Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 02:50 AM GMT

PDB ID	:	2IW2
Title	:	CRYSTAL STRUCTURE OF HUMAN PROLIDASE
Authors	:	Mueller, U.; Niesen, F.H.; Roske, Y.; Goetz, F.; Behlke, J.; Buessow, K.; Heinemann, U.
Deposited on	:	2006-06-24
Resolution	:	1.82 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

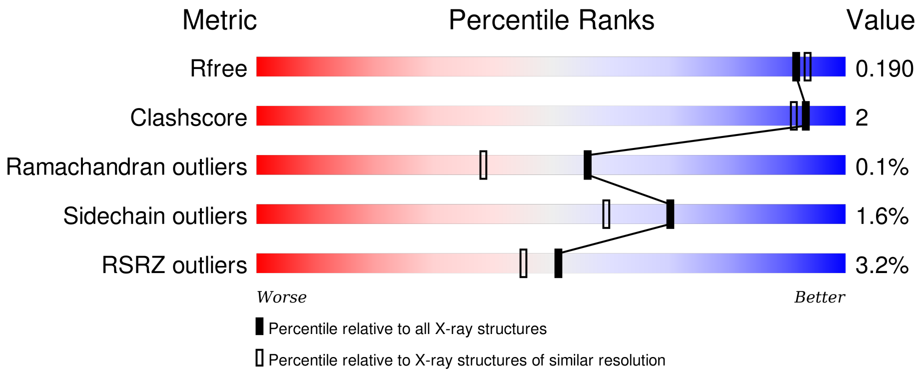
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 1.82 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}	91344	5422 (1.84-1.80)
Clashscore	102246	6347 (1.84-1.80)
Ramachandran outliers	100387	6276 (1.84-1.80)
Sidechain outliers	100360	6276 (1.84-1.80)
RSRZ outliers	91569	5439 (1.84-1.80)

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	A	494	4% / 91% 5% •
1	B	494	2% / 91% 5% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8501 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 XAA-PRO DIPEPTIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	478	Total 3764	C 2372	N 662	O 699	S 31	0	13	1
1	B	478	Total 3741	C 2359	N 655	O 697	S 30	0	5	1

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	436	PHE	LEU	CONFLICT	UNP P12955
B	436	PHE	LEU	CONFLICT	UNP P12955

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total 2	Na 2	0	0
2	A	3	Total 3	Na 3	0	0

- Molecule 3 is water.

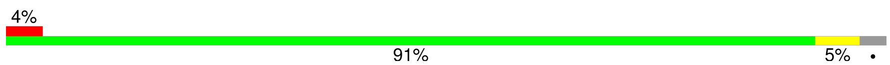
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	487	Total 487	O 487	0	0
3	B	504	Total 504	O 504	0	0

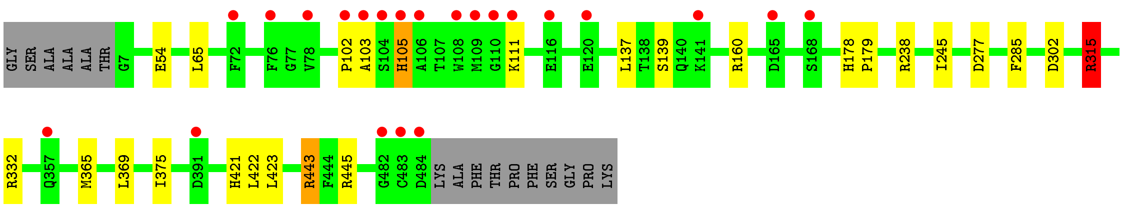
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: XAA-PRO DIPEPTIDASE

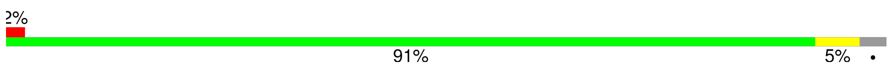
Chain A:

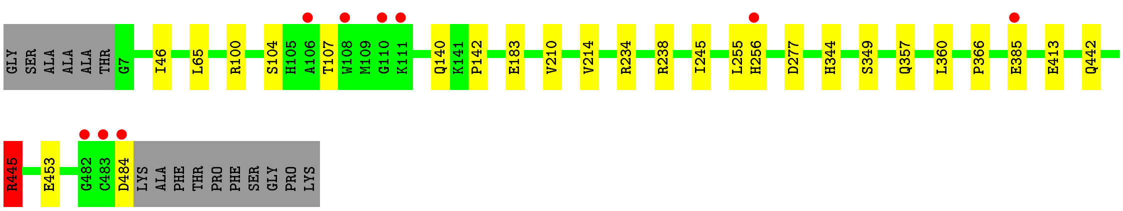




- Molecule 1: XAA-PRO DIPEPTIDASE

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	103.58Å 108.52Å 211.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.51 – 1.82 29.51 – 1.82	Depositor EDS
% Data completeness (in resolution range)	97.5 (29.51-1.82) 97.5 (29.51-1.82)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.45 (at 1.82Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.157 , 0.192 0.158 , 0.190	Depositor DCC
R_{free} test set	5187 reflections (5.26%)	DCC
Wilson B-factor (Å^2)	20.7	Xtriage
Anisotropy	0.220	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 48.5	EDS
Estimated twinning fraction	0.025 for -k,-h,-l	Xtriage
L-test for twinning[2]	$< \|L\| >$ = 0.49, $< L^2 >$ = 0.32	Xtriage
Outliers	0 of 103802 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8501	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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

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	0.69	0/3914	0.73	4/5292 (0.1%)
1	B	0.72	0/3847	0.70	1/5204 (0.0%)
All	All	0.70	0/7761	0.72	5/10496 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	443	ARG	NE-CZ-NH2	-10.94	114.83	120.30
1	A	443	ARG	NE-CZ-NH1	9.88	125.24	120.30
1	A	315[A]	ARG	NE-CZ-NH1	5.63	123.12	120.30
1	A	315[B]	ARG	NE-CZ-NH1	5.63	123.12	120.30
1	B	445	ARG	NE-CZ-NH1	5.09	122.84	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	3764	0	3670	17	0
1	B	3741	0	3646	13	1
2	A	3	0	0	0	0
2	B	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	487	0	0	4	0
3	B	504	0	0	2	0
All	All	8501	0	7316	28	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 2.

All (28) 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:B:385:GLU:O	1:B:385:GLU:HG3	1.98	0.63
1:B:245:ILE:HB	1:B:277:ASP:HB3	1.82	0.60
1:A:302:ASP:OD2	1:A:443:ARG:HD2	2.03	0.59
1:A:315[B]:ARG:NH1	3:A:2357:HOH:O	2.35	0.59
1:A:285:PHE:CE1	1:A:375[A]:ILE:HD11	2.38	0.58
1:A:102:PRO:O	1:A:105:HIS:HB2	2.03	0.58
1:A:365[A]:MET:CE	1:A:369:LEU:HA	2.34	0.57
1:B:65:LEU:HD21	1:B:238:ARG:HD3	1.89	0.54
1:A:245:ILE:HB	1:A:277:ASP:HB3	1.89	0.53
1:A:445:ARG:HD2	3:A:2455:HOH:O	2.11	0.51
1:A:365[A]:MET:HE3	1:A:369:LEU:HA	1.93	0.49
1:A:65:LEU:HD21	1:A:238:ARG:HD3	1.95	0.49
1:B:210:VAL:O	1:B:214:VAL:HG23	2.12	0.49
1:A:422:LEU:HD11	1:B:107:THR:HG23	1.97	0.47
1:B:484:ASP:N	3:B:2504:HOH:O	2.49	0.45
1:B:255:LEU:HD11	1:B:366:PRO:HG2	1.98	0.45
1:A:332:ARG:NH2	3:A:2372:HOH:O	2.50	0.44
1:B:46:ILE:HD12	1:B:142:PRO:HB3	2.01	0.43
1:B:413:GLU:HB3	1:B:453:GLU:HG3	2.00	0.42
1:B:349:SER:HB2	3:B:2463:HOH:O	2.18	0.42
1:A:178:HIS:HB3	1:A:179:PRO:HD3	2.01	0.42
1:B:442:GLN:HA	1:B:445:ARG:HG3	2.03	0.41
1:A:160:ARG:NE	3:A:2203:HOH:O	2.53	0.41
1:A:365[A]:MET:HE1	1:A:369:LEU:HA	2.04	0.40
1:A:421:HIS:CE1	1:B:104:SER:HA	2.56	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	Interatomic distance (Å)	Clash overlap (Å)
1:B:100[A]:ARG:NH2	1:B:344:HIS:O[5_455]	2.08	0.12

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	489/494 (99%)	478 (98%)	10 (2%)	1 (0%)	52	35
1	B	481/494 (97%)	471 (98%)	10 (2%)	0	100	100
All	All	970/988 (98%)	949 (98%)	20 (2%)	1 (0%)	56	38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	103	ALA

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	412/410 (100%)	403 (98%)	9 (2%)	60	45
1	B	404/410 (98%)	398 (98%)	6 (2%)	72	62
All	All	816/820 (100%)	801 (98%)	15 (2%)	70	54

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	GLU
1	A	105	HIS
1	A	111	LYS
1	A	137	LEU
1	A	139	SER
1	A	315[A]	ARG
1	A	315[B]	ARG
1	A	423[A]	LEU
1	A	423[B]	LEU
1	B	140	GLN
1	B	183	GLU
1	B	234	ARG
1	B	357	GLN
1	B	360	LEU
1	B	445	ARG

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

Mol	Chain	Res	Type
1	A	56	GLN
1	A	113	HIS
1	A	421	HIS
1	B	42	GLN

5.3.3 RNA (i)

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 5 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 (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	478/494 (96%)	-0.08	22 (4%)	36	30	11, 15, 25, 42	0
1	B	478/494 (96%)	-0.17	9 (1%)	70	65	9, 15, 26, 38	0
All	All	956/988 (96%)	-0.13	31 (3%)	51	45	9, 15, 26, 42	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	103	ALA	7.8
1	A	484	ASP	6.1
1	B	256[A]	HIS	6.0
1	B	484	ASP	5.1
1	A	104	SER	4.9
1	A	102	PRO	4.9
1	A	483	CYS	4.2
1	A	106	ALA	4.0
1	B	111	LYS	3.9
1	A	105	HIS	3.7
1	B	110	GLY	3.6
1	B	106	ALA	3.4
1	A	110	GLY	3.2
1	A	109[A]	MET	3.2
1	B	483	CYS	3.1
1	A	78	VAL	2.7
1	A	357[A]	GLN	2.7
1	A	165	ASP	2.6
1	B	385	GLU	2.5
1	A	111	LYS	2.4
1	A	76	PHE	2.4
1	A	391	ASP	2.4
1	B	108	TRP	2.3
1	A	141	LYS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	482	GLY	2.3
1	A	108	TRP	2.2
1	A	72	PHE	2.1
1	A	116	GLU	2.0
1	A	120	GLU	2.0
1	A	168	SER	2.0
1	B	482	GLY	2.0

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} 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	LLDF	B-factors($\AA^2$)	Q<0.9
2	NA	A	1486	1/1	0.99	0.12	1.13	21,21,21,21	0
2	NA	B	1484	1/1	0.83	0.12	0.65	39,39,39,39	0
2	NA	A	1485	1/1	0.91	0.11	-0.74	38,38,38,38	0
2	NA	B	1485	1/1	0.98	0.07	-1.72	25,25,25,25	0
2	NA	A	1484	1/1	0.99	0.05	-1.79	21,21,21,21	0

6.5 Other polymers (i)

There are no such residues in this entry.