



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 04:21 PM GMT

PDB ID : 4A2W  
Title : STRUCTURE OF FULL-LENGTH DUCK RIG-I  
Authors : Kowalinski, E.; Lunardi, T.; Mccarthy, A.A.; Cusack, S.  
Deposited on : 2011-09-29  
Resolution : 3.70 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

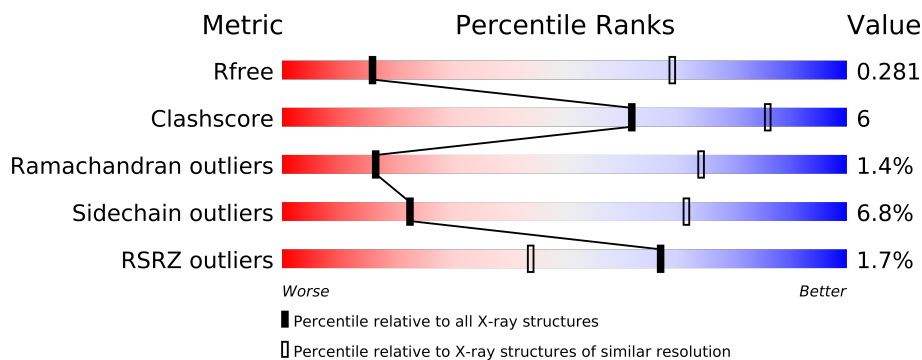
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.70 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))
$R_{free}$	66092	1098 (4.00-3.40)
Clashscore	79885	1009 (3.94-3.46)
Ramachandran outliers	78287	1016 (3.98-3.42)
Sidechain outliers	78261	1014 (3.98-3.42)
RSRZ outliers	66119	1099 (4.00-3.40)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	936	
1	B	936	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 10880 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

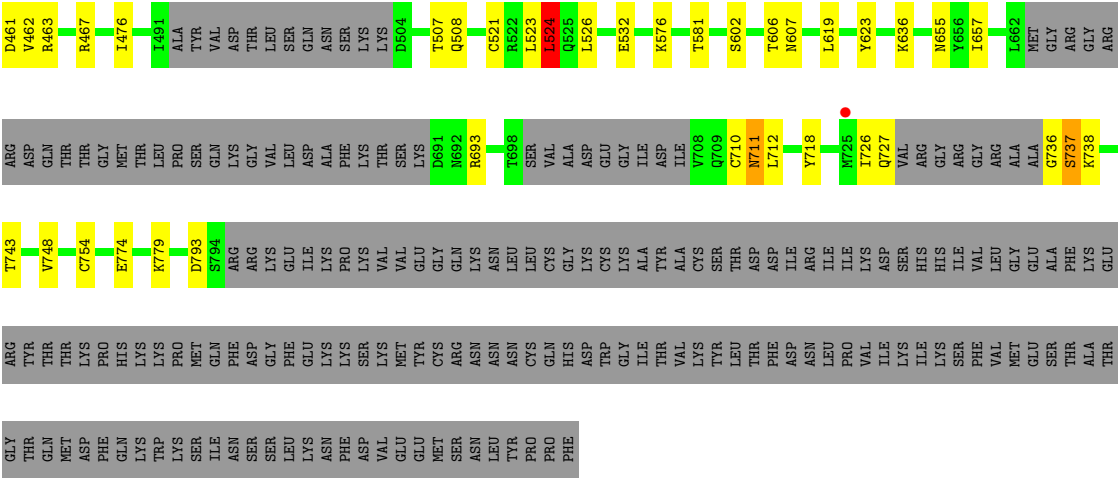
- Molecule 1 is a protein called RETINOIC ACID INDUCIBLE PROTEIN I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	678	Total	C	N	O	S	0	0	0
			5454	3446	940	1035	33			
1	B	674	Total	C	N	O	S	0	0	0
			5426	3427	936	1030	33			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP D3TI84
A	-1	ALA	-	EXPRESSION TAG	UNP D3TI84
A	0	MET	-	EXPRESSION TAG	UNP D3TI84
A	1	GLY	-	EXPRESSION TAG	UNP D3TI84
B	-2	GLY	-	EXPRESSION TAG	UNP D3TI84
B	-1	ALA	-	EXPRESSION TAG	UNP D3TI84
B	0	MET	-	EXPRESSION TAG	UNP D3TI84
B	1	GLY	-	EXPRESSION TAG	UNP D3TI84





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	163.32Å 163.32Å 152.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.77 – 3.70 42.77 – 3.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (42.77-3.70) 99.9 (42.77-3.70)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.49 (at 3.66Å)	Xtriage
Refinement program	REFMAC 5.6.0116	Depositor
R, $R_{free}$	0.218 , 0.274 0.227 , 0.281	Depositor DCC
$R_{free}$ test set	1160 reflections (5.41%)	DCC
Wilson B-factor (Å <sup>2</sup> )	135.3	Xtriage
Anisotropy	0.131	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 117.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 22592 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10880	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	174.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.45% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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	0.48	0/5539	0.62	0/7473
1	B	0.46	0/5511	0.58	0/7434
All	All	0.47	0/11050	0.60	0/14907

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.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	699	SER	Peptide

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5454	0	0	36	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	5426	0	0	27	1
All	All	10880	0	0	62	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (62) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:717:GLU:OE2	1:A:744:SER:OG	1.98	0.81
1:A:468:ARG:NH1	1:A:556:SER:O	2.31	0.64
1:B:44:GLU:OE1	1:B:47:ARG:NH1	2.33	0.61
1:A:466:LYS:NZ	1:A:614:GLU:OE2	2.34	0.60
1:B:508:GLN:NE2	1:B:636:LYS:O	2.35	0.60
1:A:710:CYS:SG	1:A:711:ASN:N	2.75	0.59
1:B:268:PRO:O	1:B:271:SER:OG	2.21	0.59
1:B:524:LEU:CD2	1:B:526:LEU:CD1	2.82	0.57
1:A:524:LEU:CD2	1:A:526:LEU:CD1	2.82	0.57
1:A:44:GLU:OE1	1:A:47:ARG:NH1	2.38	0.57
1:A:268:PRO:O	1:A:271:SER:OG	2.23	0.55
1:B:113:MET:O	1:B:116:VAL:N	2.40	0.54
1:B:108:ARG:NH2	1:B:532:GLU:OE2	2.40	0.54
1:B:710:CYS:SG	1:B:711:ASN:N	2.82	0.52
1:A:723:THR:O	1:A:727:GLN:NE2	2.42	0.52
1:B:328:SER:CA	1:B:353:ILE:CD1	2.87	0.52
1:A:108:ARG:NH2	1:A:532:GLU:OE2	2.43	0.52
1:A:113:MET:O	1:A:116:VAL:N	2.44	0.51
1:A:328:SER:CA	1:A:353:ILE:CD1	2.90	0.50
1:B:37:LEU:CD2	1:B:59:ALA:CB	2.91	0.49
1:A:431:SER:O	1:A:434:SER:OG	2.31	0.48
1:A:711:ASN:ND2	1:A:711:ASN:C	2.67	0.48
1:A:185:TRP:O	1:A:186:ASP:C	2.53	0.47
1:A:456:ASN:N	1:A:456:ASN:OD1	2.48	0.47
1:A:37:LEU:CD2	1:A:59:ALA:CB	2.93	0.47
1:A:473:PHE:CD2	1:A:598:LEU:CD2	2.97	0.47
1:A:606:THR:CG2	1:A:606:THR:O	2.63	0.47
1:A:465:VAL:CG2	1:A:742:VAL:CG1	2.94	0.46
1:B:113:MET:O	1:B:114:LEU:C	2.52	0.46
1:A:574:ASN:O	1:A:577:ASN:N	2.48	0.46
1:B:726:ILE:CG2	1:B:727:GLN:N	2.77	0.46
1:A:559:ALA:O	1:A:609:ASN:ND2	2.48	0.46
1:B:736:GLY:O	1:B:737:SER:CB	2.63	0.46
1:B:25:VAL:CG2	1:B:75:ALA:CB	2.94	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:431:SER:O	1:B:434:SER:OG	2.34	0.46
1:A:595:GLU:OE1	1:B:576:LYS:NZ	2.49	0.46
1:B:438:ILE:CD1	1:B:438:ILE:N	2.79	0.45
1:B:185:TRP:O	1:B:186:ASP:C	2.54	0.45
1:B:402:SER:OG	1:B:403:GLN:N	2.48	0.45
1:B:461:ASP:OD2	1:B:463:ARG:NE	2.50	0.44
1:B:416:VAL:O	1:B:416:VAL:CG1	2.65	0.44
1:A:25:VAL:CG2	1:A:75:ALA:CB	2.95	0.44
1:A:468:ARG:NH2	1:A:607:ASN:O	2.51	0.44
1:B:623:TYR:CE2	1:B:657:ILE:CD1	3.00	0.44
1:A:722:VAL:O	1:A:723:THR:C	2.56	0.44
1:B:264:LEU:CD1	1:B:377:CYS:SG	3.06	0.44
1:A:438:ILE:N	1:A:438:ILE:CD1	2.81	0.43
1:B:297:VAL:CG2	1:B:369:PHE:CG	3.02	0.43
1:A:726:ILE:CD1	1:A:739:CYS:SG	3.07	0.43
1:A:736:GLY:O	1:A:737:SER:CB	2.67	0.43
1:B:108:ARG:NH1	1:B:521:CYS:O	2.52	0.42
1:B:743:THR:CG2	1:B:748:VAL:CG1	2.97	0.42
1:B:109:ILE:O	1:B:109:ILE:CG2	2.68	0.42
1:A:113:MET:O	1:A:114:LEU:C	2.58	0.42
1:A:297:VAL:CG2	1:A:369:PHE:CG	3.03	0.42
1:A:623:TYR:CE2	1:A:657:ILE:CD1	3.03	0.42
1:A:402:SER:OG	1:A:403:GLN:N	2.53	0.41
1:A:574:ASN:O	1:A:575:VAL:C	2.59	0.41
1:A:773:ASP:O	1:A:774:GLU:C	2.59	0.41
1:A:381:THR:CG2	1:A:382:GLY:N	2.84	0.41
1:B:421:ASN:N	1:B:421:ASN:ND2	2.69	0.40
1:A:109:ILE:O	1:A:112:THR:N	2.55	0.40

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	Distance(Å)	Clash(Å)
1:A:397:LYS:O	1:B:779:LYS:NZ[1.565]	2.18	0.02

## 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	
1	A	664/936 (71%)	594 (90%)	61 (9%)	9 (1%)	16	74
1	B	660/936 (70%)	595 (90%)	56 (8%)	9 (1%)	16	74
All	All	1324/1872 (71%)	1189 (90%)	117 (9%)	18 (1%)	16	74

All (18) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	524	LEU
1	A	737	SER
1	B	524	LEU
1	B	737	SER
1	A	66	ARG
1	B	66	ARG
1	B	774	GLU
1	A	19	SER
1	A	738	LYS
1	A	774	GLU
1	B	19	SER
1	B	246	ALA
1	B	738	LYS
1	A	523	LEU
1	A	699	SER
1	B	163	GLU
1	B	523	LEU
1	A	575	VAL

### 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	606/831 (73%)	563 (93%)	43 (7%)	21	70
1	B	602/831 (72%)	563 (94%)	39 (6%)	24	73
All	All	1208/1662 (73%)	1126 (93%)	82 (7%)	22	71

All (82) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	8	SER
1	A	45	GLU
1	A	60	VAL
1	A	64	GLU
1	A	113	MET
1	A	131	ASP
1	A	146	LYS
1	A	164	HIS
1	A	171	LEU
1	A	174	ASP
1	A	186	ASP
1	A	245	LYS
1	A	253	LEU
1	A	264	LEU
1	A	278	ILE
1	A	297	VAL
1	A	301	THR
1	A	306	TYR
1	A	342	GLU
1	A	367	SER
1	A	421	ASN
1	A	438	ILE
1	A	453	ARG
1	A	456	ASN
1	A	462	VAL
1	A	466	LYS
1	A	476	ILE
1	A	507	THR
1	A	524	LEU
1	A	581	THR
1	A	602	SER
1	A	606	THR
1	A	607	ASN
1	A	619	LEU
1	A	655	ASN
1	A	693	ARG
1	A	711	ASN
1	A	712	LEU
1	A	713	VAL
1	A	721	ASN
1	A	726	ILE
1	A	743	THR
1	A	754	CYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	8	SER
1	B	45	GLU
1	B	60	VAL
1	B	64	GLU
1	B	113	MET
1	B	146	LYS
1	B	164	HIS
1	B	171	LEU
1	B	174	ASP
1	B	186	ASP
1	B	245	LYS
1	B	253	LEU
1	B	264	LEU
1	B	278	ILE
1	B	297	VAL
1	B	301	THR
1	B	306	TYR
1	B	342	GLU
1	B	367	SER
1	B	421	ASN
1	B	438	ILE
1	B	456	ASN
1	B	462	VAL
1	B	467	ARG
1	B	476	ILE
1	B	507	THR
1	B	524	LEU
1	B	581	THR
1	B	602	SER
1	B	606	THR
1	B	607	ASN
1	B	619	LEU
1	B	655	ASN
1	B	693	ARG
1	B	711	ASN
1	B	712	LEU
1	B	718	TYR
1	B	754	CYS
1	B	793	ASP

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 chains 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, 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	678/936 (72%)	0.14	5 (0%) 84 66	85, 157, 218, 263	0
1	B	674/936 (72%)	0.32	18 (2%) 52 32	95, 185, 265, 315	0
All	All	1352/1872 (72%)	0.23	23 (1%) 67 44	85, 169, 253, 315	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	66	ARG	3.9
1	B	26	LEU	3.4
1	A	37	LEU	3.3
1	B	37	LEU	3.2
1	B	56	PHE	3.0
1	B	725	MET	2.9
1	B	40	ARG	2.8
1	B	22	PRO	2.6
1	A	57	LEU	2.6
1	B	95	LYS	2.4
1	B	57	LEU	2.4
1	A	56	PHE	2.4
1	B	84	LEU	2.3
1	B	327	ILE	2.3
1	B	42	ARG	2.2
1	B	341	ILE	2.2
1	B	39	GLU	2.2
1	A	53	ALA	2.2
1	B	173	LEU	2.2
1	A	736	GLY	2.1
1	B	93	PHE	2.1
1	B	346	ILE	2.1
1	B	41	ILE	2.0

## 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 ⓘ

There are no such residues in this entry.