wwPDB X-ray Structure Validation Summary Report (i)

May 12, 2020 – 11:38 pm BST

PDB ID	:	1P27
Title	:	Crystal Structure of the Human Y14/Magoh complex
Authors	:	Lau, C.K.; Diem, M.D.; Dreyfuss, G.; Van Duyne, G.D.
Deposited on	:	2003-04-14
Resolution	:	2.00 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:

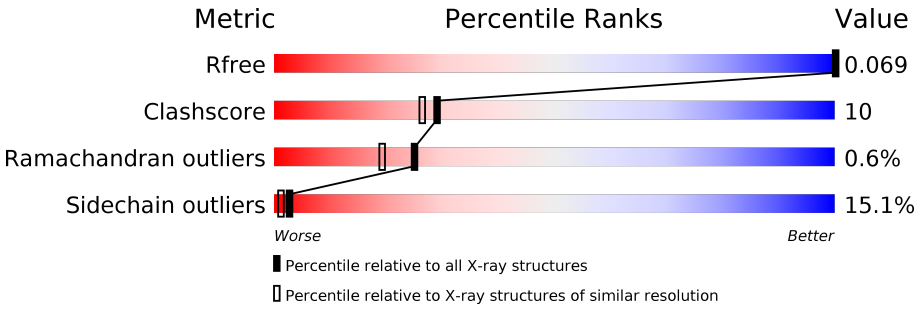
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	144	73% 21% 5% •
1	C	144	75% 19% 6%
2	B	106	70% 13% • 13%
2	D	106	51% 29% 7% 13%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3966 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 Mago nashi protein homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	144	Total	C	N	O	S	0	0	0
			1196	771	200	222	3			
1	C	144	Total	C	N	O	S	0	0	0
			1196	771	200	222	3			

- Molecule 2 is a protein called RNA-binding protein 8A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	92	Total	C	N	O	S	0	0	0
			737	468	123	143	3			
2	D	92	Total	C	N	O	S	0	0	0
			737	468	123	143	3			

- Molecule 3 is water.

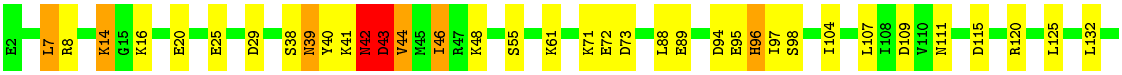
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	33	Total	O	0	0
			33	33		
3	B	12	Total	O	0	0
			12	12		
3	C	41	Total	O	0	0
			41	41		
3	D	14	Total	O	0	0
			14	14		

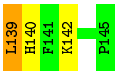
3 Residue-property plots (i)

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. 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. 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: Mago nashi protein homolog

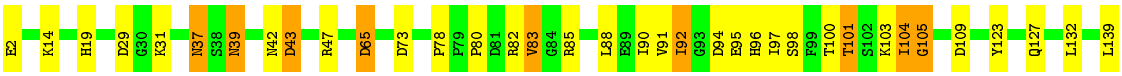
Chain A:





- Molecule 1: Mago nashi protein homolog

Chain C:



- Molecule 2: RNA-binding protein 8A

Chain B:

- Molecule 2: RNA-binding protein 8A

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	47.15Å 108.45Å 50.94Å 90.00° 90.24° 90.00°	Depositor
Resolution (Å)	46.11 – 2.00 46.11 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.8 (46.11-2.00) 97.7 (46.11-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	4.21 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.220 , 0.268 0.390 , 0.069	Depositor DCC
R_{free} test set	1701 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	38.8	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 51.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.64$, $< L^2 > = 0.51$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	3966	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.74% 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	0.80	0/1225	0.97	6/1649 (0.4%)
1	C	0.73	0/1225	0.98	7/1649 (0.4%)
2	B	0.78	0/756	0.87	1/1024 (0.1%)
2	D	0.78	0/756	0.89	2/1024 (0.2%)
All	All	0.77	0/3962	0.94	16/5346 (0.3%)

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	94	ASP	CB-CG-OD2	6.63	124.26	118.30
1	C	85	ARG	NE-CZ-NH1	6.29	123.44	120.30
1	A	109	ASP	CB-CG-OD2	6.26	123.93	118.30
1	A	73	ASP	CB-CG-OD2	6.21	123.89	118.30
2	D	138	ASP	CB-CG-OD2	6.17	123.86	118.30

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	1196	0	1177	22	1
1	C	1196	0	1177	23	1
2	B	737	0	697	12	0
2	D	737	0	697	23	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	33	0	0	1	0
3	B	12	0	0	0	0
3	C	41	0	0	0	0
3	D	14	0	0	1	0
All	All	3966	0	3748	73	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 10.

The worst 5 of 73 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:8:ARG:NH1	1:A:89:GLU:OE1	1.86	1.08
1:C:90:ILE:O	1:C:96:HIS:HD2	1.46	0.99
1:C:90:ILE:O	1:C:96:HIS:CD2	2.34	0.81
1:C:37:ASN:O	1:C:37:ASN:ND2	2.15	0.80
1:C:91:VAL:HG22	1:C:96:HIS:CD2	2.20	0.76

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:A:61:LYS:NZ	1:C:109:ASP:OD1[1_556]	1.98	0.22

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	142/144 (99%)	136 (96%)	4 (3%)	2 (1%)	11 5
1	C	142/144 (99%)	134 (94%)	7 (5%)	1 (1%)	22 16

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	B	90/106 (85%)	89 (99%)	1 (1%)	0	100 100
2	D	90/106 (85%)	88 (98%)	2 (2%)	0	100 100
All	All	464/500 (93%)	447 (96%)	14 (3%)	3 (1%)	25 19

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	42	ASN
1	A	43	ASP
1	C	105	GLY

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	132/132 (100%)	108 (82%)	24 (18%)	1 1
1	C	132/132 (100%)	115 (87%)	17 (13%)	4 2
2	B	77/90 (86%)	68 (88%)	9 (12%)	5 3
2	D	77/90 (86%)	64 (83%)	13 (17%)	2 1
All	All	418/444 (94%)	355 (85%)	63 (15%)	3 1

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	98	GLU
1	C	37	ASN
2	D	109	ARG
2	B	100	LYS
2	B	118	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	103	HIS
2	B	135	ASN
1	C	140	HIS
2	B	90	HIS
2	D	81	HIS

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.