wwPDB X-ray Structure Validation Summary Report (i)

Nov 21, 2023 – 12:10 AM JST

PDB ID	:	7E2M
Title	:	Crystal structure of the RWD domain of human GCN2 - 2
Authors	:	Hei, Z.; Zhou, J.; Liu, Z.; Wang, J.; Fang, P.
Deposited on	:	2021-02-05
Resolution	:	2.35 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

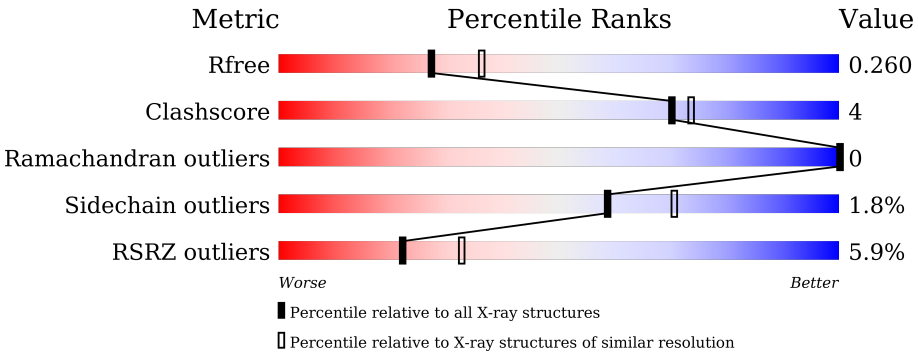
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	135	4% 87% 9%
1	B	135	4% 81% 13%
1	C	135	7% 86% 10%
1	D	135	4% 90%
1	E	135	8% 84% 10%
1	F	135	7% 89% 7%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6235 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 eIF-2-alpha kinase GCN2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	130	Total	C	N	O	S	0	0	0
			1002	640	177	182	3			
1	B	129	Total	C	N	O	S	0	1	0
			1010	643	175	188	4			
1	C	129	Total	C	N	O	S	0	0	0
			990	631	174	181	4			
1	D	129	Total	C	N	O	S	0	0	0
			981	629	172	176	4			
1	E	129	Total	C	N	O	S	0	0	0
			984	627	171	183	3			
1	F	129	Total	C	N	O	S	0	0	0
			985	630	173	179	3			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	MET	-	initiating methionine	UNP Q9P2K8
A	6	GLY	-	expression tag	UNP Q9P2K8
A	7	SER	-	expression tag	UNP Q9P2K8
A	8	SER	-	expression tag	UNP Q9P2K8
A	9	HIS	-	expression tag	UNP Q9P2K8
A	10	HIS	-	expression tag	UNP Q9P2K8
A	11	HIS	-	expression tag	UNP Q9P2K8
A	12	HIS	-	expression tag	UNP Q9P2K8
A	13	HIS	-	expression tag	UNP Q9P2K8
A	14	HIS	-	expression tag	UNP Q9P2K8
A	15	SER	-	expression tag	UNP Q9P2K8
A	16	GLY	-	expression tag	UNP Q9P2K8
B	5	MET	-	initiating methionine	UNP Q9P2K8
B	6	GLY	-	expression tag	UNP Q9P2K8
B	7	SER	-	expression tag	UNP Q9P2K8
B	8	SER	-	expression tag	UNP Q9P2K8
B	9	HIS	-	expression tag	UNP Q9P2K8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	10	HIS	-	expression tag	UNP Q9P2K8
B	11	HIS	-	expression tag	UNP Q9P2K8
B	12	HIS	-	expression tag	UNP Q9P2K8
B	13	HIS	-	expression tag	UNP Q9P2K8
B	14	HIS	-	expression tag	UNP Q9P2K8
B	15	SER	-	expression tag	UNP Q9P2K8
B	16	GLY	-	expression tag	UNP Q9P2K8
C	5	MET	-	initiating methionine	UNP Q9P2K8
C	6	GLY	-	expression tag	UNP Q9P2K8
C	7	SER	-	expression tag	UNP Q9P2K8
C	8	SER	-	expression tag	UNP Q9P2K8
C	9	HIS	-	expression tag	UNP Q9P2K8
C	10	HIS	-	expression tag	UNP Q9P2K8
C	11	HIS	-	expression tag	UNP Q9P2K8
C	12	HIS	-	expression tag	UNP Q9P2K8
C	13	HIS	-	expression tag	UNP Q9P2K8
C	14	HIS	-	expression tag	UNP Q9P2K8
C	15	SER	-	expression tag	UNP Q9P2K8
C	16	GLY	-	expression tag	UNP Q9P2K8
D	5	MET	-	initiating methionine	UNP Q9P2K8
D	6	GLY	-	expression tag	UNP Q9P2K8
D	7	SER	-	expression tag	UNP Q9P2K8
D	8	SER	-	expression tag	UNP Q9P2K8
D	9	HIS	-	expression tag	UNP Q9P2K8
D	10	HIS	-	expression tag	UNP Q9P2K8
D	11	HIS	-	expression tag	UNP Q9P2K8
D	12	HIS	-	expression tag	UNP Q9P2K8
D	13	HIS	-	expression tag	UNP Q9P2K8
D	14	HIS	-	expression tag	UNP Q9P2K8
D	15	SER	-	expression tag	UNP Q9P2K8
D	16	GLY	-	expression tag	UNP Q9P2K8
E	5	MET	-	initiating methionine	UNP Q9P2K8
E	6	GLY	-	expression tag	UNP Q9P2K8
E	7	SER	-	expression tag	UNP Q9P2K8
E	8	SER	-	expression tag	UNP Q9P2K8
E	9	HIS	-	expression tag	UNP Q9P2K8
E	10	HIS	-	expression tag	UNP Q9P2K8
E	11	HIS	-	expression tag	UNP Q9P2K8
E	12	HIS	-	expression tag	UNP Q9P2K8
E	13	HIS	-	expression tag	UNP Q9P2K8
E	14	HIS	-	expression tag	UNP Q9P2K8
E	15	SER	-	expression tag	UNP Q9P2K8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	16	GLY	-	expression tag	UNP Q9P2K8
F	5	MET	-	initiating methionine	UNP Q9P2K8
F	6	GLY	-	expression tag	UNP Q9P2K8
F	7	SER	-	expression tag	UNP Q9P2K8
F	8	SER	-	expression tag	UNP Q9P2K8
F	9	HIS	-	expression tag	UNP Q9P2K8
F	10	HIS	-	expression tag	UNP Q9P2K8
F	11	HIS	-	expression tag	UNP Q9P2K8
F	12	HIS	-	expression tag	UNP Q9P2K8
F	13	HIS	-	expression tag	UNP Q9P2K8
F	14	HIS	-	expression tag	UNP Q9P2K8
F	15	SER	-	expression tag	UNP Q9P2K8
F	16	GLY	-	expression tag	UNP Q9P2K8

- Molecule 2 is water.

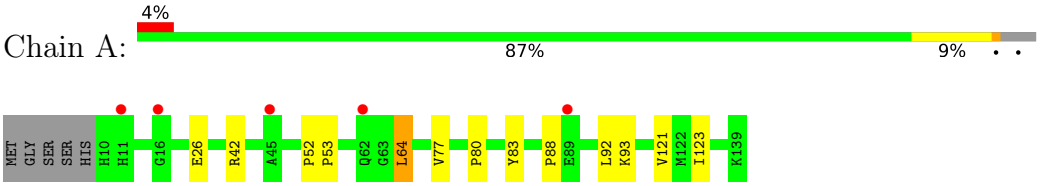
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	56	Total O 56 56	0	0
2	B	65	Total O 65 65	0	0
2	C	41	Total O 41 41	0	0
2	D	39	Total O 39 39	0	0
2	E	39	Total O 39 39	0	0
2	F	43	Total O 43 43	0	0

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.

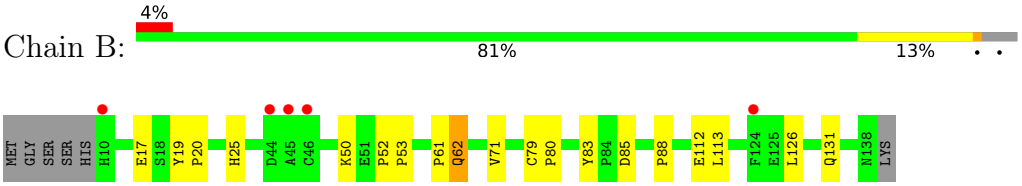
- Molecule 1: eIF-2-alpha kinase GCN2

Chain A:



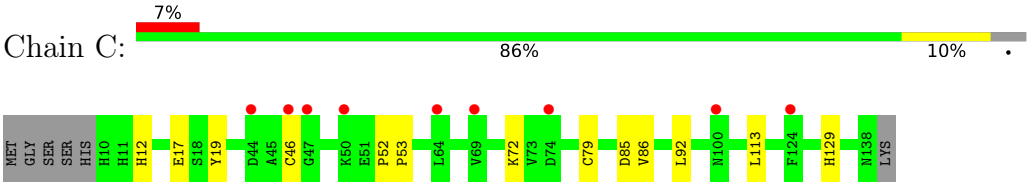
- Molecule 1: eIF-2-alpha kinase GCN2

Chain B:



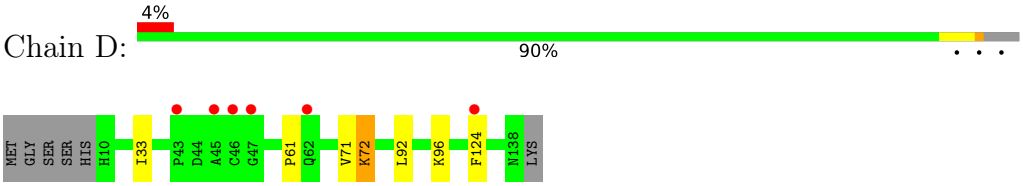
- Molecule 1: eIF-2-alpha kinase GCN2

Chain C:



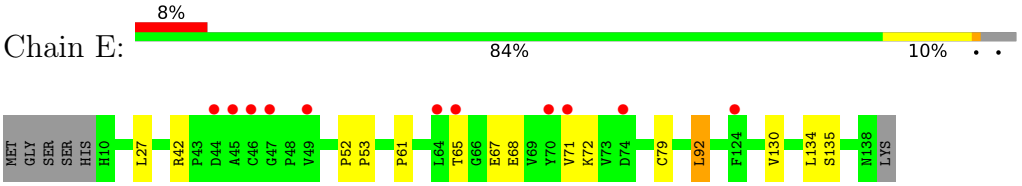
- Molecule 1: eIF-2-alpha kinase GCN2

Chain D:



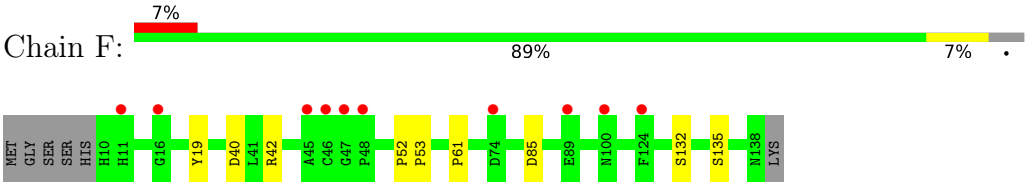
- Molecule 1: eIF-2-alpha kinase GCN2

Chain E:



- Molecule 1: eIF-2-alpha kinase GCN2

Chain F:


7%
89%
7%
MET
GLY
SER
SER
HIS
H10
H11
G16
Y19
D40
L41
R42
A45
C46
G47
P48
P52
P53
P61
D74
D85
E89
N100
F124
S132
S135
N138
LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α, β, γ	158.44Å 158.44Å 54.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.90 – 2.35 44.90 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.0 (44.90-2.35) 99.1 (44.90-2.35)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.61 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.13	Depositor
R, R_{free}	0.234 , 0.261 0.235 , 0.260	Depositor DCC
R_{free} test set	1565 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	36.1	Xtriage
Anisotropy	0.072	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 31.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.035 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6235	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 65.34 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.1880e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.24	0/1029	0.37	0/1402
1	B	0.24	0/1036	0.38	0/1410
1	C	0.24	0/1017	0.38	0/1388
1	D	0.24	0/1007	0.38	0/1373
1	E	0.24	0/1010	0.39	0/1380
1	F	0.26	0/1011	0.39	0/1379
All	All	0.24	0/6110	0.38	0/8332

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	1002	0	955	8	0
1	B	1010	0	964	12	0
1	C	990	0	934	6	0
1	D	981	0	934	3	0
1	E	984	0	919	10	0
1	F	985	0	935	4	0
2	A	56	0	0	1	0
2	B	65	0	0	1	0
2	C	41	0	0	0	0
2	D	39	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	39	0	0	0	0
2	F	43	0	0	0	0
All	All	6235	0	5641	42	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 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)
1:A:93:LYS:NZ	2:A:202:HOH:O	2.36	0.58
1:E:65:THR:HG22	1:E:68:GLU:CD	2.26	0.56
1:E:27:LEU:HD11	1:E:53:PRO:HB3	1.88	0.56
1:A:42:ARG:HD3	1:A:53:PRO:HA	1.88	0.55
1:A:77:VAL:HG11	1:A:123:ILE:HD12	1.89	0.54

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	128/135 (95%)	127 (99%)	1 (1%)	0	100	100
1	B	128/135 (95%)	128 (100%)	0	0	100	100
1	C	127/135 (94%)	127 (100%)	0	0	100	100
1	D	127/135 (94%)	127 (100%)	0	0	100	100
1	E	127/135 (94%)	122 (96%)	5 (4%)	0	100	100
1	F	127/135 (94%)	124 (98%)	3 (2%)	0	100	100
All	All	764/810 (94%)	755 (99%)	9 (1%)	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	
1	A	104/120 (87%)	102 (98%)	2 (2%)	57	68
1	B	107/120 (89%)	105 (98%)	2 (2%)	57	68
1	C	103/120 (86%)	100 (97%)	3 (3%)	42	52
1	D	100/120 (83%)	98 (98%)	2 (2%)	55	66
1	E	101/120 (84%)	99 (98%)	2 (2%)	55	66
1	F	101/120 (84%)	100 (99%)	1 (1%)	76	85
All	All	616/720 (86%)	604 (98%)	12 (2%)	59	68

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	72	LYS
1	D	92	LEU
1	F	132	SER
1	E	72	LYS
1	B	62[B]	GLN

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

Mol	Chain	Res	Type
1	E	39	GLN

5.3.3 RNA

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 (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

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	130/135 (96%)	0.19	5 (3%)	40	53	21, 29, 49, 72	57 (43%)
1	B	129/135 (95%)	0.21	5 (3%)	39	52	23, 32, 47, 72	52 (40%)
1	C	129/135 (95%)	0.39	9 (6%)	16	24	25, 39, 58, 115	49 (37%)
1	D	129/135 (95%)	0.24	6 (4%)	31	44	27, 35, 66, 79	41 (31%)
1	E	129/135 (95%)	0.43	11 (8%)	10	16	25, 34, 65, 107	59 (45%)
1	F	129/135 (95%)	0.39	10 (7%)	13	19	26, 34, 60, 120	44 (34%)
All	All	775/810 (95%)	0.31	46 (5%)	22	33	21, 34, 63, 120	302 (38%)

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	48	PRO	7.8
1	E	46	CYS	6.7
1	F	46	CYS	5.9
1	E	47	GLY	4.4
1	E	45	ALA	4.3

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 (i)

There are no such residues in this entry.