



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:15 AM GMT

PDB ID : 2GBF
Title : rat dpp-IV with alkynyl cyanopyrrolidine #1
Authors : Longenecker, K.L.; Jakob, C.G.; Fry, E.H.; Wilk, S.
Deposited on : 2006-03-10
Resolution : 3.10 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
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) : 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

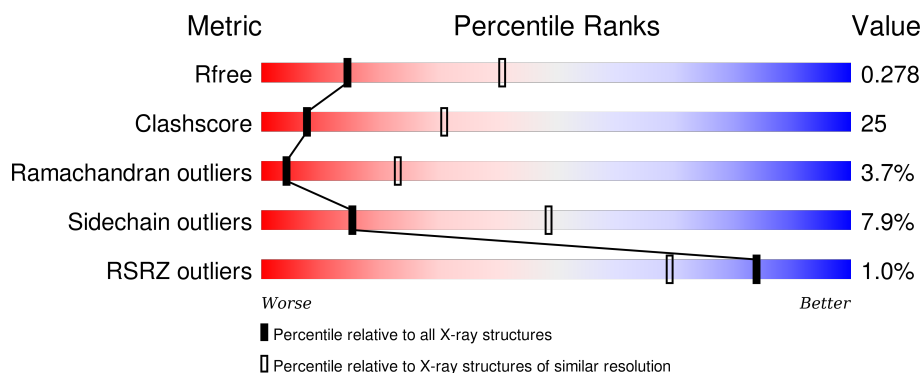
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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))
R_{free}	91344	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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 $\leq 5\%$. 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	A	730	<div> <div></div> <div>54%</div> <div>40%</div> <div>6%</div> </div>
1	B	730	<div> <div></div> <div>54%</div> <div>39%</div> <div>6%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	AIA	A	768	-	-	-	X

2 Entry composition [i](#)

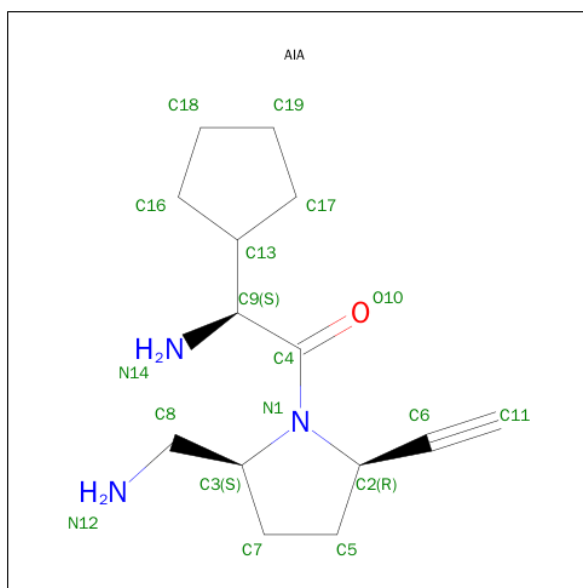
There are 2 unique types of molecules in this entry. The entry contains 11858 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	730	Total	C	N	O	S	0	0	0
			5920	3789	981	1124	26			
1	B	730	Total	C	N	O	S	0	0	0
			5920	3789	981	1124	26			

- Molecule 2 is (1S)-2-[(2S,5R)-2-(AMINOMETHYL)-5-ETHYNYLPYRROLIDIN-1-YL]-1-CYCLOPENTYL-2-OXOETHANAMINE (three-letter code: AIA) (formula: $C_{14}H_{23}N_3O$).

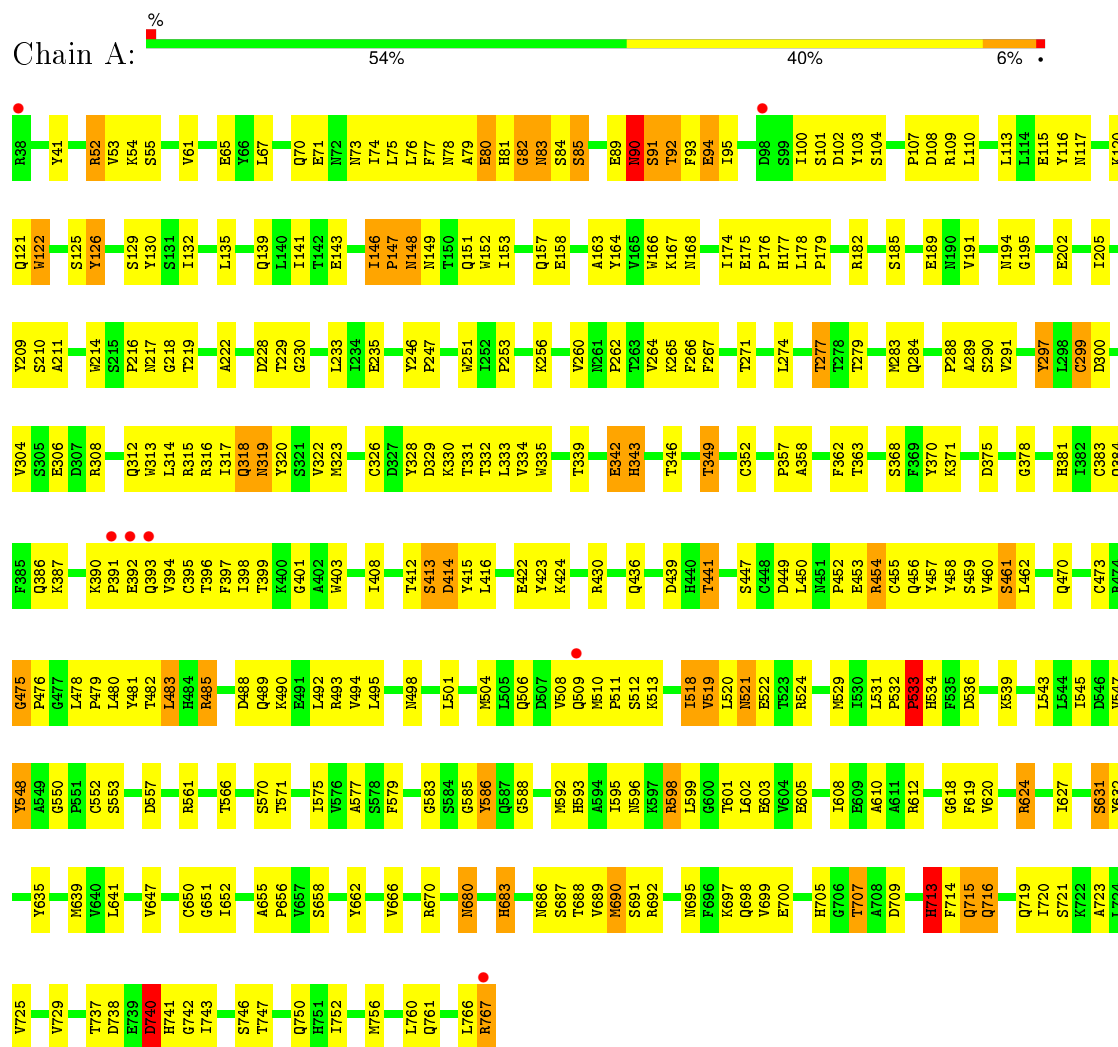


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			18	14	3	1		

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 ($\text{RSRZ} > 2$). 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: Dipeptidyl peptidase 4



Q719	K623	I544	Q384	L298	G195	Y116
I720	R624	I545	F385	C299	E202	N117
S721	D546	D546	K386	D300		K120
K722	V547	V547	G475	K387		Q121
A723	I627	Y548	P476	K390	I205	W122
		A549		P391	Y209	S125
D726	Y632	G550	P391	E306	S210	Y126
	Y635	F551	E392	D307	A211	
V729	Y639	S553	Q393	R308		S129
D738	M639	S553	T482	Q312	W214	S215
E739	V640	D557	L483	C395	S215	Y130
D740	L641	R561	H484	T396	P216	S131
H741		R561	R485	F397	N217	I132
G742	V647	T566	S486	I398		L135
I743	G650	T566	T487	T399		
	G651	S570	K400	K400	T219	
I752	I652	T571	Q401	Q318	A222	Q139
	A655	I575	A402	N319		
M756		V576	W403	Y320		T142
Q761	S658	A577	I408	S321	D228	E143
L766	R659	S578	T412	K323	G230	E144
K767	Y662	F579	D414	C326	L233	K145
	D663		Y415	D327	I234	I146
	D664		L416	Y328	E235	P147
	S665	G583	L501	D329		N149
	V666	G585	M504	K330	W251	T150
	R670	Y586	L505	T331	I251	Q151
		G587	Q506	T332	P253	W152
		G588	D507	L333		I153
	M680	M592	V508	V334	K256	
	H683	H593	Q509	W335		Q157
	M686	A594	P510	T339	V260	E158
S687	T688	I595	P511	E342	N261	K161
V689	M690	I596	D439	H243	P262	L162
S691	G600	R597	H440		V264	A163
R692	T601	L599	T441	T346	K265	Y164
K697	L602	G600	K444	T349	F266	V165
Q698	V604	N521	S447		I268	W166
V699	E603	E522	C448	P357		K167
E700	E605	R524	D449	A358	T271	N168
	I608	Y527	L450	T363		I174
H705	E609	Q528	N451		L274	E175
G706	A610	M529	P452	S368	T277	P176
T707	A611	P532	E453	F369	L178	H177
A708	R612	H534	R454	I370	T279	P179
D709	L615	F535	C455	K371		R182
	K616	D536	Q456	D375	M283	
H713	H617	D536	Y457		Q284	S185
F714	G618	D536	Y458	G378	I285	
Q715	F619	K539	S459	V460	A289	E189
	Q716	L462	S461	V460	S290	H190
		L543	L462	H381	V291	Y191
			Q470	I382		
				C383		N194

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α , β , γ	208.17Å 208.17Å 208.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.10 19.94 – 3.11	Depositor EDS
% Data completeness (in resolution range)	98.8 (20.00-3.10) 99.3 (19.94-3.11)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.57 (at 3.09Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.250 , 0.289 0.243 , 0.278	Depositor DCC
R_{free} test set	2743 reflections (5.10%)	DCC
Wilson B-factor (Å ²)	73.1	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 59.3	EDS
Estimated twinning fraction	0.000 for l,-k,h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.54$, $\langle L^2 \rangle = 0.38$	Xtriage
Outliers	1 of 53760 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	11858	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.78% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: AIA

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	0.46	2/6088 (0.0%)	0.71	2/8278 (0.0%)
1	B	0.43	0/6088	0.70	2/8278 (0.0%)
All	All	0.45	2/12176 (0.0%)	0.71	4/16556 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	631	SER	C-O	8.88	1.40	1.23
1	A	352	CYS	CB-SG	-5.14	1.73	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	713	HIS	N-CA-C	7.29	130.67	111.00
1	A	713	HIS	N-CA-C	7.22	130.50	111.00
1	A	90	ASN	N-CA-C	-5.15	97.10	111.00
1	B	90	ASN	N-CA-C	-5.12	97.18	111.00

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	A	5920	0	5631	294	0
1	B	5920	0	5632	292	0
2	A	18	0	21	0	0
All	All	11858	0	11284	584	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 25.

The worst 5 of 584 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:761:GLN:HB3	1:A:767:ARG:HB3	1.39	1.03
1:B:349:THR:HB	1:B:593:HIS:HD2	1.26	0.99
1:A:460:VAL:HG22	1:A:461:SER:H	1.28	0.99
1:A:349:THR:HB	1:A:593:HIS:HD2	1.25	0.98
1:B:761:GLN:HB3	1:B:767:ARG:HB3	1.42	0.98

There are no symmetry-related clashes.

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	728/730 (100%)	614 (84%)	87 (12%)	27 (4%)	4	23
1	B	728/730 (100%)	615 (84%)	86 (12%)	27 (4%)	4	23
All	All	1456/1460 (100%)	1229 (84%)	173 (12%)	54 (4%)	4	23

5 of 54 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	80	GLU
1	A	143	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	450	LEU
1	A	521	ASN
1	A	533	PRO

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	643/651 (99%)	594 (92%)	49 (8%)	16	51
1	B	643/651 (99%)	591 (92%)	52 (8%)	15	47
All	All	1286/1302 (99%)	1185 (92%)	101 (8%)	15	49

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	715	GLN
1	B	149	ASN
1	B	686	ASN
1	A	740	ASP
1	B	94	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	719	GLN
1	B	284	GLN
1	B	715	GLN
1	A	749	HIS
1	B	148	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

1 ligand is 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	AIA	A	768	1	17,19,19	1.40	2 (11%)	15,26,26	1.36	2 (13%)

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	AIA	A	768	1	-	0/14/36/36	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	768	AIA	C8-N12	-4.70	1.35	1.47
2	A	768	AIA	C4-N1	2.01	1.39	1.35

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	768	AIA	C7-C3-N1	2.27	104.32	101.93
2	A	768	AIA	C13-C9-C4	2.66	116.56	110.86

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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	730/730 (100%)	-0.35	7 (0%) 84 69	32, 58, 91, 165	0
1	B	730/730 (100%)	-0.23	7 (0%) 84 69	35, 65, 104, 161	0
All	All	1460/1460 (100%)	-0.29	14 (0%) 84 69	32, 61, 99, 165	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	767	ARG	7.1
1	A	393	GLN	3.1
1	B	392	GLU	3.0
1	B	393	GLN	2.8
1	A	98	ASP	2.6

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, 95th 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	AIA	A	768	18/18	0.92	0.29	3.15	65,65,65,65	0

6.5 Other polymers ⓘ

There are no such residues in this entry.