



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 04:26 AM BST

PDB ID : 1J2E  
Title : Crystal structure of Human Dipeptidyl peptidase IV  
Authors : Hiramatsu, H.; Kyono, K.; Higashiyama, Y.; Fukushima, C.; Shima, H.;  
Sugiyama, S.; Inaka, K.; Yamamoto, A.; Shimizu, R.  
Deposited on : 2002-12-30  
Resolution : 2.60 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](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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

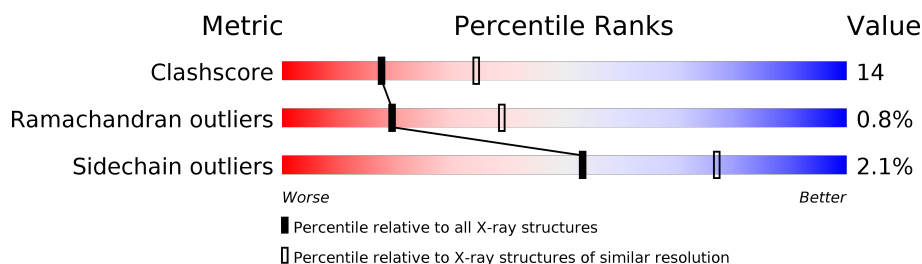
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

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 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	740	
1	B	740	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12355 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 Dipeptidyl peptidase IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	729	Total	C	N	O	S	0	0	0
			5971	3831	983	1131	26			
1	B	729	Total	C	N	O	S	0	0	0
			5971	3831	983	1131	26			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	767	HIS	-	expression tag	UNP P27487
A	768	HIS	-	expression tag	UNP P27487
A	769	HIS	-	expression tag	UNP P27487
A	770	HIS	-	expression tag	UNP P27487
A	771	HIS	-	expression tag	UNP P27487
A	772	HIS	-	expression tag	UNP P27487
B	767	HIS	-	expression tag	UNP P27487
B	768	HIS	-	expression tag	UNP P27487
B	769	HIS	-	expression tag	UNP P27487
B	770	HIS	-	expression tag	UNP P27487
B	771	HIS	-	expression tag	UNP P27487
B	772	HIS	-	expression tag	UNP P27487

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is water.

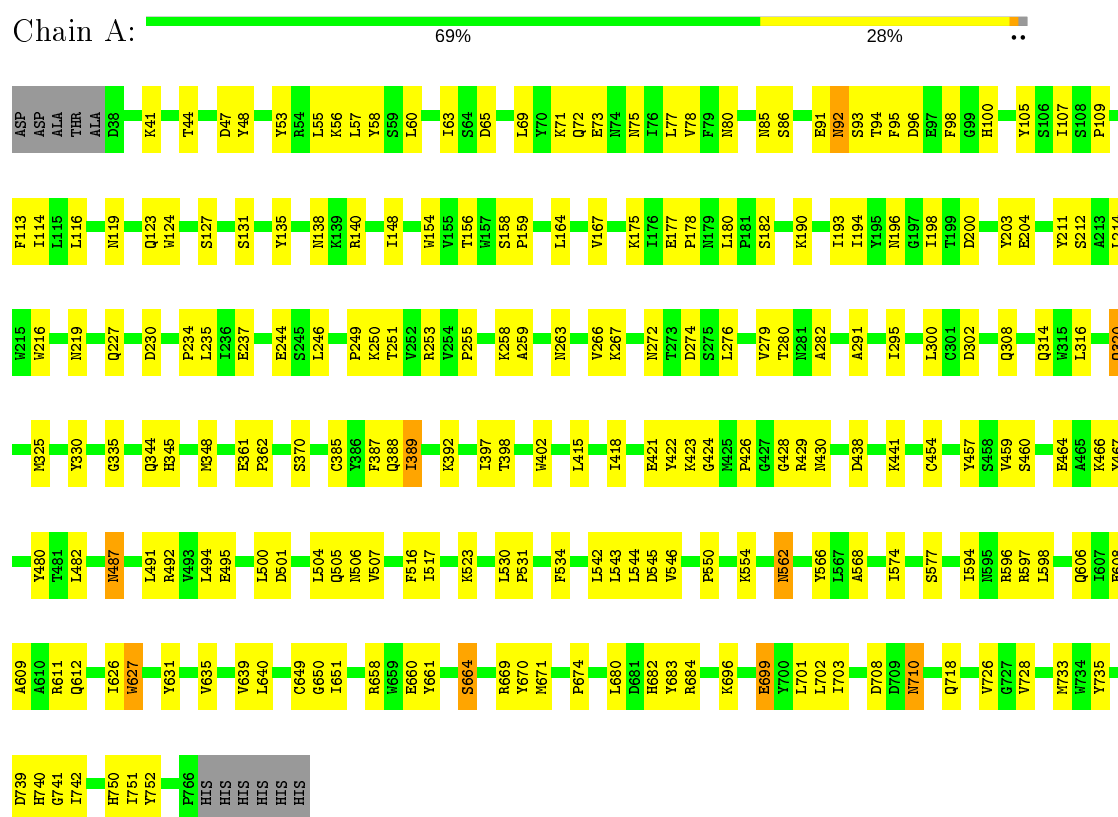
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	136	Total	O	0	0
			136	136		
3	B	137	Total	O	0	0
			137	137		

### 3 Residue-property plots

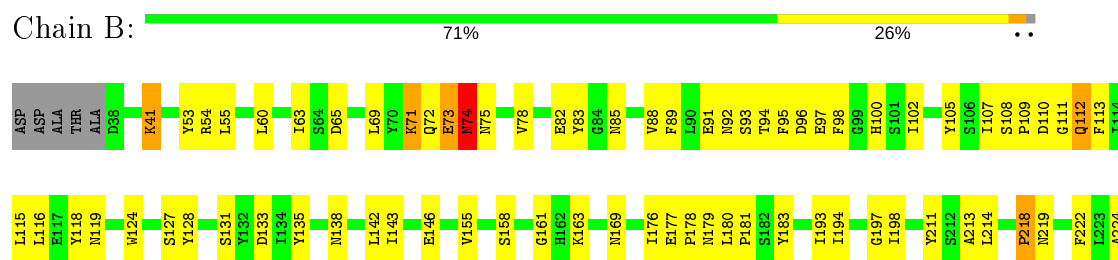
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

Note EDS was not executed.

#### • Molecule 1: Dipeptidyl peptidase IV



#### • Molecule 1: Dipeptidyl peptidase IV





## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.04Å 125.92Å 136.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.60	Depositor
% Data completeness (in resolution range)	94.4 (10.00-2.60)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNX	Depositor
R, $R_{free}$	0.249 , 0.302	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12355	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

## 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: NAG

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	0.35	0/6143	0.62	1/8355 (0.0%)
1	B	0.35	0/6143	0.62	2/8355 (0.0%)
All	All	0.35	0/12286	0.62	3/16710 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	458	SER	N-CA-C	-5.16	97.08	111.00
1	B	300	LEU	N-CA-C	-5.14	97.11	111.00
1	A	300	LEU	N-CA-C	-5.13	97.16	111.00

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	5971	0	5684	167	0
1	B	5971	0	5683	164	0
2	A	70	0	65	2	0
2	B	70	0	65	3	0
3	A	136	0	0	0	0
3	B	137	0	0	2	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	12355	0	11497	322	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 14.

The worst 5 of 322 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:219:ASN:H	1:B:308:GLN:HE22	1.07	0.99
1:B:72:GLN:HG3	1:B:73:GLU:H	1.27	0.98
1:A:219:ASN:H	1:A:308:GLN:HE22	1.11	0.93
1:B:429:ARG:HH11	1:B:429:ARG:HG3	1.39	0.87
1:A:428:GLY:O	1:A:429:ARG:HD2	1.78	0.83

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	727/740 (98%)	671 (92%)	52 (7%)	4 (1%)	25	47
1	B	727/740 (98%)	678 (93%)	41 (6%)	8 (1%)	14	30
All	All	1454/1480 (98%)	1349 (93%)	93 (6%)	12 (1%)	19	39

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	92	ASN
1	A	320	GLN
1	B	73	GLU
1	B	279	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	342	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	654/663 (99%)	642 (98%)	12 (2%)	59	80
1	B	654/663 (99%)	638 (98%)	16 (2%)	49	74
All	All	1308/1326 (99%)	1280 (98%)	28 (2%)	53	77

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	41	LYS
1	B	112	GLN
1	B	710	ASN
1	B	71	LYS
1	B	74	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	694	ASN
1	B	119	ASN
1	B	679	ASN
1	A	697	GLN
1	B	138	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

10 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	NAG	B	910	1	14,14,15	0.53	0	17,19,21	0.85	1 (5%)
2	NAG	A	905	1	14,14,15	0.59	0	17,19,21	1.33	2 (11%)
2	NAG	B	907	1	14,14,15	0.49	0	17,19,21	0.94	1 (5%)
2	NAG	A	901	1	14,14,15	0.51	0	17,19,21	0.69	0
2	NAG	B	908	1	14,14,15	0.48	0	17,19,21	0.72	1 (5%)
2	NAG	A	904	1	14,14,15	0.53	0	17,19,21	0.90	1 (5%)
2	NAG	B	909	1	14,14,15	0.54	0	17,19,21	0.56	0
2	NAG	A	903	1	14,14,15	0.44	0	17,19,21	0.72	0
2	NAG	B	906	1	14,14,15	0.47	0	17,19,21	0.67	0
2	NAG	A	902	1	14,14,15	0.49	0	17,19,21	0.71	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	B	910	1	-	2/6/23/26	0/1/1/1
2	NAG	A	905	1	-	2/6/23/26	0/1/1/1
2	NAG	B	907	1	-	3/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	901	1	-	3/6/23/26	0/1/1/1
2	NAG	B	908	1	-	2/6/23/26	0/1/1/1
2	NAG	A	904	1	-	0/6/23/26	0/1/1/1
2	NAG	B	909	1	-	4/6/23/26	0/1/1/1
2	NAG	A	903	1	-	3/6/23/26	0/1/1/1
2	NAG	B	906	1	-	3/6/23/26	0/1/1/1
2	NAG	A	902	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	905	NAG	C4-C3-C2	-3.59	105.75	111.02
2	B	907	NAG	C2-N2-C7	-2.65	119.13	122.90
2	A	905	NAG	C2-N2-C7	-2.61	119.19	122.90
2	B	910	NAG	C2-N2-C7	-2.34	119.58	122.90
2	B	908	NAG	C2-N2-C7	-2.15	119.84	122.90

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	910	NAG	C8-C7-N2-C2
2	B	910	NAG	O7-C7-N2-C2
2	A	901	NAG	C8-C7-N2-C2
2	A	901	NAG	O7-C7-N2-C2
2	B	908	NAG	C8-C7-N2-C2

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	910	NAG	1	0
2	A	905	NAG	1	0
2	B	909	NAG	2	0
2	A	902	NAG	1	0

## 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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.