W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 03:54 PM GMT

PDB ID : 4DTC
Title : Crystal Structure of DPP-IV with Compound C5
Authors : Xiong, B.; Zhu, L.R.; Chen, D.Q.; Zhao, Y.L.; Jiang, F.; Shen, J.K.
Deposited on : 2012-02-21
Resolution : 3.00 Å(reported)

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 3.00 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.
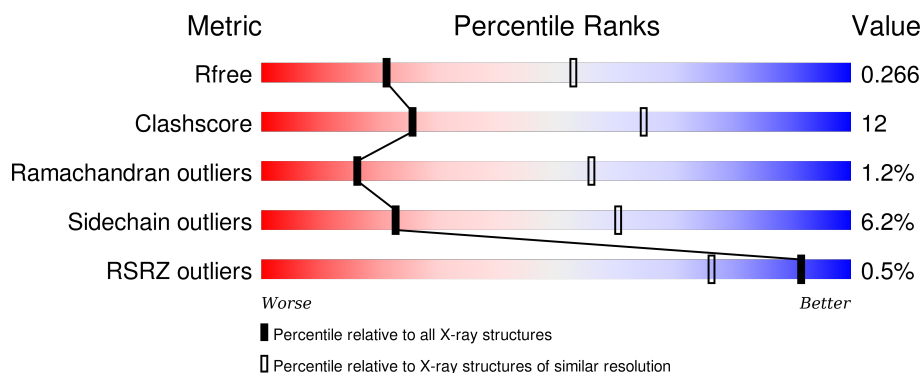
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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	728	67% ... 31% ... •
1	B	728	% ... 67% ... 30% ... •

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	D5C	A	801	-	-	-	X

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11992 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	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 N-[(3R)-3-AMINO-4-(2,4,5-TRIFLUOROPHENYL)BUTYL]-3'-(TRIFLUOROMETHYL)BIPHENYL-4-CARBOXAMIDE (three-letter code: D5C) (formula: $C_{24}H_{20}F_6N_2O$).
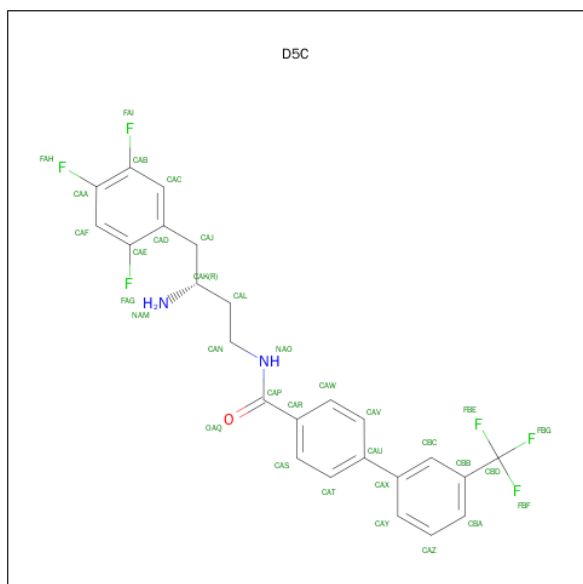
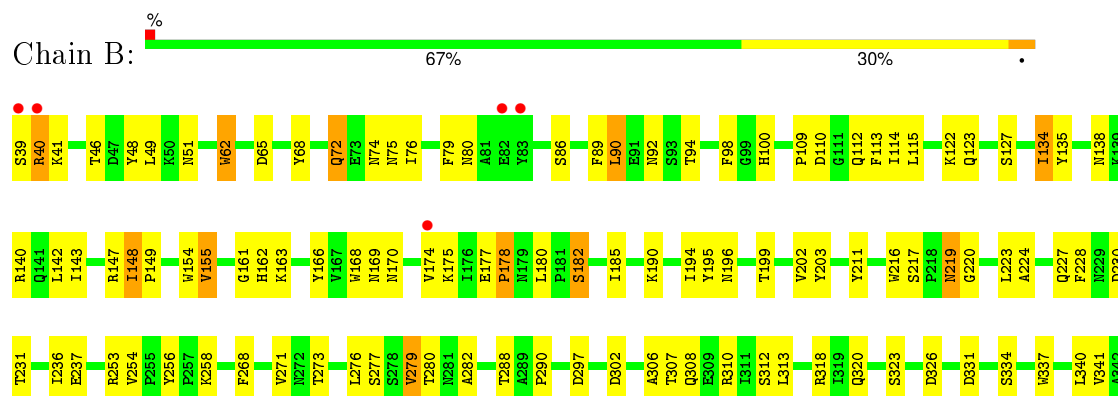| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			33	24	6	2	1		
2	B	1	Total	C	F	N	O	0	0
			33	24	6	2	1		

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 (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

- Molecule 1: Dipeptidyl peptidase 4

R343 E347 T350 T351 G352 W353 R358 E361 F364 S370 F371 E378 R382 H383 I384 C385 Y386 I389 T398 W402 E403 V404 Y417 I418 S419 N420 E421 Y422 K423 P426 N430 L431 Y432 K433 I434 Q435 D438 K441 V442 T443 C444 L445 L449

N450 P451 C454 Y457 S458 V459 K463 E464 A465 K466 L477 Y480 T481 L482 H483 N487 D488 L494 L500 D501 N506 D515 Y526 P532 L542 D545 V546 Y547 A548 G549 S552 Q553 K554 N562 Y566 S569 F578 D579 G580 S583

G584 Y585 Q586 G587 I594 N595 R596 E602 V603 A609 F618 I624 A625 I626 W627 Y631 Y634 V635 T636 S637 M638 V639 L640 S644 G645 V646 C649 G650 A654 W659 E660 Y661 Y662 V665 R669 T675 D678 N679 L680 Y683 R684 N685 R691

E699 I703 H704 H712 K721 V724 D725 V726 G727 V728 M733 W734 Y735 G741 I742 H750 I751 K760 S764 L765 P766

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	80.47Å　80.47Å　290.90Å 90.00°　90.00°　120.00°	Depositor
Resolution (Å)	29.89　–　3.00 29.89　–　3.00	Depositor EDS
% Data completeness (in resolution range)	98.3 (29.89-3.00) 98.4 (29.89-3.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.83 (at 3.00Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.206　,　0.268 0.202　,　0.266	Depositor DCC
R_{free} test set	2126 reflections (5.40%)	DCC
Wilson B-factor (Å^2)	88.4	Xtriage
Anisotropy	0.176	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 52.9	EDS
Estimated twinning fraction	0.001 for -h,-k,l 0.099 for h,-h-k,-l 0.028 for -k,-h,-l	Xtriage
L-test for twinning [2]	$<\|L\|> = 0.50$, $<L^2> = 0.33$	Xtriage
Outliers	0 of 41436 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11992	wwPDB-VP
Average B, all atoms (Å^2)	107.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% 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: D5C

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	0.66	7/6135 (0.1%)	0.74	2/8344 (0.0%)
1	B	0.64	6/6135 (0.1%)	0.73	2/8344 (0.0%)
All	All	0.65	13/12270 (0.1%)	0.74	4/16688 (0.0%)

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	337	TRP	CD2-CE2	5.96	1.48	1.41
1	B	659	TRP	CD2-CE2	5.41	1.47	1.41
1	A	337	TRP	CD2-CE2	5.38	1.47	1.41
1	B	154	TRP	CD2-CE2	5.18	1.47	1.41
1	B	353	TRP	CD2-CE2	5.17	1.47	1.41

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	358	ARG	NE-CZ-NH2	8.11	124.35	120.30
1	B	358	ARG	NE-CZ-NH1	-7.57	116.52	120.30
1	A	358	ARG	NE-CZ-NH1	7.45	124.02	120.30
1	A	358	ARG	NE-CZ-NH2	-7.36	116.62	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	5963	0	5685	149	0
1	B	5963	0	5685	149	0
2	A	33	0	20	4	0
2	B	33	0	20	2	0
All	All	11992	0	11410	292	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 12.

The worst 5 of 292 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:74:ASN:HB2	1:A:92:ASN:HB2	1.55	0.87
1:B:74:ASN:HB2	1:B:92:ASN:HB2	1.56	0.87
1:B:352:GLY:HA2	1:B:595:ASN:ND2	1.95	0.81
1:B:352:GLY:HA2	1:B:595:ASN:HD22	1.45	0.80
1:A:420:ASN:HD22	1:A:426:PRO:HA	1.48	0.78

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	726/728 (100%)	642 (88%)	75 (10%)	9 (1%)	16	56
1	B	726/728 (100%)	650 (90%)	68 (9%)	8 (1%)	17	58
All	All	1452/1456 (100%)	1292 (89%)	143 (10%)	17 (1%)	16	56

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	140	ARG
1	B	140	ARG
1	A	178	PRO
1	A	279	VAL
1	A	645	GLY

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%)	614 (94%)	39 (6%)	24	62
1	B	653/653 (100%)	611 (94%)	42 (6%)	22	59
All	All	1306/1306 (100%)	1225 (94%)	81 (6%)	23	60

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	685	ASN
1	B	94	THR
1	B	545	ASP
1	A	751	ILE
1	B	46	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	ASN
1	B	74	ASN
1	B	750	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)

2 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	D5C	A	801	-	35,35,35	2.20	9 (25%)	45,50,50	1.39	5 (11%)
2	D5C	B	801	-	35,35,35	2.17	9 (25%)	45,50,50	1.18	6 (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	D5C	A	801	-	-	0/24/24/24	0/3/3/3
2	D5C	B	801	-	-	0/24/24/24	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	D5C	CAJ-CAD	-7.29	1.41	1.51
2	B	801	D5C	CAJ-CAD	-7.11	1.41	1.51
2	A	801	D5C	FAG-CAE	-5.56	1.22	1.35
2	A	801	D5C	CAR-CAP	-4.09	1.41	1.50
2	B	801	D5C	CBD-CBB	-3.93	1.41	1.49

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	801	D5C	CAT-CAU-CAX	-3.07	115.88	121.39
2	A	801	D5C	CAF-CAA-CAB	-3.05	118.37	121.13
2	B	801	D5C	CAN-CAL-CAK	-2.94	108.61	114.68
2	A	801	D5C	CAS-CAR-CAP	-2.38	113.06	120.60
2	A	801	D5C	CAF-CAE-CAD	-2.31	121.19	124.13

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801	D5C	4	0
2	B	801	D5C	2	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 (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	728/728 (100%)	-0.21	2 (0%)	94	84	60, 102, 151, 200	0
1	B	728/728 (100%)	-0.22	5 (0%)	89	70	57, 106, 156, 228	0
All	All	1456/1456 (100%)	-0.21	7 (0%)	91	76	57, 104, 155, 228	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	83	TYR	6.5
1	B	82	GLU	3.0
1	B	40	ARG	2.9
1	A	39	SER	2.9
1	B	174	VAL	2.2

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(Å^2)	Q<0.9
2	D5C	A	801	33/33	0.92	0.30	2.12	74,105,139,143	0
2	D5C	B	801	33/33	0.93	0.24	0.35	80,110,147,151	0

6.5   Other polymers (i)

There are no such residues in this entry.