



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:10 PM GMT

PDB ID : 1NU6  
Title : Crystal structure of human Dipeptidyl Peptidase IV (DPP-IV)  
Authors : Hennig, M.; Stihle, M.; Thoma, R.; Ruf, A.  
Deposited on : 2003-01-31  
Resolution : 2.10 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

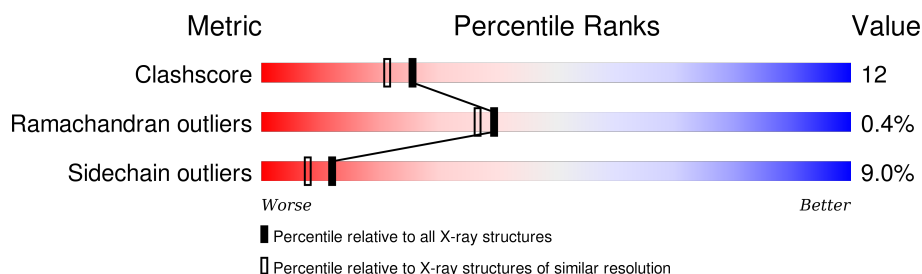
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	728	 72% 22% 5%
1	B	728	 70% 24% 5%

## 2 Entry composition [i](#)

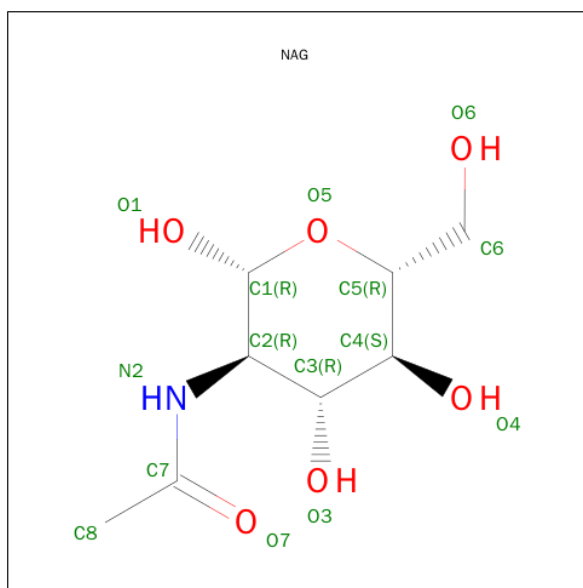
There are 5 unique types of molecules in this entry. The entry contains 12366 atoms, of which 0 are hydrogens and 0 are deuteriums.

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	728	Total	C	N	O	S	0	0	0
			5963	3827	982	1128	26			
1	B	728	Total	C	N	O	S	0	0	0
			5963	3827	982	1128	26			

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



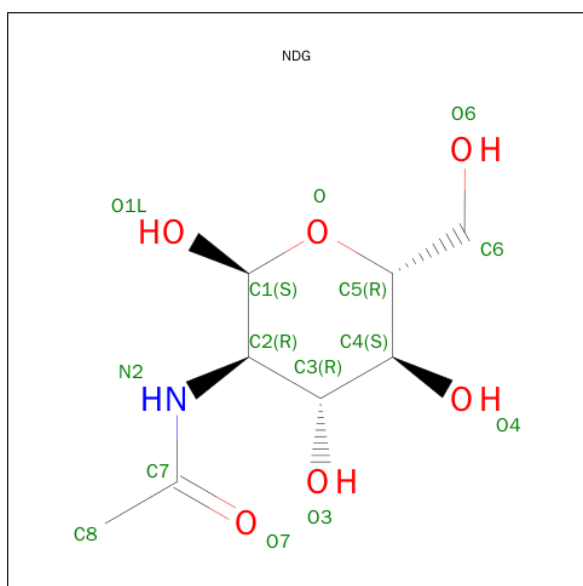
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
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 SUGAR (2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE) (three-letter code: NDG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 4 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	3	Total	Hg	0	0
			3	3		
4	A	3	Total	Hg	0	0
			3	3		

- Molecule 5 is water.

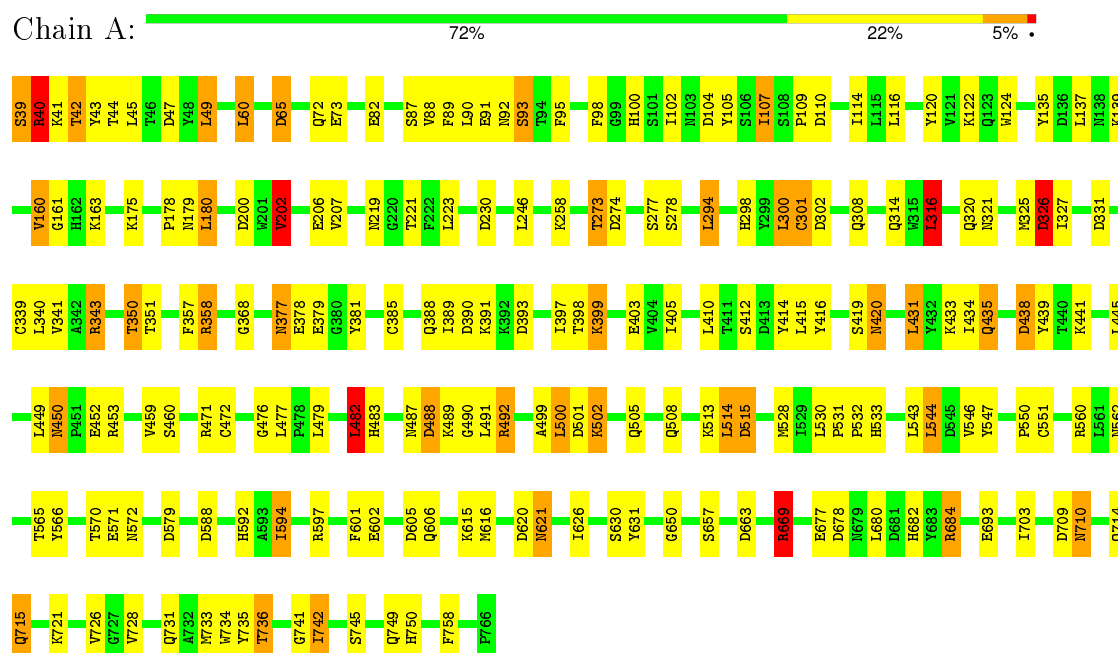
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	152	Total 152	O 152	0	0
5	B	170	Total 170	O 170	0	0

### 3 Residue-property plots

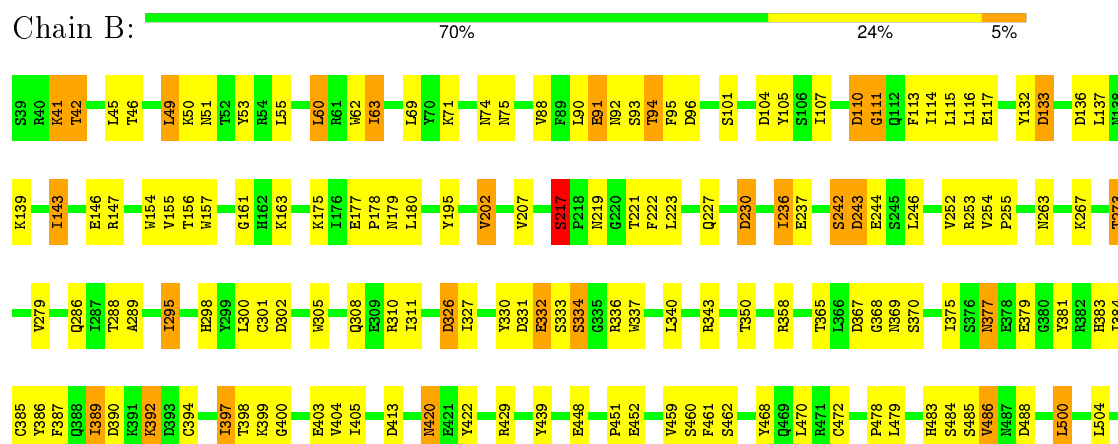
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

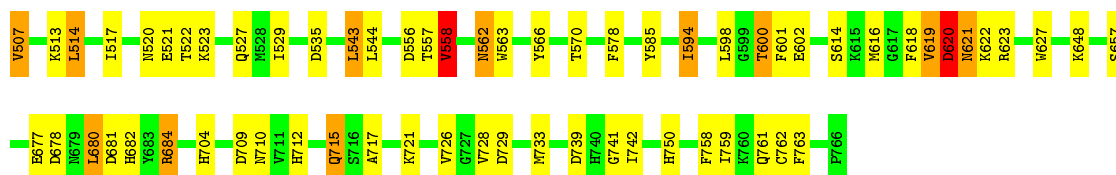
Note EDS was not executed.

#### • Molecule 1: Dipeptidyl peptidase IV



#### • Molecule 1: Dipeptidyl peptidase IV





## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.50 Å 68.24 Å 419.29 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 2.10	Depositor
% Data completeness (in resolution range)	83.0 (12.00-2.10)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5	Depositor
R, $R_{free}$	0.215 , 0.266	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12366	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	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, NDG, HG

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	0.84	2/6135 (0.0%)	1.00	29/8344 (0.3%)
1	B	0.86	4/6135 (0.1%)	1.01	26/8344 (0.3%)
All	All	0.85	6/12270 (0.0%)	1.00	55/16688 (0.3%)

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	1	0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	301	CYS	CB-SG	9.21	1.98	1.82
1	B	301	CYS	CB-SG	8.79	1.97	1.82
1	B	558	VAL	CB-CG1	-5.87	1.40	1.52
1	A	301	CYS	CA-CB	5.63	1.66	1.53
1	B	721	LYS	CD-CE	5.18	1.64	1.51

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	110	ASP	CB-CG-OD2	9.34	126.70	118.30
1	B	620	ASP	CB-CG-OD2	7.71	125.23	118.30
1	A	620	ASP	CB-CG-OD2	7.26	124.84	118.30
1	B	111	GLY	N-CA-C	7.26	131.24	113.10
1	A	202	VAL	CB-CA-C	-6.78	98.53	111.40

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	40	ARG	CA

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	5963	0	5678	133	1
1	B	5963	0	5679	158	0
2	A	56	0	52	2	0
2	B	42	0	39	3	0
3	B	14	0	13	0	0
4	A	3	0	0	0	0
4	B	3	0	0	0	0
5	A	152	0	0	7	0
5	B	170	0	0	4	0
All	All	12366	0	11461	290	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 12.

The worst 5 of 290 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:160:VAL:HG12	1:A:160:VAL:O	1.59	1.01
1:B:620:ASP:O	1:B:620:ASP:OD1	1.91	0.88
1:B:543:LEU:HD23	1:B:544:LEU:N	1.91	0.86
1:B:267:LYS:HD2	1:B:286:GLN:NE2	1.90	0.85
1:B:600:THR:CG2	1:B:601:PHE:H	1.90	0.84

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:A:41:LYS:NZ	1:A:341:VAL:O[1_655]	2.12	0.08

## 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	726/728 (100%)	683 (94%)	39 (5%)	4 (1%)	30	24
1	B	726/728 (100%)	683 (94%)	41 (6%)	2 (0%)	46	45
All	All	1452/1456 (100%)	1366 (94%)	80 (6%)	6 (0%)	39	37

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	488	ASP
1	B	74	ASN
1	B	111	GLY
1	A	393	ASP
1	A	160	VAL

### 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	653/653 (100%)	597 (91%)	56 (9%)	13	9
1	B	653/653 (100%)	591 (90%)	62 (10%)	11	7
All	All	1306/1306 (100%)	1188 (91%)	118 (9%)	12	8

5 of 118 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	710	ASN
1	B	93	SER
1	B	620	ASP
1	A	715	GLN
1	B	51	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	715	GLN
1	B	66	HIS
1	B	710	ASN
1	A	731	GLN
1	A	748	HIS

### 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 14 ligands modelled in this entry, 6 are monoatomic - leaving 8 for Mogul analysis.

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	A	793	1	14,14,15	0.59	0	15,19,21	1.61	3 (20%)
2	NAG	A	794	1	14,14,15	0.69	0	15,19,21	1.40	1 (6%)
2	NAG	A	795	1	14,14,15	0.55	0	15,19,21	1.31	2 (13%)
2	NAG	A	796	1	14,14,15	0.69	0	15,19,21	1.50	3 (20%)
2	NAG	B	793	1	14,14,15	0.41	0	15,19,21	2.34	5 (33%)
2	NAG	B	794	1	14,14,15	0.57	0	15,19,21	1.37	2 (13%)
2	NAG	B	796	1	14,14,15	0.93	1 (7%)	15,19,21	1.35	2 (13%)
3	NDG	B	797	1	14,14,15	0.65	0	15,19,21	1.92	4 (26%)

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	A	793	1	-	0/6/23/26	0/1/1/1
2	NAG	A	794	1	-	0/6/23/26	0/1/1/1
2	NAG	A	795	1	-	0/6/23/26	0/1/1/1
2	NAG	A	796	1	-	0/6/23/26	0/1/1/1
2	NAG	B	793	1	-	0/6/23/26	0/1/1/1
2	NAG	B	794	1	-	0/6/23/26	0/1/1/1
2	NAG	B	796	1	-	0/6/23/26	0/1/1/1
3	NDG	B	797	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	796	NAG	O5-C1	-2.91	1.38	1.43

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	793	NAG	C4-C3-C2	-4.08	104.89	111.23
3	B	797	NDG	C3-C4-C5	-3.71	103.73	110.20
2	A	795	NAG	C4-C3-C2	-3.34	106.04	111.23
2	B	796	NAG	O6-C6-C5	-3.03	101.33	111.33
3	B	797	NDG	C2-N2-C7	-2.99	119.20	123.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	793	NAG	2	0
2	B	793	NAG	3	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 will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.