	0	0	0
			4096	2653	702	720	21			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1560	GLY	-	expression tag	UNP A5YKK6
A	1561	HIS	-	expression tag	UNP A5YKK6
A	1562	MET	-	expression tag	UNP A5YKK6
A	1563	LEU	-	expression tag	UNP A5YKK6
A	1564	GLU	-	expression tag	UNP A5YKK6

- Molecule 2 is a protein called CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	191	Total	C	N	O	S	0	0	0
			1558	1008	259	283	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	340	GLY	-	expression tag	UNP Q9NZN8
B	341	PRO	-	expression tag	UNP Q9NZN8
B	342	HIS	-	expression tag	UNP Q9NZN8
B	343	MET	-	expression tag	UNP Q9NZN8

- Molecule 3 is a protein called CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	142	Total	C	N	O	S	0	0	0
			1210	789	198	217	6			

There are 19 discrepancies between the modelled and reference sequences:

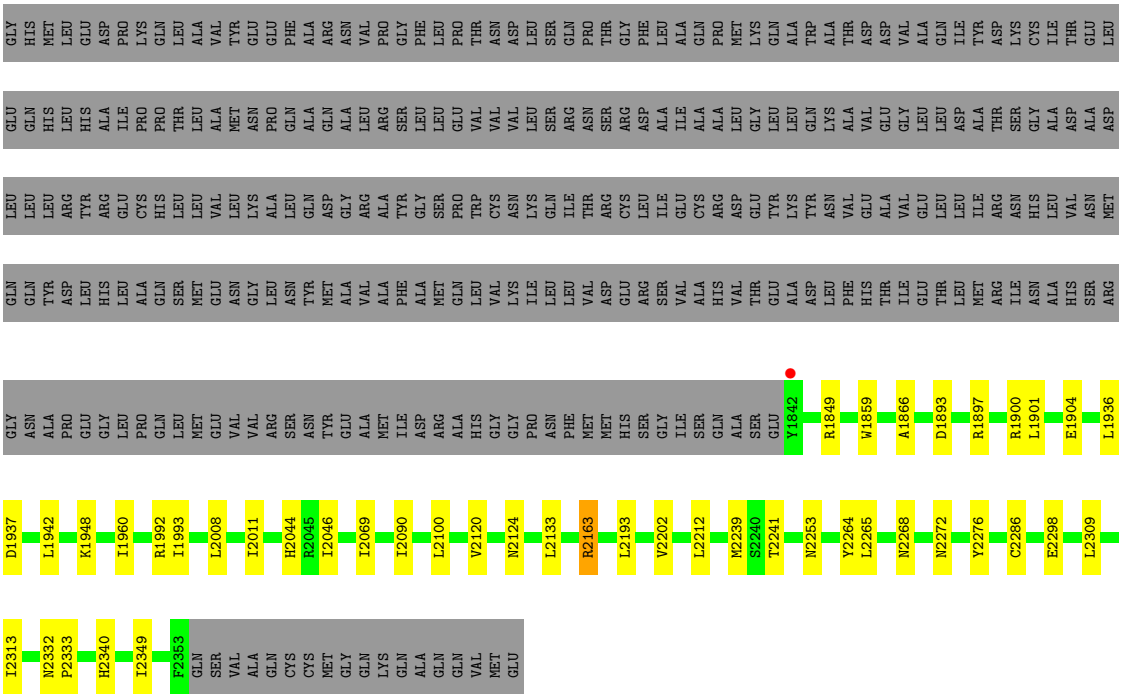
Chain	Residue	Modelled	Actual	Comment	Reference
C	588	MET	-	expression tag	UNP O75175
C	589	GLY	-	expression tag	UNP O75175
C	590	SER	-	expression tag	UNP O75175
C	591	SER	-	expression tag	UNP O75175
C	592	HIS	-	expression tag	UNP O75175
C	593	HIS	-	expression tag	UNP O75175
C	594	HIS	-	expression tag	UNP O75175
C	595	HIS	-	expression tag	UNP O75175
C	596	HIS	-	expression tag	UNP O75175
C	597	HIS	-	expression tag	UNP O75175
C	598	SER	-	expression tag	UNP O75175
C	599	SER	-	expression tag	UNP O75175
C	600	GLY	-	expression tag	UNP O75175
C	601	THR	-	expression tag	UNP O75175
C	602	GLY	-	expression tag	UNP O75175
C	603	SER	-	expression tag	UNP O75175
C	604	GLY	-	expression tag	UNP O75175
C	605	HIS	-	expression tag	UNP O75175
C	606	MET	-	expression tag	UNP O75175

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.

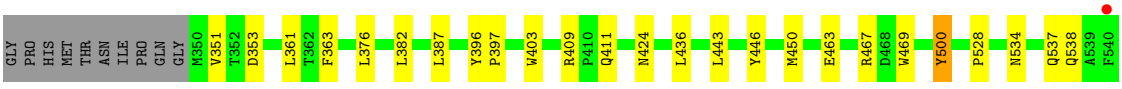
• Molecule 1: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1

Chain A:



• Molecule 2: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 2

Chain B:



• Molecule 3: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 3

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	91.57Å 165.92Å 78.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.48 – 3.20 48.48 – 3.20	Depositor EDS
% Data completeness (in resolution range)	98.8 (48.48-3.20) 98.9 (48.48-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	2.48 (at 3.19Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.224 , 0.273 0.220 , 0.270	Depositor DCC
R_{free} test set	1035 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	62.8	Xtriage
Anisotropy	0.751	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 39.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6864	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 3.90% 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	0.22	0/4206	0.36	0/5738
2	B	0.21	0/1606	0.37	0/2186
3	C	0.21	0/1259	0.34	0/1714
All	All	0.21	0/7071	0.36	0/9638

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

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	4096	0	4038	24	0
2	B	1558	0	1489	16	0
3	C	1210	0	1090	10	0
All	All	6864	0	6617	43	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.

All (43) 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:B:436:LEU:HB3	2:B:463:GLU:HG3	1.80	0.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:424:ASN:HD21	3:C:636:GLN:HG2	1.66	0.60
3:C:674:ILE:HG23	3:C:678:LEU:HD12	1.83	0.59
1:A:2163:ARG:NH1	1:A:2276:TYR:OH	2.36	0.58
2:B:534:ASN:HB3	2:B:537:GLN:HG3	1.88	0.55
1:A:1960:ILE:HD11	1:A:2011:ILE:HB	1.88	0.54
2:B:443:LEU:HD13	3:C:670:THR:HA	1.89	0.53
1:A:2264:TYR:O	1:A:2268:ASN:ND2	2.37	0.53
2:B:467:ARG:HB2	2:B:469:TRP:HD1	1.75	0.52
1:A:2100:LEU:HD21	2:B:382:LEU:HD13	1.91	0.51
2:B:463:GLU:O	2:B:467:ARG:HG2	2.10	0.51
2:B:351:VAL:HG12	2:B:353:ASP:H	1.76	0.50
1:A:2163:ARG:NE	1:A:2163:ARG:O	2.44	0.48
2:B:528:PRO:HG2	3:C:643:CYS:HB2	1.96	0.48
3:C:709:HIS:HB2	3:C:725:ILE:HG23	1.97	0.47
1:A:1900:ARG:NH1	1:A:1904:GLU:OE2	2.48	0.46
2:B:409:ARG:CZ	2:B:411:GLN:HE21	2.27	0.46
3:C:702:TYR:HE2	3:C:740:PHE:HE1	1.63	0.46
1:A:1849:ARG:HA	1:A:1901:LEU:HD13	1.97	0.46
1:A:1859:TRP:NE1	1:A:1942:LEU:HB2	2.30	0.46
1:A:2044:HIS:ND1	1:A:2046:ILE:HG22	2.32	0.45
1:A:1948:LYS:HG2	1:A:2008:LEU:HD11	1.99	0.45
3:C:646:PRO:O	3:C:651:GLN:NE2	2.50	0.44
1:A:1893:ASP:O	1:A:1897:ARG:HG2	2.16	0.44
1:A:2298:GLU:HG3	1:A:2340:HIS:HD2	1.83	0.44
1:A:2309:LEU:O	1:A:2313:ILE:HG12	2.18	0.44
3:C:726:TYR:CZ	3:C:735:ARG:HB2	2.53	0.44
1:A:2090:ILE:HD12	3:C:622:TRP:HE1	1.83	0.43
1:A:1936:LEU:HD22	1:A:1993:ILE:HG13	2.01	0.43
2:B:446:TYR:CE1	2:B:450:MET:HG3	2.54	0.43
1:A:2212:LEU:HD21	1:A:2265:LEU:HD22	2.00	0.42
1:A:1937:ASP:OD1	1:A:1992:ARG:NH1	2.51	0.42
1:A:2349:ILE:HD11	2:B:361:LEU:HD23	2.02	0.41
1:A:2133:LEU:HD21	1:A:2272:ASN:HA	2.03	0.41
2:B:396:TYR:N	2:B:397:PRO:HD2	2.36	0.41
3:C:708:ARG:HA	3:C:724:TYR:HD1	1.84	0.41
1:A:2193:LEU:HD13	1:A:2253:ASN:HB2	2.03	0.41
1:A:2332:ASN:HA	1:A:2333:PRO:HD3	1.84	0.41
1:A:2069:ILE:HG23	1:A:2120:VAL:HG21	2.03	0.40
2:B:467:ARG:HB2	2:B:469:TRP:CD1	2.55	0.40
2:B:500:TYR:HD1	2:B:500:TYR:HA	1.79	0.40
1:A:2124:ASN:HB2	2:B:376:LEU:HD22	2.04	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2239:MET:HE1	1:A:2286:CYS:HA	2.03	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	510/812 (63%)	488 (96%)	21 (4%)	1 (0%)	47 79
2	B	189/201 (94%)	181 (96%)	8 (4%)	0	100 100
3	C	140/166 (84%)	134 (96%)	6 (4%)	0	100 100
All	All	839/1179 (71%)	803 (96%)	35 (4%)	1 (0%)	51 83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1866	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	441/711 (62%)	438 (99%)	3 (1%)	84 94
2	B	164/175 (94%)	159 (97%)	5 (3%)	41 73

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	C	123/151 (82%)	123 (100%)	0	100 100
All	All	728/1037 (70%)	720 (99%)	8 (1%)	73 88

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2163	ARG
1	A	2202	VAL
1	A	2241	THR
2	B	363	PHE
2	B	387	LEU
2	B	403	TRP
2	B	500	TYR
2	B	538	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1934	HIS
1	A	2127	GLN
1	A	2199	ASN
1	A	2303	GLN
2	B	411	GLN
2	B	451	ASN
2	B	538	GLN

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 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	512/812 (63%)	-0.21	1 (0%)	95	94	36, 58, 111, 145	0
2	B	191/201 (95%)	-0.05	1 (0%)	91	86	54, 77, 112, 132	0
3	C	142/166 (85%)	0.03	2 (1%)	75	63	46, 80, 145, 157	0
All	All	845/1179 (71%)	-0.13	4 (0%)	91	86	36, 65, 121, 157	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1842	TYR	5.0
2	B	540	PHE	2.9
3	C	719	PHE	2.3
3	C	748	GLU	2.1

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.