wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 11:21 PM EDT

PDB ID	:	1HDU
Title	:	Crystal structure of bovine pancreatic carboxypeptidase A complexed with a minocarbonylphenylalanine at 1.75 A
Authors	:	Cho, J.H.; Ha, N.-C.; Chung, S.J.; Kim, D.H.; Choi, K.Y.; Oh, B.-H.
Deposited on	:	2000-11-17
Resolution	:	1.75 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

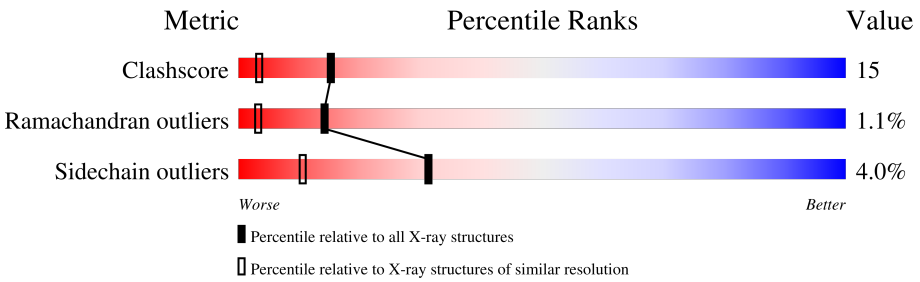
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 1.75 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	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	307	77% 20% •
1	B	307	76% 23% •
1	D	307	75% 23% •
1	E	307	73% 25% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10399 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 CARBOXYPEPTIDASE A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	307	2436	1561	406	464	5	0	0	0
1	B	307	2436	1561	406	464	5	0	0	0
1	D	307	2436	1561	406	464	5	0	0	0
1	E	307	2436	1561	406	464	5	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	GLN	GLU	conflict	UNP P00730
A	31	GLU	GLN	conflict	UNP P00730
A	89	ASN	ASP	conflict	UNP P00730
A	93	ASN	ASP	conflict	UNP P00730
A	114	ASN	ASP	conflict	UNP P00730
A	122	GLU	GLN	conflict	UNP P00730
A	185	ASN	ASP	conflict	UNP P00730
A	228	ALA	GLU	conflict	UNP P00730
A	305	VAL	LEU	conflict	UNP P00730
B	28	GLN	GLU	conflict	UNP P00730
B	31	GLU	GLN	conflict	UNP P00730
B	89	ASN	ASP	conflict	UNP P00730
B	93	ASN	ASP	conflict	UNP P00730
B	114	ASN	ASP	conflict	UNP P00730
B	122	GLU	GLN	conflict	UNP P00730
B	185	ASN	ASP	conflict	UNP P00730
B	228	ALA	GLU	conflict	UNP P00730
B	305	VAL	LEU	conflict	UNP P00730
D	28	GLN	GLU	conflict	UNP P00730
D	31	GLU	GLN	conflict	UNP P00730
D	89	ASN	ASP	conflict	UNP P00730

Continued on next page...

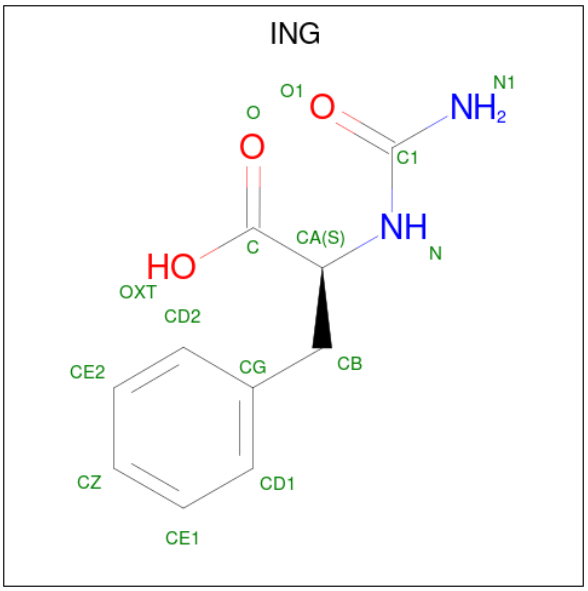
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	93	ASN	ASP	conflict	UNP P00730
D	114	ASN	ASP	conflict	UNP P00730
D	122	GLU	GLN	conflict	UNP P00730
D	185	ASN	ASP	conflict	UNP P00730
D	228	ALA	GLU	conflict	UNP P00730
D	305	VAL	LEU	conflict	UNP P00730
E	28	GLN	GLU	conflict	UNP P00730
E	31	GLU	GLN	conflict	UNP P00730
E	89	ASN	ASP	conflict	UNP P00730
E	93	ASN	ASP	conflict	UNP P00730
E	114	ASN	ASP	conflict	UNP P00730
E	122	GLU	GLN	conflict	UNP P00730
E	185	ASN	ASP	conflict	UNP P00730
E	228	ALA	GLU	conflict	UNP P00730
E	305	VAL	LEU	conflict	UNP P00730

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	1	1	1	0	0
2	B	1	1	1	0	0
2	D	1	1	1	0	0
2	E	1	1	1	0	0

- Molecule 3 is D-[(AMINO)CARBONYL]PHENYLALANINE (three-letter code: ING) (formula: $C_{10}H_{12}N_2O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	15	10	2	3	0	0
			Total	C	N	O		
3	B	1	15	10	2	3	0	0
			Total	C	N	O		
3	D	1	15	10	2	3	0	0
			Total	C	N	O		
3	E	1	15	10	2	3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	147	147	147	0	0
			Total	O		
4	B	144	144	144	0	0
			Total	O		
4	D	149	149	149	0	0
			Total	O		
4	E	151	151	151	0	0

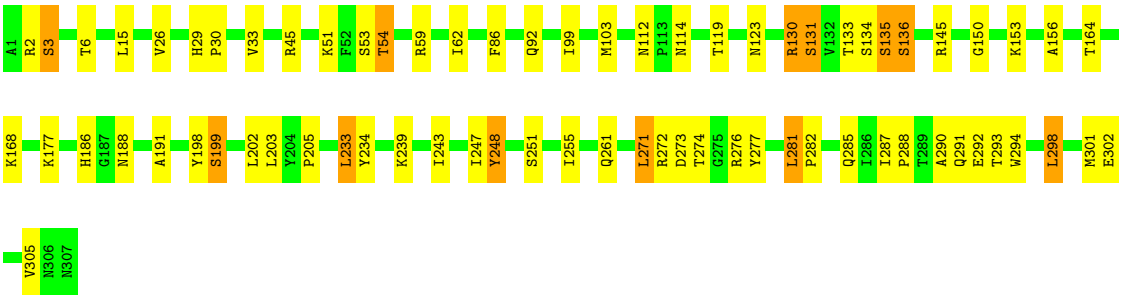
3 Residue-property plots (i)

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.

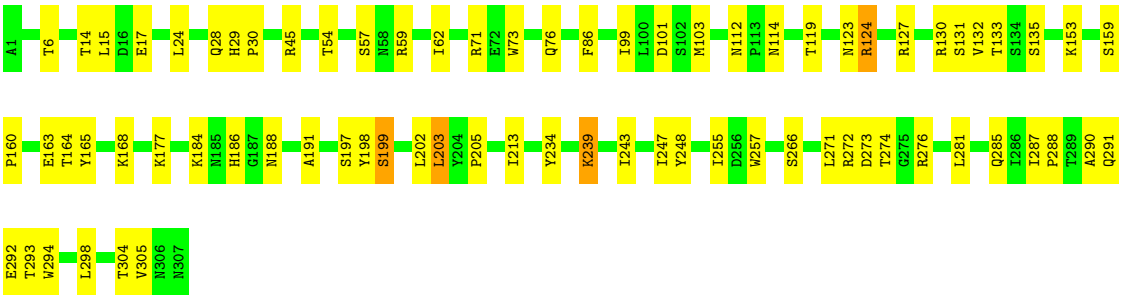
• Molecule 1: CARBOXYPEPTIDASE A

Chain A: 77% 20% •



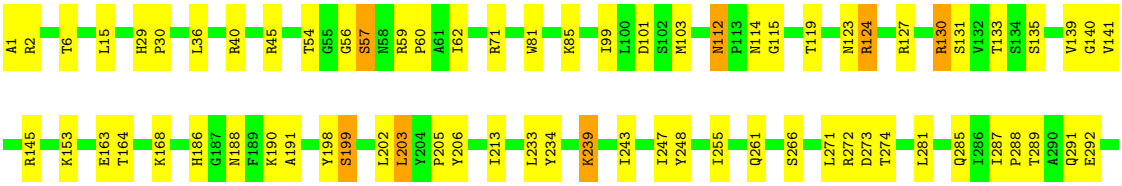
• Molecule 1: CARBOXYPEPTIDASE A

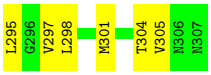
Chain B: 76% 23% •



• Molecule 1: CARBOXYPEPTIDASE A

Chain D: 75% 23% •





• Molecule 1: CARBOXYPEPTIDASE A

Chain E:

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	65.57Å 60.52Å 74.41Å 90.00° 97.84° 90.00°	Depositor
Resolution (Å)	100.00 – 1.75	Depositor
% Data completeness (in resolution range)	81.6 (100.00-1.75)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.198 , 0.229	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10399	wwPDB-VP
Average B, all atoms (Å^2)	15.0	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: ZN, ING

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	0.33	0/2501	0.59	0/3399
1	B	0.32	0/2501	0.60	0/3399
1	D	0.32	0/2501	0.59	0/3399
1	E	0.33	0/2501	0.59	0/3399
All	All	0.33	0/10004	0.59	0/13596

There are no bond length outliers.

There are no bond angle outliers.

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	2436	0	2351	67	0
1	B	2436	0	2351	70	0
1	D	2436	0	2351	72	0
1	E	2436	0	2351	81	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
3	A	15	0	11	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	15	0	11	0	0
3	D	15	0	11	0	0
3	E	15	0	11	0	0
4	A	147	0	0	9	0
4	B	144	0	0	12	0
4	D	149	0	0	14	0
4	E	151	0	0	13	0
All	All	10399	0	9448	288	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 15.

The worst 5 of 288 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:131:SER:HB2	4:B:2087:HOH:O	1.49	1.12
1:E:131:SER:HB2	4:E:2094:HOH:O	1.50	1.11
1:D:131:SER:HB2	4:D:2091:HOH:O	1.51	1.07
1:B:273:ASP:C	1:B:274:THR:N	2.11	1.04
1:A:54:THR:HG22	1:A:59:ARG:HH12	1.18	1.03

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	303/307 (99%)	287 (95%)	11 (4%)	5 (2%)	9 1
1	B	303/307 (99%)	292 (96%)	10 (3%)	1 (0%)	41 22
1	D	303/307 (99%)	291 (96%)	9 (3%)	3 (1%)	15 4
1	E	303/307 (99%)	287 (95%)	12 (4%)	4 (1%)	12 2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1212/1228 (99%)	1157 (96%)	42 (4%)	13 (1%)	14 3

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	SER
1	A	199	SER
1	B	199	SER
1	D	199	SER
1	E	3	SER

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	263/263 (100%)	250 (95%)	13 (5%)	25 7
1	B	263/263 (100%)	256 (97%)	7 (3%)	44 22
1	D	263/263 (100%)	253 (96%)	10 (4%)	33 11
1	E	263/263 (100%)	251 (95%)	12 (5%)	27 8
All	All	1052/1052 (100%)	1010 (96%)	42 (4%)	31 10

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	281	LEU
1	E	233	LEU
1	D	298	LEU
1	E	130	ARG
1	E	248	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	249	GLN
1	E	171	ASN
1	D	261	GLN
1	E	37	GLN
1	E	249	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
3	ING	E	1309	2	15,15,15	1.32	0	18,19,19	0.60	1 (5%)
3	ING	A	1309	2	15,15,15	1.30	0	18,19,19	0.61	1 (5%)
3	ING	D	1309	2	15,15,15	1.30	0	18,19,19	0.62	1 (5%)
3	ING	B	1309	2	15,15,15	1.34	1 (6%)	18,19,19	0.63	1 (5%)

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
3	ING	E	1309	2	-	0/12/12/12	0/1/1/1
3	ING	A	1309	2	-	0/12/12/12	0/1/1/1
3	ING	D	1309	2	-	0/12/12/12	0/1/1/1
3	ING	B	1309	2	-	0/12/12/12	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1309	ING	CD2-CG	2.12	1.43	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1309	ING	O1-C1-N1	-2.22	119.40	123.22
3	D	1309	ING	O1-C1-N1	-2.13	119.56	123.22
3	E	1309	ING	O1-C1-N1	-2.11	119.59	123.22
3	A	1309	ING	O1-C1-N1	-2.06	119.69	123.22

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	E	1
1	A	1
1	D	1
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	273:ASP	C	274:THR	N	2.17
1	A	273:ASP	C	274:THR	N	2.16
1	D	273:ASP	C	274:THR	N	2.12
1	B	273:ASP	C	274:THR	N	2.11

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.