wwPDB X-ray Structure Validation Summary Report (i)

Jul 31, 2023 – 10:46 PM EDT

PDB ID	:	2H1P
Title	:	THE THREE-DIMENSIONAL STRUCTURES OF A POLYSACCHARIDE BINDING ANTIBODY TO CRYPTOCOCCUS NEOFORMANS AND ITS COMPLEX WITH A PEPTIDE FROM A PHAGE DISPLAY LIBRARY: IMPLICATIONS FOR THE IDENTIFICATION OF PEPTIDE MIMOTOPES
Authors	:	Young, A.C.M.; Valadon, P.; Casadevall, A.; Scharff, M.D.; Sacchettini, J.C.
Deposited on	:	1997-11-12
Resolution	:	2.40 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:

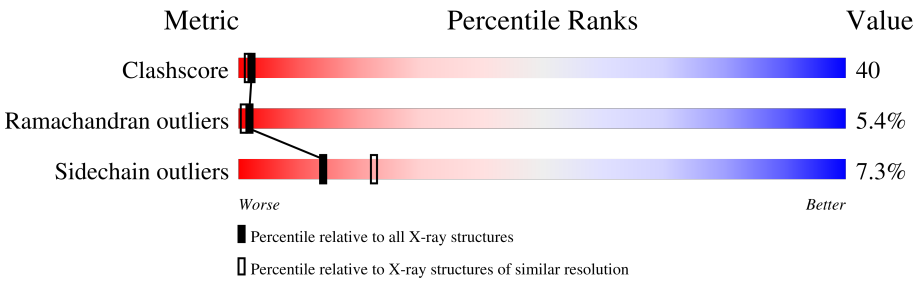
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.40 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	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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	L	219	42% 50% 5% •
2	H	220	39% 52% 8% •
3	P	12	17% 58% 17% 8%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3509 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 2H1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	219	1696	1059	287	343	7	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	18	PRO	GLN	conflict	PIR S16112
L	39	HIS	TYR	conflict	PIR S16112
L	51	LEU	PRO	conflict	PIR S16112
L	55	LYS	ARG	conflict	PIR S16112
L	66	LYS	ARG	conflict	PIR S16112
L	88	GLN	LEU	conflict	PIR S16112
L	94	SER	PHE	conflict	PIR S16112
L	96	SER	GLY	conflict	PIR S16112
L	101	TRP	TYR	conflict	PIR S16112
L	108	LYS	ARG	conflict	PIR S16112
L	171	GLU	GLN	conflict	PIR S16112

- Molecule 2 is a protein called 2H1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	220	1665	1052	276	330	7	0	0	0

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	305	VAL	LEU	conflict	PIR S38864
H	310	GLY	ASP	conflict	PIR S38864
H	314	LEU	PRO	conflict	PIR S38864
H	327	PHE	LEU	conflict	PIR S38864
H	333	PHE	GLY	conflict	PIR S38864

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	334	LEU	MET	conflict	PIR S38864
H	340	THR	ILE	conflict	PIR S38864
H	342	GLU	ASP	conflict	PIR S38864
H	347	LEU	TRP	conflict	PIR S38864
H	352	ASN	SER	conflict	PIR S38864
H	354	ASN	GLY	conflict	PIR S38864
H	356	ASP	THR	conflict	PIR S38864
H	357	LYS	TYR	conflict	PIR S38864
H	360	HIS	TYR	conflict	PIR S38864
H	363	THR	SER	conflict	PIR S38864
H	364	MET	VAL	conflict	PIR S38864
H	393	LEU	MET	conflict	PIR S38864
H	399	ARG	GLN	conflict	PIR S38864
H	400	ASP	GLY	conflict	PIR S38864
H	401	SER	VAL	conflict	PIR S38864
H	403	ALA	THR	conflict	PIR S38864
H	404	SER	MET	conflict	PIR S38864
H	405	LEU	ILE	conflict	PIR S38864
H	406	TYR	ARG	conflict	PIR S38864
H	408	ASP	ALA	conflict	PIR S38864
H	415	THR	LEU	conflict	PIR S38864
H	416	LEU	VAL	conflict	PIR S38864
H	420	SER	ALA	conflict	PIR S38864
H	421	ALA	GLY	conflict	PIR S38864

- Molecule 3 is a protein called PA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	P	11	86	56	13	16	1	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	L	28	28	28	0	0
4	H	31	Total 31	O 31	0	0
4	P	3	Total 3	O 3	0	0

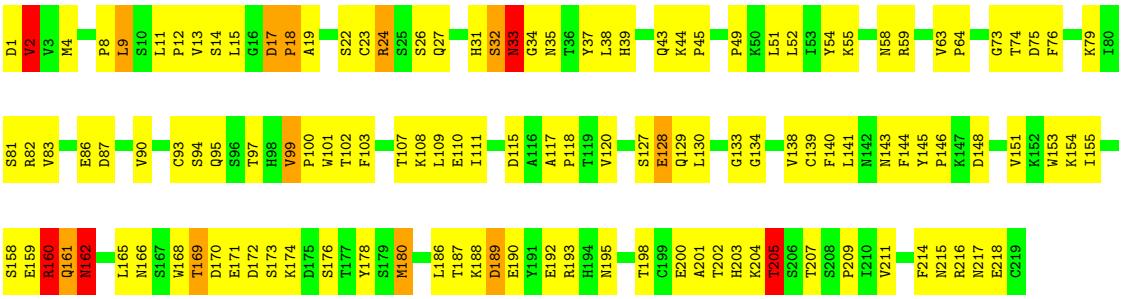
3 Residue-property plots (i)

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.

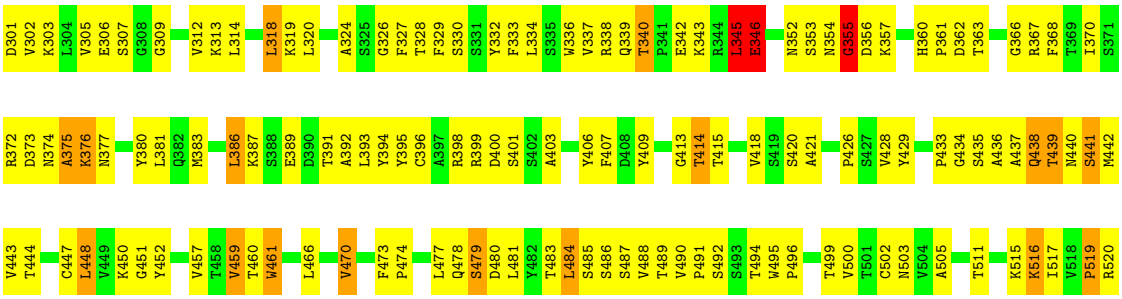
• Molecule 1: 2H1

Chain L: 42% 50% 5%



• Molecule 2: 2H1

Chain H: 39% 52% 8%



• Molecule 3: PA1

Chain P: 17% 58% 17% 8%

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	43.79Å 51.25Å 60.18Å 91.82° 98.57° 107.84°	Depositor
Resolution (Å)	20.00 – 2.40	Depositor
% Data completeness (in resolution range)	90.0 (20.00-2.40)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	TNT, X-PLOR	Depositor
R, R_{free}	0.186 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3509	wwPDB-VP
Average B, all atoms (Å^2)	34.0	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	L	0.79	0/1737	1.01	5/2358 (0.2%)
2	H	0.77	0/1705	1.08	6/2324 (0.3%)
3	P	1.29	1/89 (1.1%)	1.01	0/120
All	All	0.80	1/3531 (0.0%)	1.05	11/4802 (0.2%)

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	L	2	0
2	H	1	1
All	All	3	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	P	612	GLY	N-CA	-10.69	1.30	1.46

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	346	GLU	N-CA-C	8.49	133.93	111.00
1	L	162	ASN	N-CA-C	6.43	128.35	111.00
1	L	165	LEU	CA-CB-CG	5.82	128.68	115.30
2	H	520	ARG	N-CA-C	5.73	126.48	111.00
1	L	2	VAL	N-CA-C	-5.63	95.80	111.00

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	L	162	ASN	CA
1	L	202	THR	CB
2	H	346	GLU	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	362	ASP	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	L	1696	0	1630	117	3
2	H	1665	0	1644	147	1
3	P	86	0	72	17	0
4	H	31	0	0	2	1
4	L	28	0	0	4	0
4	P	3	0	0	0	0
All	All	3509	0	3346	270	3

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 40.

The worst 5 of 270 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)
2:H:480:ASP:OD1	2:H:481:LEU:HD13	1.58	1.04
3:P:608:TRP:O	3:P:609:MET:HB2	1.57	1.00
2:H:340:THR:CG2	2:H:342:GLU:HG2	1.94	0.98
1:L:1:ASP:O	1:L:2:VAL:CG2	2.14	0.96
2:H:340:THR:HG21	2:H:342:GLU:HG2	1.46	0.94

All (3) 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:L:159:GLU:OE1	4:H:733:HOH:O[1_455]	1.96	0.24
1:L:55:LYS:NZ	1:L:192:GLU:OE1[1_656]	2.02	0.18
1:L:174:LYS:NZ	2:H:366:GLY:O[1_565]	2.14	0.06

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	L	217/219 (99%)	197 (91%)	10 (5%)	10 (5%)	2	1
2	H	218/220 (99%)	183 (84%)	24 (11%)	11 (5%)	2	1
3	P	9/12 (75%)	5 (56%)	1 (11%)	3 (33%)	0	0
All	All	444/451 (98%)	385 (87%)	35 (8%)	24 (5%)	2	1

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	18	PRO
1	L	32	SER
1	L	33	ASN
1	L	162	ASN
1	L	205	THR

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	L	197/197 (100%)	181 (92%)	16 (8%)	11	18

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	H	191/191 (100%)	178 (93%)	13 (7%)	16	25
3	P	8/10 (80%)	8 (100%)	0	100	100
All	All	396/398 (100%)	367 (93%)	29 (7%)	14	22

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

Mol	Chain	Res	Type
1	L	207	THR
2	H	502	CYS
2	H	320	LEU
2	H	461	TRP
2	H	318	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	438	GLN
2	H	440	ASN
1	L	129	GLN
1	L	142	ASN
1	L	161	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.