wwPDB X-ray Structure Validation Summary Report (i)

Feb 4, 2024 – 10:25 PM EST

PDB ID	:	1SYQ
Title	:	Human vinculin head domain VH1, residues 1-258, in complex with human talin's vinculin binding site 1, residues 607-636
Authors	:	Izard, T.; Vonrhein, C.
Deposited on	:	2004-04-01
Resolution	:	2.42 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:

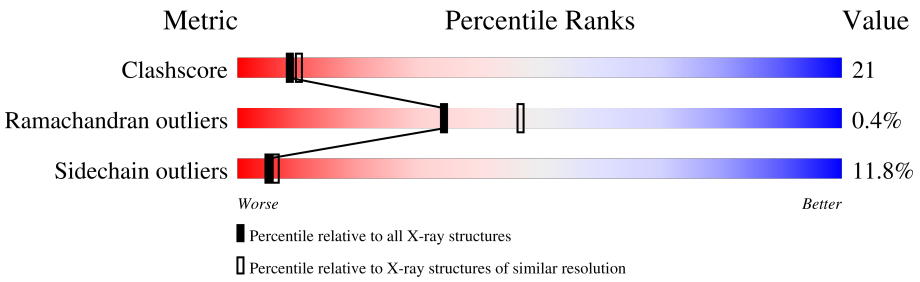
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.42 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	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)

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	264	66% 27% 6% •
2	B	25	60% 40%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2389 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 vinculin isoform VCL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	259	2030	1278	345	393	14	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	HIS	-	expression tag	UNP P18206
A	-4	HIS	-	expression tag	UNP P18206
A	-3	HIS	-	expression tag	UNP P18206
A	-2	HIS	-	expression tag	UNP P18206
A	-1	HIS	-	expression tag	UNP P18206
A	0	HIS	-	expression tag	UNP P18206

- Molecule 2 is a protein called Talin 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	25	169	106	31	32	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	173	173	173	0	0
3	B	17	17	17	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: vinculin isoform VCL

Chain A: 66% 27% 6% •

HIS HIS HIS HIS HIS H0 T8 E14 P15 V16 A17 Q18 Q19 H22 I25 M26 K35 A42 P43 V44 V62 E66 I69 L70 K71 R72 D73 M74 P75 P76 V81 T86 K87 L88 V89 Q90 A91 A92 Q93 M94 L95 Q96 S101 V102 P103 A104 R105 D106

Y107 G111 G118 L124 T125 E128 R136 I141 L142 E143 E152 T153 M154 L157 K162 M171 D176 Q179 L182 Q185 E186 H187 R188 V189 V192 L202 P203 S207 A208 M209 V213 N217 S218 K219 N220 Q221 G222 I223 E224 E225 A226

L227 K228 N229 K236 M237 I241 N242 E243 T251 S252 W253 D254 A257 W258

• Molecule 2: Talin 1

Chain B: 60% 40%

P607 L608 L609 S620 R624 S625 A626 Q627 P628 A629 S630 A631

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	141.10Å 141.10Å 104.92Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	32.25 – 2.42	Depositor
% Data completeness (in resolution range)	98.0 (32.25-2.42)	Depositor
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.188 , 0.236	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2389	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

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	0.60	0/2058	0.72	0/2788
2	B	0.70	0/170	0.69	0/229
All	All	0.61	0/2228	0.72	0/3017

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	2030	0	2081	84	1
2	B	169	0	182	11	0
3	A	173	0	0	12	1
3	B	17	0	0	4	0
All	All	2389	0	2263	92	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 21.

The worst 5 of 92 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:257:ALA:O	1:A:258:TRP:CD1	1.76	1.39

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:257:ALA:O	1:A:258:TRP:HD1	1.14	1.15
1:A:157:LEU:CD1	1:A:209:MET:CE	2.46	0.93
1:A:219:LYS:HD3	3:A:370:HOH:O	1.71	0.90
1:A:157:LEU:HD13	1:A:209:MET:CE	2.04	0.88

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:221:GLN:NE2	3:A:259:HOH:O[18_654]	1.91	0.29

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	257/264 (97%)	248 (96%)	8 (3%)	1 (0%)	34	47
2	B	23/25 (92%)	23 (100%)	0	0	100	100
All	All	280/289 (97%)	271 (97%)	8 (3%)	1 (0%)	34	47

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	222	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	230/235 (98%)	201 (87%)	29 (13%)	4 5
2	B	16/16 (100%)	16 (100%)	0	100 100
All	All	246/251 (98%)	217 (88%)	29 (12%)	5 6

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	136	ARG
1	A	253	TRP
1	A	185	GLN
1	A	229	ASN
1	A	162	LYS

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

Mol	Chain	Res	Type
1	A	18	GLN
1	A	229	ASN
1	A	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)

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)

There are no chain breaks in this entry.

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.