



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 21, 2016 – 04:46 PM EDT

PDB ID : 5ANR  
Title : 4E-T - DDX6 - CNOT1 complex  
Authors : Basquin, J.; Oezguer, S.; Conti, E.  
Deposited on : 2015-09-08  
Resolution : 2.10 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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  
Mogul : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027257  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027257

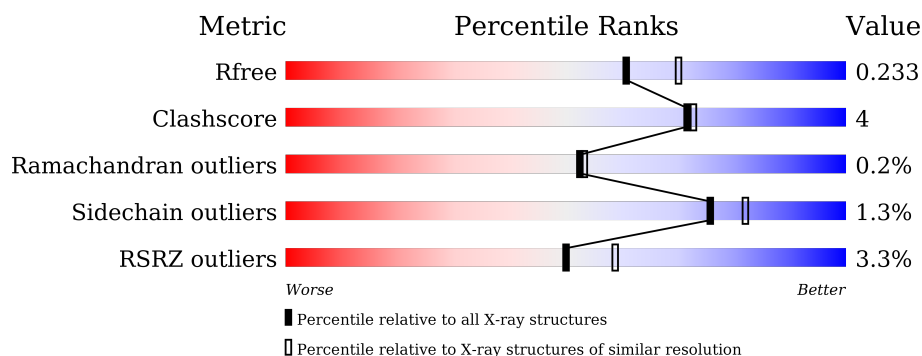
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
$R_{free}$	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	258	<div> <div>2%</div> <div> <div></div> <div>84%</div> <div>11%</div> <div>5%</div> </div> </div>
2	B	378	<div> <div></div> <div> <div>85%</div> <div>9%</div> <div>5%</div> </div> </div>
3	C	44	<div> <div>36%</div> <div> <div>43%</div> <div>9%</div> <div>48%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5205 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	245	Total	C	N	O	S	0	0	0
			1937	1248	323	358	8			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1052	GLN	-	EXPRESSION TAG	UNP A5YKK6
A	1053	GLY	-	EXPRESSION TAG	UNP A5YKK6
A	1054	PRO	-	EXPRESSION TAG	UNP A5YKK6
A	1055	ASP	-	EXPRESSION TAG	UNP A5YKK6
A	1056	SER	-	EXPRESSION TAG	UNP A5YKK6
A	1057	MET	-	EXPRESSION TAG	UNP A5YKK6

- Molecule 2 is a protein called PROBABLE ATP-DEPENDENT RNA HELICASE DDX6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	360	Total	C	N	O	S	0	0	0
			2836	1809	487	522	18			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	92	ARG	-	EXPRESSION TAG	UNP P26196
B	93	SER	-	EXPRESSION TAG	UNP P26196
B	94	MET	-	EXPRESSION TAG	UNP P26196

- Molecule 3 is a protein called EUKARYOTIC TRANSLATION INITIATION FACTOR 4E TRANSPORTER.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	23	Total	C	N	O	0	0	0
			156	99	25	32			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	196	ARG	-	EXPRESSION TAG	UNP Q9NRA8
C	197	SER	-	EXPRESSION TAG	UNP Q9NRA8
C	198	MET	-	EXPRESSION TAG	UNP Q9NRA8

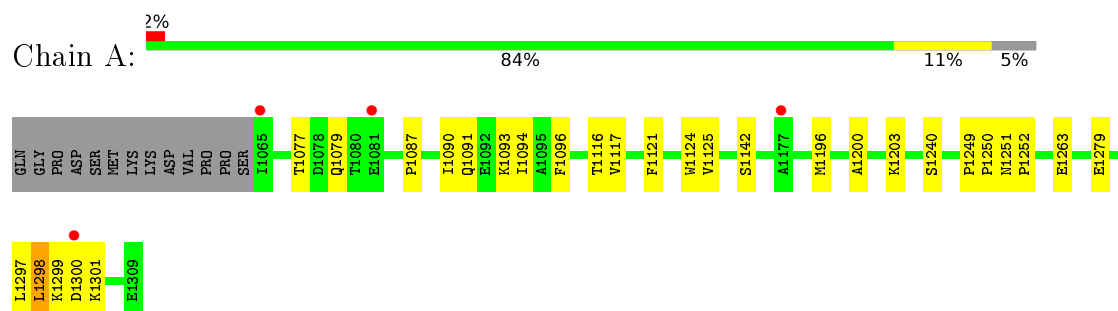
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	109	Total 109	O 109	0	0
4	B	166	Total 166	O 166	0	0
4	C	1	Total 1	O 1	0	0

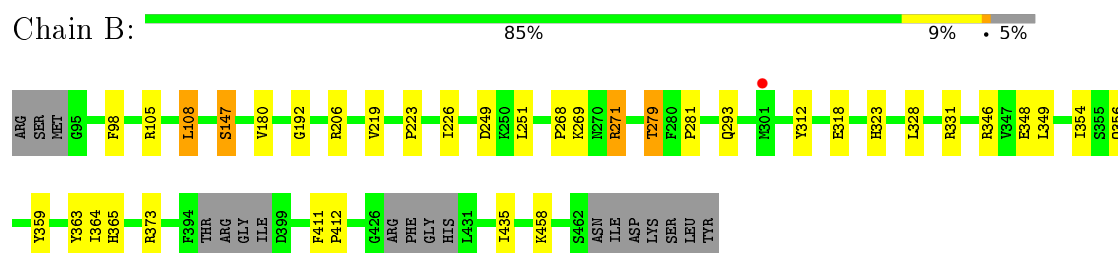
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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.

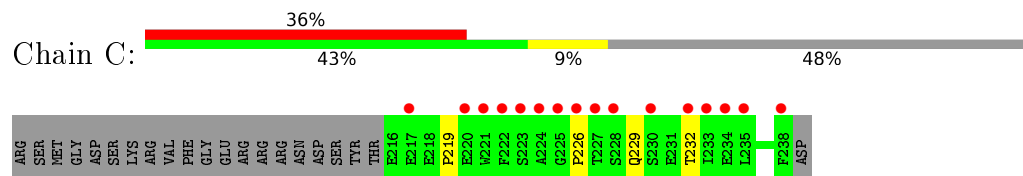
#### • Molecule 1: CCR4-NOT TRANSCRIPTION COMPLEX SUBUNIT 1



#### • Molecule 2: PROBABLE ATP-DEPENDENT RNA HELICASE DDX6



#### • Molecule 3: EUKARYOTIC TRANSLATION INITIATION FACTOR 4E TRANSPORTER



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.13Å 93.13Å 175.33Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.32 – 2.10 47.32 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.32-2.10) 99.8 (47.32-2.10)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.25 (at 2.10Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.193 , 0.232 0.195 , 0.233	Depositor DCC
$R_{free}$ test set	4924 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.8	Xtriage
Anisotropy	0.589	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 54.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.028 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5205	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.42	0/1973	0.51	1/2682 (0.0%)
2	B	0.39	0/2881	0.53	0/3892
3	C	0.31	0/160	0.42	0/220
All	All	0.40	0/5014	0.52	1/6794 (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
1	A	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1300	ASP	N-CA-C	-6.02	94.76	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1299	LYS	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	1937	0	1951	17	0
2	B	2836	0	2871	27	0
3	C	156	0	123	3	0
4	A	109	0	0	1	1
4	B	166	0	0	5	1
4	C	1	0	0	0	0
All	All	5205	0	4945	42	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 4.

The worst 5 of 42 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:318:GLU:OE2	2:B:346:ARG:NH2	2.08	0.85
4:A:2012:HOH:O	2:B:331:ARG:NH2	2.17	0.76
2:B:271:ARG:NH2	4:B:2097:HOH:O	2.25	0.69
1:A:1087:PRO:HG2	1:A:1090:ILE:HD13	1.78	0.64
2:B:279:THR:HG23	2:B:281:PRO:HD3	1.80	0.63

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:2088:HOH:O	4:B:2071:HOH:O[5_555]	2.13	0.07

## 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	243/258 (94%)	234 (96%)	8 (3%)	1 (0%)	39	37
2	B	354/378 (94%)	350 (99%)	4 (1%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	21/44 (48%)	18 (86%)	3 (14%)	0	100	100
All	All	618/680 (91%)	602 (97%)	15 (2%)	1 (0%)	52	53

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1301	LYS

### 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	216/242 (89%)	215 (100%)	1 (0%)	92	95
2	B	312/338 (92%)	307 (98%)	5 (2%)	70	76
3	C	13/39 (33%)	12 (92%)	1 (8%)	16	12
All	All	541/619 (87%)	534 (99%)	7 (1%)	76	82

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	271	ARG
3	C	232	THR
2	B	279	THR
2	B	108	LEU
2	B	356	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1207	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 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	245/258 (94%)	-0.23	4 (1%) 74 79	28, 42, 62, 78	0
2	B	360/378 (95%)	-0.25	1 (0%) 94 95	32, 43, 61, 81	0
3	C	23/44 (52%)	2.68	16 (69%) 0 0	63, 72, 84, 90	0
All	All	628/680 (92%)	-0.13	21 (3%) 50 59	28, 43, 68, 90	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	225	GLY	7.0
3	C	232	THR	4.7
3	C	223	SER	3.7
3	C	224	ALA	3.6
3	C	227	THR	3.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.