wwPDB X-ray Structure Validation Summary Report (i)

Aug 6, 2023 – 11:39 PM EDT

PDB ID	:	1KXV
Title	:	Camelid VHH Domains in Complex with Porcine Pancreatic alpha-Amylase
Authors	:	Desmyter, A.; Spinelli, S.; Payan, F.; Lauwereys, M.; Wyns, L.; Muyldermans, S.; Cambillau, C.
Deposited on	:	2002-02-01
Resolution	:	1.60 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*
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:

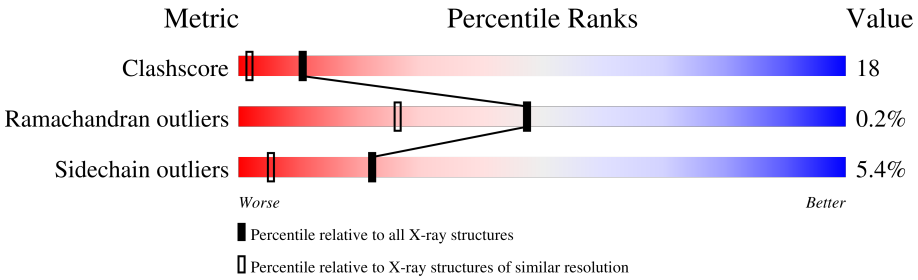
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.60 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	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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 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	496	77% 19% .
1	B	496	76% 19% .
2	C	121	82% 15% . .
2	D	121	72% 18% 8% .

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10951 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 ALPHA-AMYLASE, PANCREATIC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	496	Total	C	N	O	S	0	0	0
			3900	2465	684	730	21			
1	B	496	Total	C	N	O	S	0	0	0
			3894	2459	685	729	21			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	243	LYS	GLN	SEE REMARK 999	UNP P00690
A	310	SER	ALA	SEE REMARK 999	UNP P00690
A	323	ILE	VAL	SEE REMARK 999	UNP P00690
A	404	GLN	GLU	SEE REMARK 999	UNP P00690
B	243	LYS	GLN	SEE REMARK 999	UNP P00690
B	310	SER	ALA	SEE REMARK 999	UNP P00690
B	323	ILE	VAL	SEE REMARK 999	UNP P00690
B	404	GLN	GLU	SEE REMARK 999	UNP P00690

- Molecule 2 is a protein called CAMELID VHH DOMAIN CAB10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	119	Total	C	N	O	S	0	0	0
			882	548	152	175	7			
2	D	121	Total	C	N	O	S	0	0	0
			898	556	155	180	7			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	590	Total	O	0	0
			590	590		
3	B	484	Total	O	0	0
			484	484		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	147	Total O 147 147	0	0
3	D	156	Total O 156 156	0	0

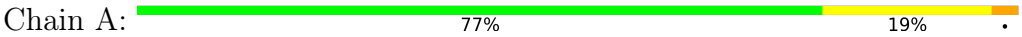
3 Residue-property plots (i)

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.

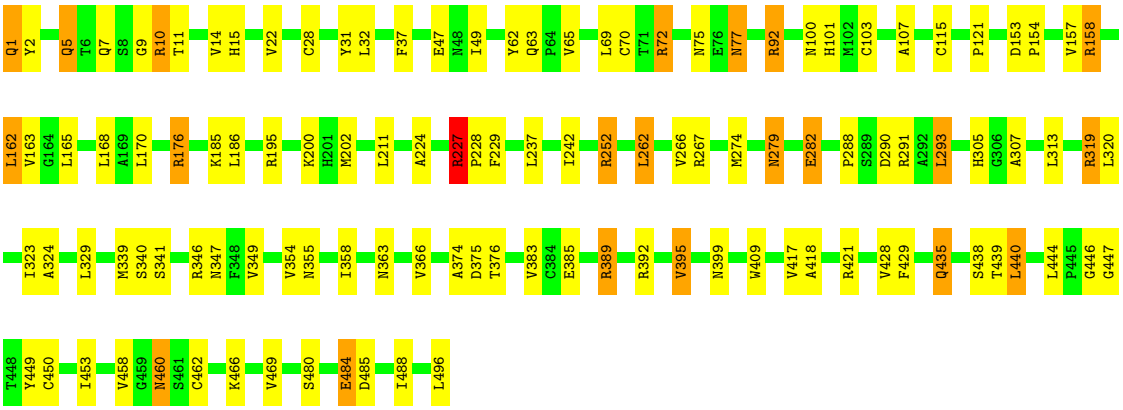
• Molecule 1: ALPHA-AMYLASE, PANCREATIC

Chain A:



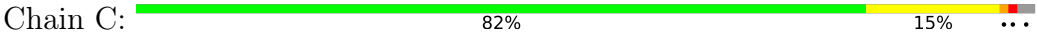
• Molecule 1: ALPHA-AMYLASE, PANCREATIC

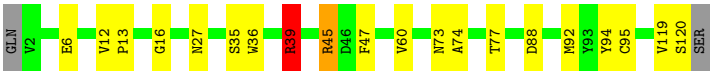
Chain B:



• Molecule 2: CAMELID VHH DOMAIN CAB10

Chain C:





• Molecule 2: CAMELID VHH DOMAIN CAB10

Chain D:

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	57.57Å 60.99Å 107.33Å 98.78° 100.88° 101.14°	Depositor
Resolution (Å)	29.34 – 1.60	Depositor
% Data completeness (in resolution range)	95.1 (29.34-1.60)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 0.9	Depositor
R, R_{free}	0.205 , 0.229	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10951	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	0.93	2/4010 (0.0%)	1.34	35/5450 (0.6%)
1	B	0.88	3/4003 (0.1%)	1.16	25/5439 (0.5%)
2	C	0.82	0/901	1.06	3/1225 (0.2%)
2	D	1.67	5/917 (0.5%)	1.56	18/1245 (1.4%)
All	All	0.99	10/9831 (0.1%)	1.27	81/13359 (0.6%)

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	0	1
2	C	0	2
All	All	0	3

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	121	SER	C-OXT	33.39	1.86	1.23
2	D	8	GLY	C-N	-23.54	0.90	1.33
2	D	1	GLN	C-N	-6.71	1.18	1.34
2	D	121	SER	N-CA	5.68	1.57	1.46
1	B	421	ARG	CD-NE	-5.67	1.36	1.46

The worst 5 of 81 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	60	VAL	CG1-CB-CG2	21.17	144.77	110.90
1	A	195	ARG	NE-CZ-NH1	-21.02	109.79	120.30
1	A	195	ARG	NE-CZ-NH2	20.61	130.61	120.30
2	D	121	SER	N-CA-CB	20.05	140.58	110.50
1	A	252	ARG	NE-CZ-NH2	19.88	130.24	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	468	TYR	Sidechain
2	C	39	ARG	Sidechain
2	C	45	ARG	Sidechain

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	3900	0	3671	143	0
1	B	3894	0	3666	164	0
2	C	882	0	848	16	0
2	D	898	0	862	21	0
3	A	590	0	0	27	0
3	B	484	0	0	24	0
3	C	147	0	0	4	0
3	D	156	0	0	8	0
All	All	10951	0	9047	344	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 18.

The worst 5 of 344 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:1:GLN:NE2	1:B:1:GLN:H1	1.19	1.40
1:B:1:GLN:NE2	1:B:1:GLN:N	1.90	1.19
1:B:282:GLU:OE1	3:B:626:HOH:O	1.66	1.14
1:A:1:GLN:CA	1:A:1:GLN:OE1	2.00	1.09
1:B:15:HIS:HB3	1:B:339:MET:HE3	1.31	1.08

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	494/496 (100%)	482 (98%)	12 (2%)	0	100 100
1	B	494/496 (100%)	480 (97%)	13 (3%)	1 (0%)	47 26
2	C	117/121 (97%)	112 (96%)	5 (4%)	0	100 100
2	D	119/121 (98%)	114 (96%)	4 (3%)	1 (1%)	19 6
All	All	1224/1234 (99%)	1188 (97%)	34 (3%)	2 (0%)	47 26

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	224	ALA
2	D	9	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	411/413 (100%)	391 (95%)	20 (5%)	25 6
1	B	410/413 (99%)	388 (95%)	22 (5%)	22 5
2	C	95/97 (98%)	91 (96%)	4 (4%)	30 9
2	D	97/97 (100%)	88 (91%)	9 (9%)	9 1
All	All	1013/1020 (99%)	958 (95%)	55 (5%)	22 5

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	211	LEU
1	B	329	LEU
2	D	115	THR
2	D	39	ARG
1	B	227	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	101	HIS
2	D	81	GLN
1	B	279	ASN
2	C	81	GLN
1	B	201	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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
2	D	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	1:GLN	C	2:VAL	N	1.18
1	D	8:GLY	C	9:GLY	N	0.90

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.