



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2023 – 09:37 PM EDT

PDB ID : 2DOI
Title : The X-ray crystallographic structure of the angiogenesis inhibitor, angiostatin, bound to a peptide from the group A streptococcus protein PAM
Authors : Cnudde, S.E.; Prorok, M.; Castellino, F.J.; Geiger, J.H.
Deposited on : 2006-04-29
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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

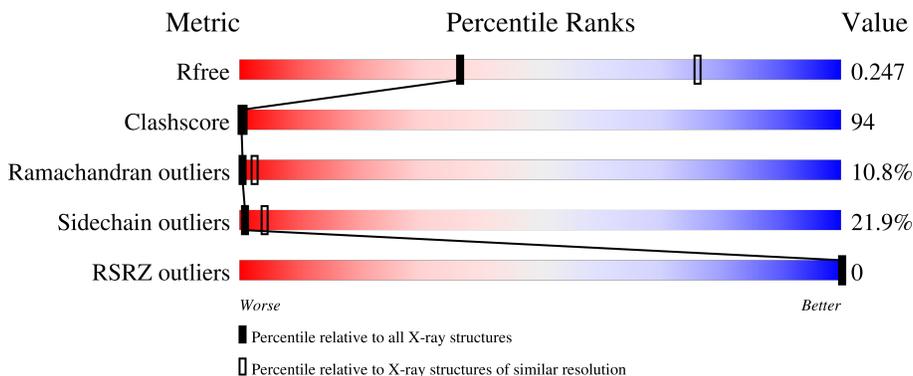
1 Overall quality at a glance

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}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 $\leq 5\%$. 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	234	
1	X	234	
2	B	30	
2	C	30	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3064 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 Angiostatin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	X	168	1347	824	242	264	17	0	0	0
1	A	164	1313	803	235	259	16	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	289	GLU	ASN	engineered mutation	UNP P00747
A	289	GLU	ASN	engineered mutation	UNP P00747

- Molecule 2 is a protein called Plasminogen-binding group A streptococcal M-like protein PAM.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	24	202	122	39	41	0	0	0
2	B	24	202	122	39	41	0	0	0

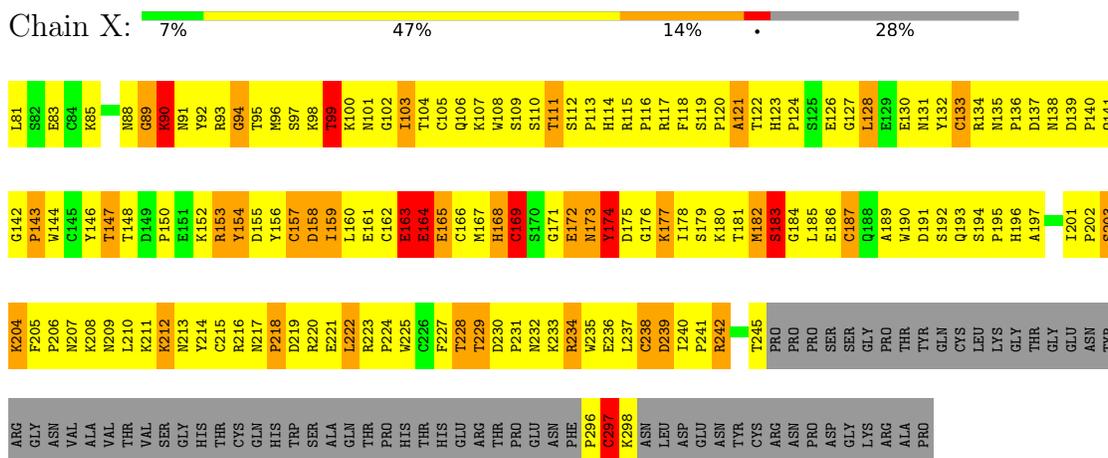
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	330	TYR	-	cloning artifact	UNP P49054
B	330	TYR	-	cloning artifact	UNP P49054

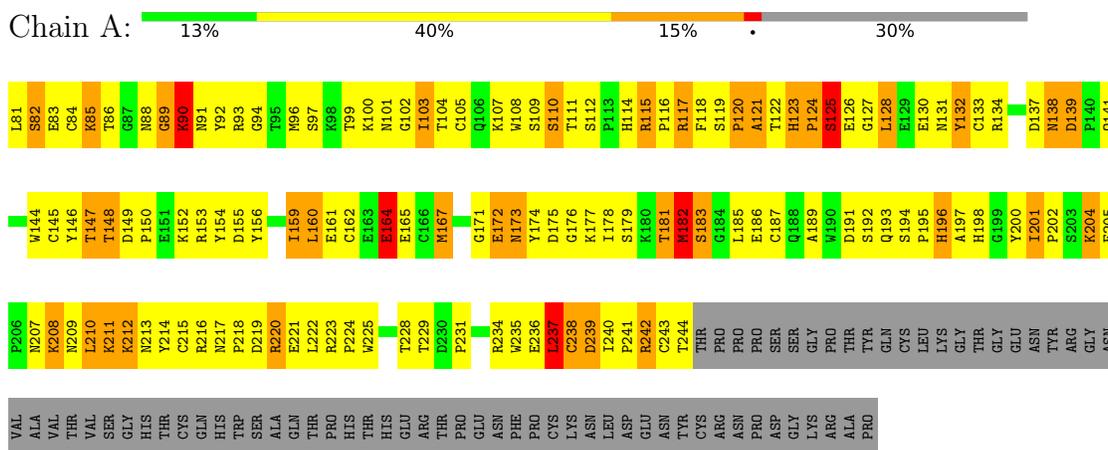
3 Residue-property plots

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 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: Angiostatin



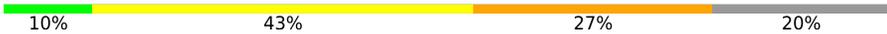
- Molecule 1: Angiostatin



- Molecule 2: Plasminogen-binding group A streptococcal M-like protein PAM



- Molecule 2: Plasminogen-binding group A streptococcal M-like protein PAM

Chain B: 

VAL	GLU	L304	LYS	T305	A306	D307	A308	E309	L310	Q311	R312	L313	K314	R317	H318	E319	E320	A321	E322	L323	E324	R325	L326	K327	SER	GLU	TYR
-----	-----	------	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----

4 Data and refinement statistics i

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α , β , γ	58.47Å 58.47Å 389.15Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.10 19.96 – 2.81	Depositor EDS
% Data completeness (in resolution range)	97.0 (20.00-3.10) 85.9 (19.96-2.81)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 2.79Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.202 , 0.296 0.203 , 0.247	Depositor DCC
R_{free} test set	1656 reflections (10.18%)	wwPDB-VP
Wilson B-factor (Å ²)	77.0	Xtrriage
Anisotropy	0.133	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 38.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	0.479 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3064	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.10% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	1.01	0/1353	1.16	5/1837 (0.3%)
1	X	1.12	2/1387 (0.1%)	1.21	5/1878 (0.3%)
2	B	1.12	0/202	1.35	1/268 (0.4%)
2	C	1.06	0/202	1.38	2/268 (0.7%)
All	All	1.07	2/3144 (0.1%)	1.21	13/4251 (0.3%)

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

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	X	133	CYS	CB-SG	-6.01	1.72	1.82
1	X	187	CYS	CB-SG	-5.07	1.73	1.81

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	326	LEU	CA-CB-CG	8.21	134.17	115.30
1	X	128	LEU	CA-CB-CG	-6.39	100.60	115.30
1	X	185	LEU	CB-CG-CD2	-6.25	100.38	111.00
2	C	304	LEU	CB-CG-CD1	6.14	121.45	111.00
1	A	181	THR	CB-CA-C	-6.07	95.22	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	208	LYS	Peptide

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	1313	0	1182	254	0
1	X	1347	0	1225	257	0
2	B	202	0	202	37	0
2	C	202	0	202	37	0
All	All	3064	0	2811	552	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 94.

The worst 5 of 552 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:187:CYS:HB2	1:A:213:ASN:ND2	1.62	1.12
1:X:103:ILE:HD12	1:X:103:ILE:H	1.08	1.11
1:X:195:PRO:O	1:X:196:HIS:HD2	1.32	1.10
1:X:224:PRO:CG	1:X:240:ILE:HD12	1.81	1.10
1:X:133:CYS:O	1:X:134:ARG:HD3	1.51	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	162/234 (69%)	113 (70%)	36 (22%)	13 (8%)	1	5
1	X	164/234 (70%)	112 (68%)	32 (20%)	20 (12%)	0	1
2	B	22/30 (73%)	14 (64%)	7 (32%)	1 (4%)	2	15
2	C	22/30 (73%)	9 (41%)	7 (32%)	6 (27%)	0	0
All	All	370/528 (70%)	248 (67%)	82 (22%)	40 (11%)	0	2

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	121	ALA
1	X	127	GLY
1	X	159	ILE
1	X	163	GLU
1	X	169	CYS

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	148/212 (70%)	117 (79%)	31 (21%)	1	5
1	X	153/212 (72%)	118 (77%)	35 (23%)	1	3
2	B	21/27 (78%)	15 (71%)	6 (29%)	0	1
2	C	21/27 (78%)	18 (86%)	3 (14%)	3	14
All	All	343/478 (72%)	268 (78%)	75 (22%)	1	4

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	182	MET
2	B	311	GLN
1	A	191	ASP
1	A	237	LEU
1	X	191	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	318	HIS
1	A	101	ASN
2	B	318	HIS
1	A	209	ASN
1	A	213	ASN

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](#)

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	164/234 (70%)	-0.94	0 100 100	10, 33, 58, 81	0
1	X	168/234 (71%)	-0.96	0 100 100	11, 33, 55, 66	0
2	B	24/30 (80%)	-0.94	0 100 100	15, 43, 52, 70	0
2	C	24/30 (80%)	-0.81	0 100 100	13, 33, 55, 60	0
All	All	380/528 (71%)	-0.94	0 100 100	10, 33, 58, 81	0

There are no RSRZ outliers to report.

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 monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.