



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 28, 2024 – 02:47 pm GMT

PDB ID : 2WXX
Title : Crystal structure of mouse angiotensinogen in the oxidised form
Authors : Zhou, A.; Wei, Z.; Carrell, R.W.; Read, R.J.
Deposited on : 2009-11-11
Resolution : 2.95 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	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

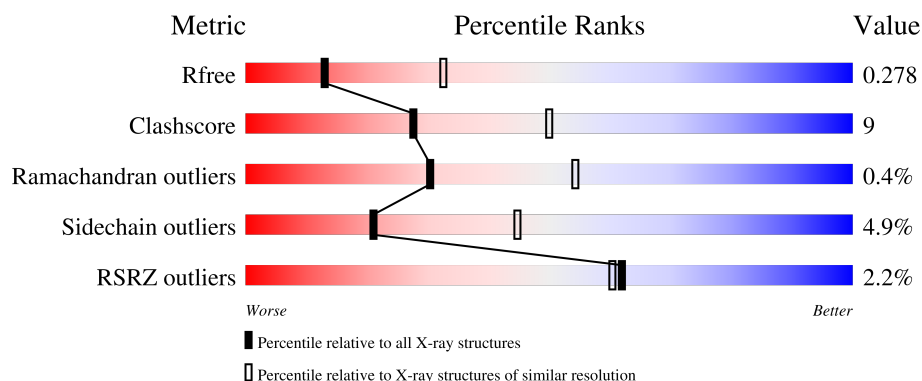
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

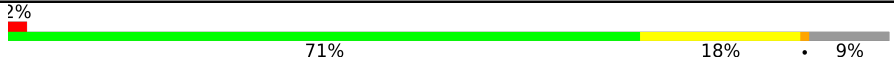



The reported resolution of this entry is 2.95 Å.

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}	164625	1044 (2.98-2.94)
Clashscore	180529	1097 (2.98-2.94)
Ramachandran outliers	177936	1049 (2.98-2.94)
Sidechain outliers	177891	1049 (2.98-2.94)
RSRZ outliers	164620	1044 (2.98-2.94)

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	453	 2% 71% 18% • 9%
1	B	453	 % 75% 18% • 6%
1	C	453	 2% 71% 20% • 7%
1	D	453	 2% 73% 17% • 9%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 12945 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 ANGIOTENSINOGEN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	412	Total	C	N	O	S	0	1	0
			3181	2036	535	600	10			
1	B	428	Total	C	N	O	S	0	2	0
			3314	2117	558	629	10			
1	C	422	Total	C	N	O	S	0	3	0
			3274	2095	550	619	10			
1	D	411	Total	C	N	O	S	0	2	0
			3176	2036	533	597	10			

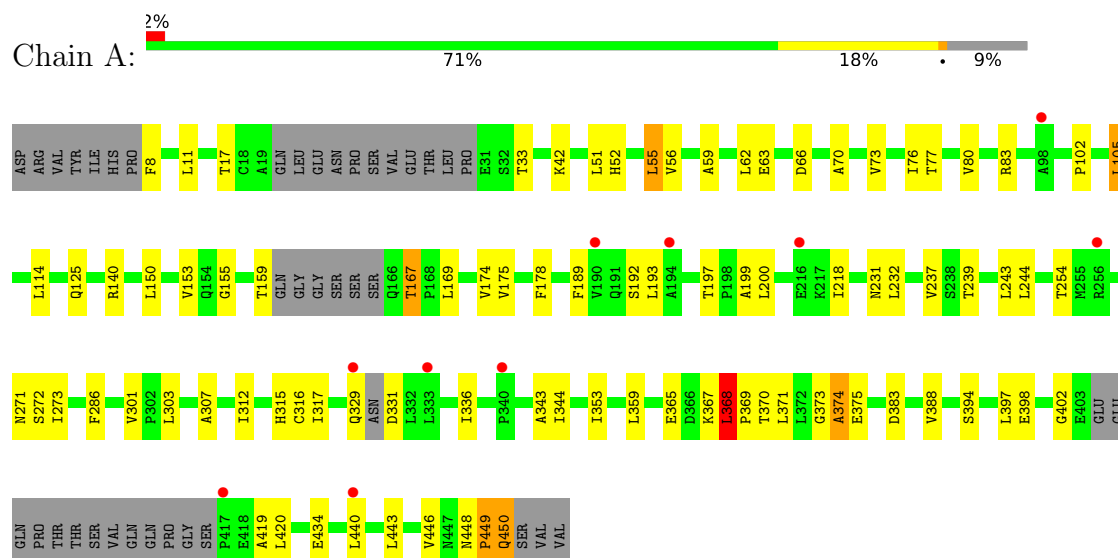
There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	THR	ALA	conflict	UNP P11859
A	174	VAL	MET	conflict	UNP P11859
A	317	ILE	THR	conflict	UNP P11859
A	329	GLN	ARG	conflict	UNP P11859
A	379	ASN	SER	conflict	UNP P11859
B	77	THR	ALA	conflict	UNP P11859
B	174	VAL	MET	conflict	UNP P11859
B	317	ILE	THR	conflict	UNP P11859
B	329	GLN	ARG	conflict	UNP P11859
B	379	ASN	SER	conflict	UNP P11859
C	77	THR	ALA	conflict	UNP P11859
C	174	VAL	MET	conflict	UNP P11859
C	317	ILE	THR	conflict	UNP P11859
C	329	GLN	ARG	conflict	UNP P11859
C	379	ASN	SER	conflict	UNP P11859
D	77	THR	ALA	conflict	UNP P11859
D	174	VAL	MET	conflict	UNP P11859
D	317	ILE	THR	conflict	UNP P11859
D	329	GLN	ARG	conflict	UNP P11859
D	379	ASN	SER	conflict	UNP P11859

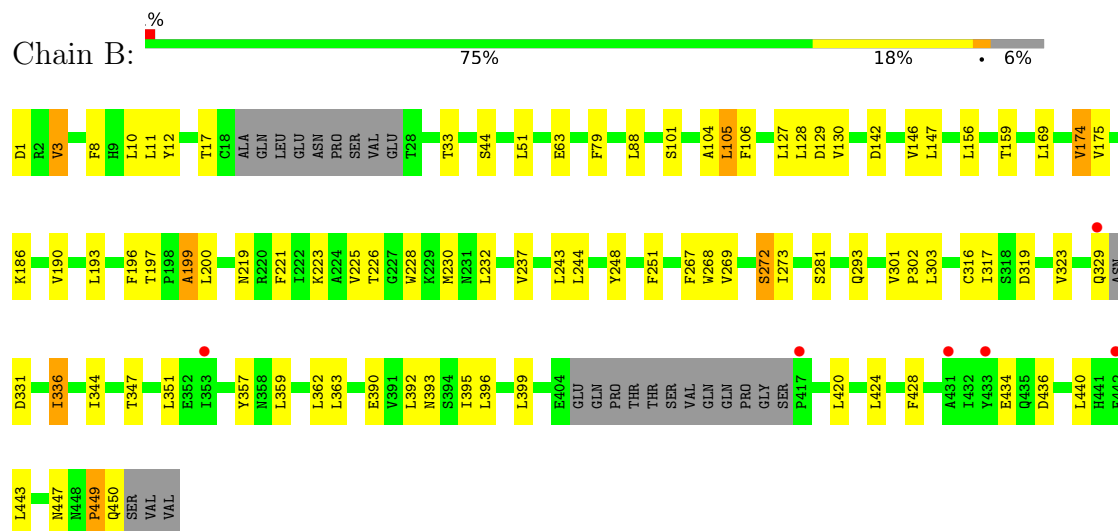
3 Residue-property plots [i](#)

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 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: ANGIOTENSINOGEN

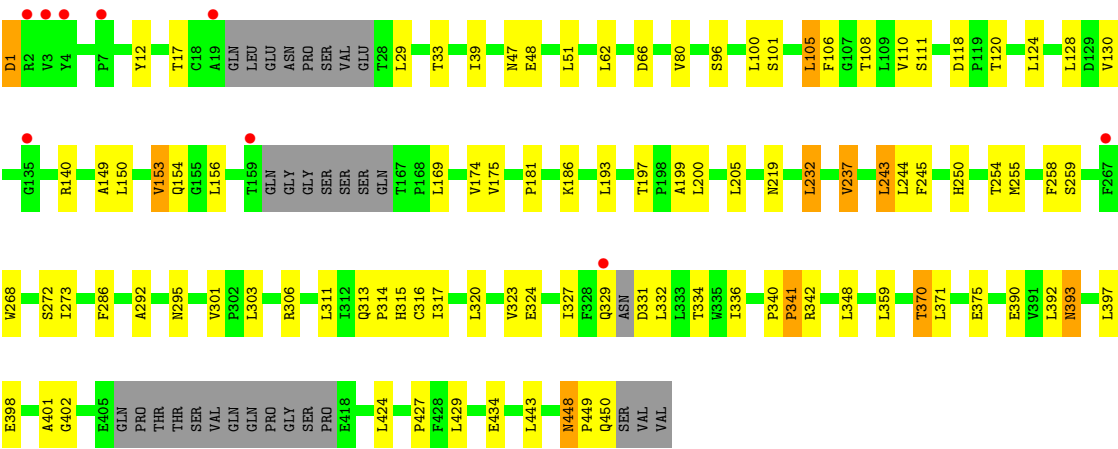


• Molecule 1: ANGIOTENSINOGEN

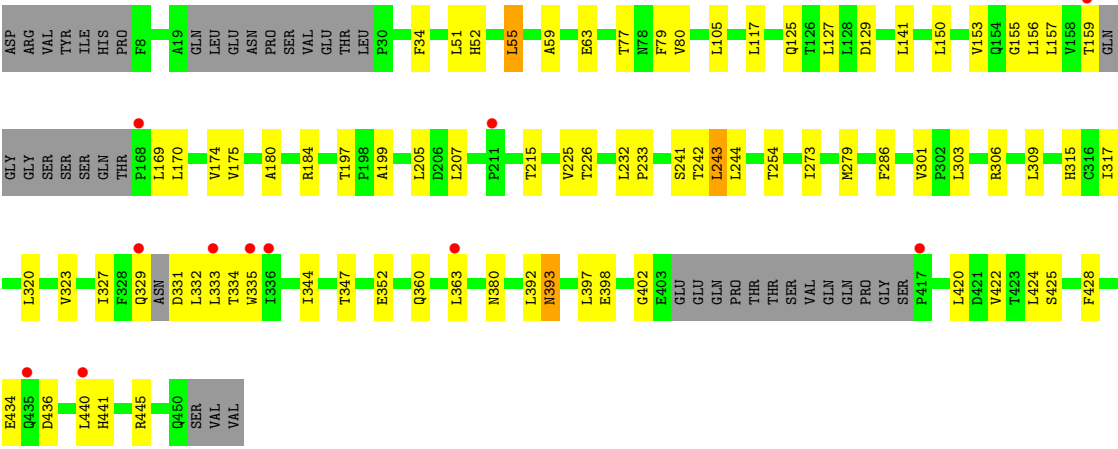


• Molecule 1: ANGIOTENSINOGEN





● Molecule 1: ANGIOTENSINOGEN



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	95.37Å 99.42Å 425.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	213.20 – 2.95 212.88 – 2.95	Depositor EDS
% Data completeness (in resolution range)	84.5 (213.20-2.95) 84.5 (212.88-2.95)	Depositor EDS
R_{merge}	0.01	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.51 (at 2.96Å)	Xtriage
Refinement program	REFMAC 5.5.0099	Depositor
R, R_{free}	0.223 , 0.271 0.230 , 0.278	Depositor DCC
R_{free} test set	1850 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	76.7	Xtriage
Anisotropy	0.474	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 60.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.055 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	12945	wwPDB-VP
Average B, all atoms (Å ²)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.82% 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

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	0.40	1/3251 (0.0%)	0.54	1/4427 (0.0%)
1	B	0.38	0/3392	0.56	0/4621
1	C	0.34	0/3353	0.50	0/4567
1	D	0.31	0/3250	0.47	0/4424
All	All	0.36	1/13246 (0.0%)	0.52	1/18039 (0.0%)

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	2	0
1	B	1	0
1	C	1	0
1	D	1	0
All	All	5	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	167	THR	CB-OG1	5.40	1.54	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	368	LEU	CA-CB-CG	-5.03	103.73	115.30

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	167	THR	CB

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom
1	A	327	ILE	CB
1	B	327	ILE	CB
1	C	327	ILE	CB
1	D	327	ILE	CB

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	3181	0	3179	66	0
1	B	3314	0	3308	56	0
1	C	3274	0	3261	67	0
1	D	3176	0	3173	46	0
All	All	12945	0	12921	230	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 9.

The worst 5 of 230 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:367:LYS:C	1:A:368:LEU:HD23	1.50	1.29
1:A:371:LEU:HD23	1:A:371:LEU:O	1.48	1.12
1:A:367:LYS:O	1:A:368:LEU:HD23	1.62	1.00
1:C:448:ASN:C	1:C:448:ASN:HD22	1.66	0.97
1:C:193:LEU:HD23	1:C:200:LEU:HD11	1.47	0.94

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	403/453 (89%)	383 (95%)	17 (4%)	3 (1%)	19	41
1	B	422/453 (93%)	402 (95%)	17 (4%)	3 (1%)	19	41
1	C	415/453 (92%)	393 (95%)	22 (5%)	0	100	100
1	D	403/453 (89%)	385 (96%)	18 (4%)	0	100	100
All	All	1643/1812 (91%)	1563 (95%)	74 (4%)	6 (0%)	30	54

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	374	ALA
1	B	449	PRO
1	B	199	ALA
1	B	336	ILE
1	A	449	PRO

5.3.2 Protein sidechains ⓘ

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	354/394 (90%)	334 (94%)	20 (6%)	17	40
1	B	371/394 (94%)	356 (96%)	15 (4%)	27	51
1	C	364/394 (92%)	342 (94%)	22 (6%)	16	37
1	D	352/394 (89%)	339 (96%)	13 (4%)	29	54
All	All	1441/1576 (91%)	1371 (95%)	70 (5%)	21	45

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	55	LEU
1	D	105	LEU
1	D	352	GLU
1	B	44	SER
1	B	17	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	448	ASN
1	D	380	ASN
1	D	393	ASN
1	B	393	ASN
1	B	447	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 oligosaccharides 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 ⓘ

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	412/453 (90%)	-0.15	10 (2%) 59 57	50, 75, 114, 130	10 (2%)
1	B	428/453 (94%)	-0.15	6 (1%) 73 71	45, 76, 113, 142	11 (2%)
1	C	422/453 (93%)	-0.09	9 (2%) 63 62	26, 88, 118, 138	12 (2%)
1	D	411/453 (90%)	0.12	11 (2%) 56 53	52, 118, 182, 230	11 (2%)
All	All	1673/1812 (92%)	-0.07	36 (2%) 62 60	26, 86, 146, 230	44 (2%)

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	135	GLY	5.3
1	C	4	TYR	4.6
1	D	329	GLN	3.9
1	C	3	VAL	3.8
1	A	194	ALA	2.9

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.