



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 13, 2025 – 01:13 pm BST

PDB ID : 9I3F / pdb\_00009i3f  
Title : Crystal structure of the AGR2 and IRE1beta\_loop complex  
Authors : Yan, Y.; Ron, D.  
Deposited on : 2025-01-22  
Resolution : 1.90 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](mailto: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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

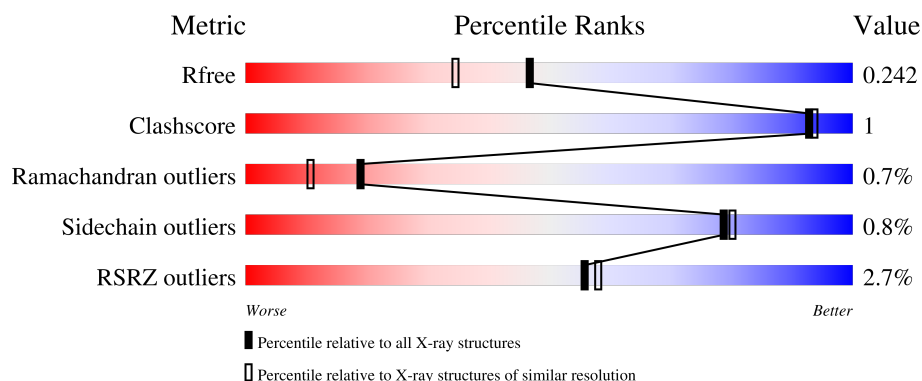
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	F	53	<div> <div>6%</div> <div>45%</div> <div>55%</div> </div>
1	G	53	<div> <div>11%</div> <div>32%</div> <div>68%</div> </div>
1	H	53	<div> <div>6%</div> <div>30%</div> <div>68%</div> </div>
1	I	53	<div> <div>4%</div> <div>34%</div> <div>66%</div> </div>
2	A	131	<div> <div>%</div> <div>94%</div> <div>5% ..</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	131	<div><div></div><div>92%</div><div>6% ..</div></div>
2	C	131	<div>%<div><div></div><div>95%</div><div>. ..</div></div></div>
2	D	131	<div><div></div><div>94%</div><div>5% .</div></div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5208 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 Serine/threonine-protein kinase/endoribonuclease IRE2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	I	18	Total	C	N	O	0	0	0
			136	83	22	31			
1	F	24	Total	C	N	O	0	0	0
			171	106	28	37			
1	G	17	Total	C	N	O	0	0	0
			120	74	18	28			
1	H	17	Total	C	N	O	0	0	0
			128	79	21	28			

- Molecule 2 is a protein called Anterior gradient protein 2 homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	130	Total	C	N	O	S	0	3	0
			1086	697	181	203	5			
2	B	130	Total	C	N	O	S	0	2	0
			1072	686	180	202	4			
2	C	130	Total	C	N	O	S	0	1	0
			1065	683	179	199	4			
2	D	130	Total	C	N	O	S	0	0	0
			1058	679	178	197	4			

- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		
3	C	1	Total	Mg	0	0
			1	1		

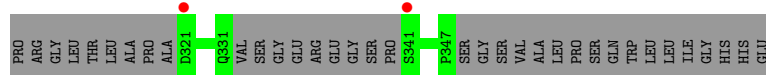
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	I	15	Total 15	O 15	0	0
4	A	77	Total 77	O 77	0	0
4	B	76	Total 76	O 76	0	0
4	C	75	Total 75	O 75	0	0
4	D	83	Total 83	O 83	0	0
4	F	16	Total 16	O 16	0	0
4	G	12	Total 12	O 12	0	0
4	H	15	Total 15	O 15	0	0

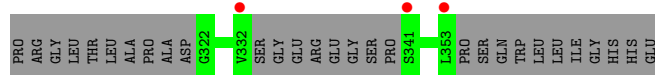
### 3 Residue-property plots [i](#)

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.

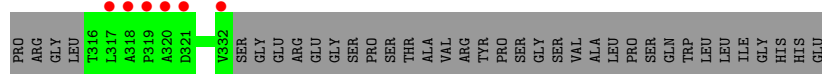
- Molecule 1: Serine/threonine-protein kinase/endoribonuclease IRE2



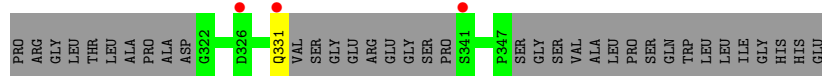
- Molecule 1: Serine/threonine-protein kinase/endoribonuclease IRE2



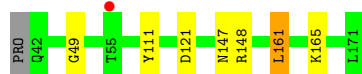
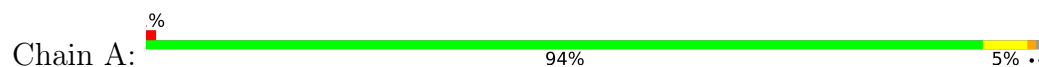
- Molecule 1: Serine/threonine-protein kinase/endoribonuclease IRE2




- Molecule 1: Serine/threonine-protein kinase/endoribonuclease IRE2

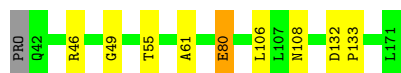


- Molecule 2: Anterior gradient protein 2 homolog

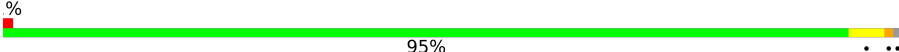


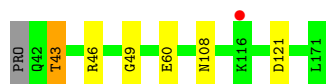
- Molecule 2: Anterior gradient protein 2 homolog

Chain B:  92% 6% ..



- Molecule 2: Anterior gradient protein 2 homolog

Chain C:  95% ..



- Molecule 2: Anterior gradient protein 2 homolog

Chain D:  94% 5% •



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.20Å 81.82Å 120.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	67.76 – 1.90 67.76 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.6 (67.76-1.90) 99.5 (67.76-1.90)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.75 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0430	Depositor
R, $R_{free}$	0.199 , 0.234 0.206 , 0.242	Depositor DCC
$R_{free}$ test set	2863 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.2	Xtriage
Anisotropy	0.936	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5208	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	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 34.13 % 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 7.1589e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	F	0.55	0/172	0.89	0/235
1	G	0.51	0/121	1.09	0/168
1	H	0.50	0/129	0.95	0/176
1	I	0.51	0/137	0.96	0/187
2	A	0.48	0/1110	0.93	1/1503 (0.1%)
2	B	0.47	0/1095	0.89	0/1483
2	C	0.48	0/1088	0.90	1/1475 (0.1%)
2	D	0.49	0/1081	0.94	1/1465 (0.1%)
All	All	0.48	0/4933	0.92	3/6692 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	121	ASP	CA-CB-CG	6.68	119.28	112.60
2	C	121	ASP	CA-CB-CG	5.24	117.83	112.60
2	A	121	ASP	CA-CB-CG	5.15	117.75	112.60

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	F	171	0	169	0	0
1	G	120	0	116	0	0
1	H	128	0	122	0	0
1	I	136	0	126	0	0
2	A	1086	0	1091	4	0
2	B	1072	0	1079	5	0
2	C	1065	0	1075	3	0
2	D	1058	0	1069	3	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	77	0	0	0	0
4	B	76	0	0	0	0
4	C	75	0	0	1	0
4	D	83	0	0	1	0
4	F	16	0	0	0	0
4	G	12	0	0	0	0
4	H	15	0	0	0	0
4	I	15	0	0	0	0
All	All	5208	0	4847	14	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 1.

All (14) 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:55[B]:THR:HG21	2:B:61:ALA:HB2	1.72	0.71
2:B:55[B]:THR:HG22	2:B:106:LEU:O	2.08	0.53
2:D:155:ALA:HB3	4:D:263:HOH:O	2.10	0.52
2:C:43:THR:HG23	4:C:363:HOH:O	2.09	0.51
2:B:132:ASP:CG	2:B:133:PRO:HD2	2.37	0.50
2:A:161:LEU:CD1	2:A:165:LYS:HE3	2.44	0.47
2:A:111[B]:TYR:CE2	2:C:60:GLU:OE1	2.70	0.45
2:D:127:ARG:HG3	2:D:150:TYR:HB2	1.99	0.44
2:D:46:ARG:HD2	2:D:108:ASN:OD1	2.18	0.44
2:B:46:ARG:HD2	2:B:108:ASN:OD1	2.18	0.43
2:A:147:ASN:C	2:A:148:ARG:HG3	2.45	0.42
2:C:46:ARG:HD2	2:C:108:ASN:OD1	2.20	0.41
2:B:80:GLU:H	2:B:80:GLU:CD	2.29	0.40
2:A:161:LEU:HD13	2:A:165:LYS:HE3	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	F	20/53 (38%)	20 (100%)	0	0	100	100
1	G	15/53 (28%)	13 (87%)	2 (13%)	0	100	100
1	H	13/53 (24%)	13 (100%)	0	0	100	100
1	I	14/53 (26%)	14 (100%)	0	0	100	100
2	A	131/131 (100%)	128 (98%)	2 (2%)	1 (1%)	16	8
2	B	130/131 (99%)	129 (99%)	0	1 (1%)	16	8
2	C	129/131 (98%)	128 (99%)	0	1 (1%)	16	8
2	D	128/131 (98%)	127 (99%)	0	1 (1%)	16	8
All	All	580/736 (79%)	572 (99%)	4 (1%)	4 (1%)	19	11

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	49	GLY
2	B	49	GLY
2	D	49	GLY
2	C	49	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	F	20/43 (46%)	20 (100%)	0	100	100
1	G	14/43 (33%)	14 (100%)	0	100	100
1	H	15/43 (35%)	14 (93%)	1 (7%)	13	6
1	I	16/43 (37%)	16 (100%)	0	100	100
2	A	120/118 (102%)	119 (99%)	1 (1%)	79	80
2	B	119/118 (101%)	118 (99%)	1 (1%)	79	80
2	C	118/118 (100%)	117 (99%)	1 (1%)	79	80
2	D	117/118 (99%)	117 (100%)	0	100	100
All	All	539/644 (84%)	535 (99%)	4 (1%)	79	83

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	161	LEU
2	B	80	GLU
2	C	43	THR
1	H	331	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	42	GLN
2	D	42	GLN
2	D	98	GLN
1	H	331	GLN

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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<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	F	24/53 (45%)	0.45	3 (12%) 9 9	16, 21, 40, 44	0
1	G	17/53 (32%)	1.39	6 (35%) 1 1	16, 20, 60, 62	0
1	H	17/53 (32%)	0.94	3 (17%) 4 4	19, 29, 46, 47	0
1	I	18/53 (33%)	1.01	2 (11%) 12 12	17, 30, 58, 62	0
2	A	130/131 (99%)	0.05	1 (0%) 82 84	9, 20, 33, 40	3 (2%)
2	B	130/131 (99%)	0.01	0 100 100	9, 20, 33, 36	2 (1%)
2	C	130/131 (99%)	-0.03	1 (0%) 82 84	9, 20, 30, 40	1 (0%)
2	D	130/131 (99%)	-0.02	0 100 100	13, 20, 30, 39	0
All	All	596/736 (80%)	0.12	16 (2%) 56 58	9, 20, 35, 62	6 (1%)

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	320	ALA	5.7
1	F	353	LEU	4.1
1	I	321	ASP	4.0
1	G	317	LEU	3.8
1	G	321	ASP	3.5
1	I	341	SER	3.2
1	G	318	ALA	2.9
2	A	55[A]	THR	2.8
1	F	341	SER	2.8
1	H	326	ASP	2.8
1	G	319	PRO	2.7
1	H	341	SER	2.6
1	F	332	VAL	2.5
1	G	332	VAL	2.5
2	C	116	LYS	2.1
1	H	331	GLN	2.1

## 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 oligosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MG	A	201	1/1	0.98	0.02	14,14,14,14	0
3	MG	C	201	1/1	0.98	0.06	15,15,15,15	0
3	MG	B	201	1/1	0.99	0.03	15,15,15,15	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.