WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 12, 2017 – 08:25 pm GMT

PDB ID : 6PAX

Title : CRYSTAL STRUCTURE OF THE HUMAN PAX-6 PAIRED DOMAIN-DNA COMPLEX REVEALS A GENERAL MODEL FOR PAX PROTEIN-DNA INTERACTIONS

Authors : Xu, H.E.; Rould, M.A.; Xu, W.; Epstein, J.A.; Maas, R.L.; Pabo, C.O.

Deposited on : 1999-04-22

Resolution : 2.50 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
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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

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 2.50 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	Percentile Ranks	Value
Clashscore		31
Ramachandran outliers		0
Sidechain outliers		7.0%
RSRZ outliers		11.9%

Worse　　　　　　　　　　　　　　　　　*Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
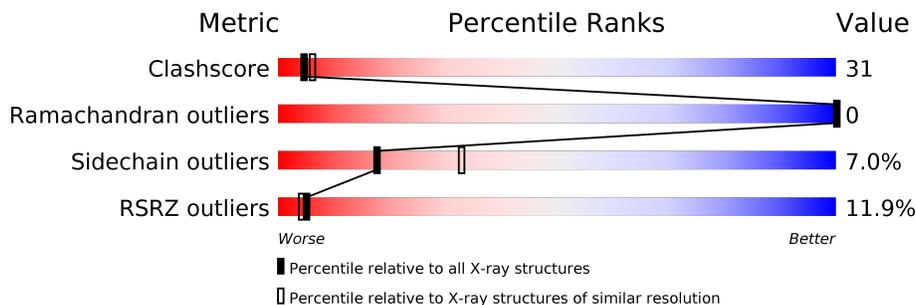
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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. 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	B	26	19% 27% 65% 8%
2	C	26	4% 15% 77% 8%
3	A	133	12% 75% 21% •

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2528 atoms, of which 367 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 DNA chain called 26 NUCLEOTIDE DNA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	26	Total	C	H	N	O	P	0	0	0
			585	254	54	100	152	25			

- Molecule 2 is a DNA chain called 26 NUCLEOTIDE DNA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	26	Total	C	H	N	O	P	0	0	0
			581	254	52	94	156	25			

- Molecule 3 is a protein called HOMEOBOX PROTEIN PAX-6.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	A	133	Total	C	H	N	O	S	0	0	0
			1278	627	261	196	190	4			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	ARG	LYS	CONFLICT	UNP P26367
A	59	ALA	GLU	CONFLICT	UNP P26367
A	89	GLN	ARG	CONFLICT	UNP P26367

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	40	Total	O	0	0
			40	40		
4	B	15	Total	O	0	0
			15	15		
4	C	29	Total	O	0	0
			29	29		

3 Residue-property plots ⓘ

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: 26 NUCLEOTIDE DNA

Chain B:
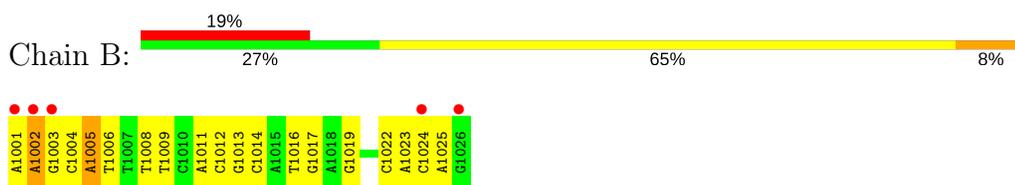
• Molecule 2: 26 NUCLEOTIDE DNA

Chain C:

• Molecule 3: HOMEOBOX PROTEIN PAX-6

Chain A:
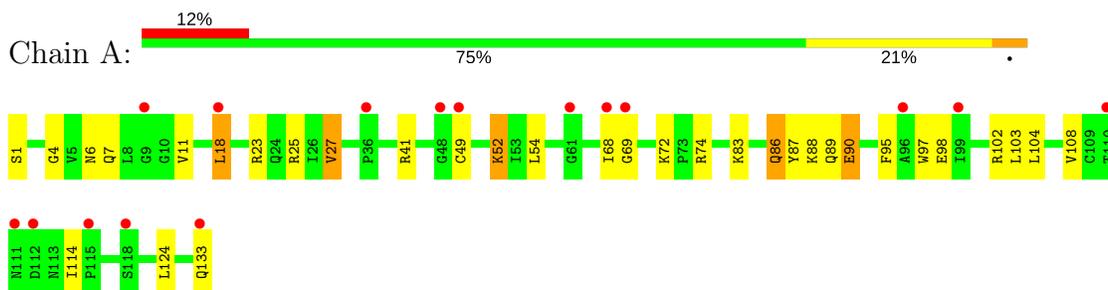
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	33.84Å 61.69Å 171.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 19.82 – 2.00	Depositor EDS
% Data completeness (in resolution range)	96.3 (20.00-2.50) 76.6 (19.82-2.00)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.233 , 0.256 (Not available) , (Not available)	Depositor DCC
R_{free} test set	NotAvailable	DCC
Wilson B-factor (Å^2)	(Not available)	Xtriage
Anisotropy	(Not available)	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 78.7	EDS
L-test for twinning[1]	$< |L| > =$ (Not available), $< L^2 > =$ (Not available)	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	2528	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *(Not available)*

---

[1]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	B	0.84	0/596	1.13	1/918 (0.1%)
2	C	0.80	0/592	0.96	0/912
3	A	0.57	0/1033	0.66	0/1397
All	All	0.71	0/2221	0.90	1/3227 (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	B	0	1
2	C	0	3
All	All	0	4

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1002	DA	C5'-C4'-C3'	-8.91	98.06	114.10

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1005	DA	Sidechain
2	C	2003	DC	Sidechain
2	C	2009	DA	Sidechain
2	C	2022	DA	Sidechain

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	B	531	54	294	40	1
2	C	529	52	296	54	1
3	A	1017	261	1040	30	0
4	A	40	0	0	2	0
4	B	15	0	0	3	0
4	C	29	0	0	5	0
All	All	2161	367	1630	111	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 31.

The worst 5 of 111 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:C:2002:DT:H2"	2:C:2003:DC:C5'	1.61	1.30
2:C:2001:DT:H2"	2:C:2002:DT:C5'	1.66	1.25
2:C:2023:DA:H2"	2:C:2024:DT:C5'	1.78	1.14
2:C:2002:DT:H2"	2:C:2003:DC:H5"	1.23	1.10
2:C:2023:DA:H2"	2:C:2024:DT:H5"	1.18	1.09

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)
1:B:1002:DA:N1	2:C:2001:DT:H3[2_555]	1.49	0.11

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	
3	A	131/133 (98%)	128 (98%)	3 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	
3	A	114/114 (100%)	106 (93%)	8 (7%)	18	33

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	86	GLN
3	A	133	GLN
3	A	90	GLU
3	A	52	LYS
3	A	89	GLN

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

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 carbohydrates 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, 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(Å²)	Q<0.9
1	B	26/26 (100%)	1.30	5 (19%)	1	1	41, 49, 70, 81	0
2	C	26/26 (100%)	1.23	1 (3%)	41	43	37, 53, 59, 63	0
3	A	133/133 (100%)	1.23	16 (12%)	5	4	32, 46, 64, 82	0
All	All	185/185 (100%)	1.24	22 (11%)	5	4	32, 48, 66, 82	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	110	THR	3.8
1	B	1002	DA	3.5
1	B	1003	DG	3.5
3	A	61	GLY	3.1
1	B	1024	DC	3.0

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 (i)

There are no such residues in this entry.