

# Full wwPDB X-ray Structure Validation Report (i)

#### Oct 5, 2024 – 12:34 PM EDT

PDB ID : 2H0E

Title : Crystal Structure of PucM in the absence of substrate

Authors : Rhee, S. Deposited on : 2006-05-14

Resolution : 2.20 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (1)) were used in the production of this report:

MolProbity : 4.02b-467

Mogul : 2022.3.0, CSD as543be (2022)

Xtriage (Phenix) : 1.20.1 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)

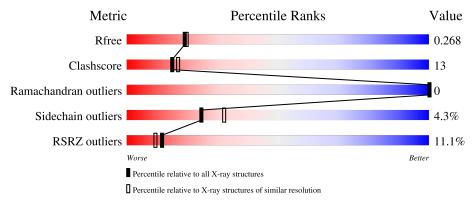
Validation Pipeline (wwPDB-VP) : 2.39

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X- $RAY\ DIFFRACTION$ 

The reported resolution of this entry is 2.20 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 $(\# \mathrm{Entries})$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries},\ {\rm resolution\ range}({\rm \AA})) \end{array}$
$R_{free}$	164625	5791 (2.20-2.20)
Clashscore	180529	6634 (2.20-2.20)
Ramachandran outliers	177936	6560 (2.20-2.20)
Sidechain outliers	177891	6561 (2.20-2.20)
RSRZ outliers	164620	5791 (2.20-2.20)

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% 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	121	12%	21%	• 7%			
1	В	121	7%	22%	• 6%			

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:



	Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
ſ	2	GOL	В	201	-	-	-	X



# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1859 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 Transthyretin-like protein pucM.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace			
1	A	112	Total 892	C 571	N 151	O 165	S 1	Se 4	0	2	0
1	В	114	Total 888	C 568	N 148	O 166	S 1	Se 5	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP O32142
A	8	MSE	MET	modified residue	UNP O32142
A	39	MSE	MET	modified residue	UNP O32142
A	61	MSE	MET	modified residue	UNP O32142
A	67	MSE	MET	modified residue	UNP O32142
A	80	MSE	MET	modified residue	UNP O32142
В	1	MSE	MET	modified residue	UNP O32142
В	8	MSE	MET	modified residue	UNP O32142
В	39	MSE	MET	modified residue	UNP O32142
В	61	MSE	MET	modified residue	UNP O32142
В	67	MSE	MET	modified residue	UNP O32142
В	80	MSE	MET	modified residue	UNP O32142

• Molecule 2 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0
2	В	1	Total C O 6 3 3	0	0

### • Molecule 3 is water.

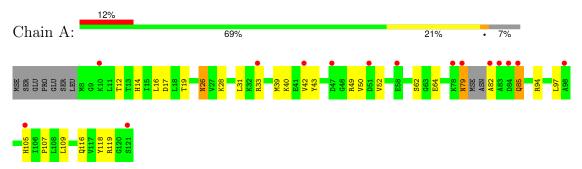
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	34	Total O 34 34	0	0
3	В	33	Total O 33 33	0	0



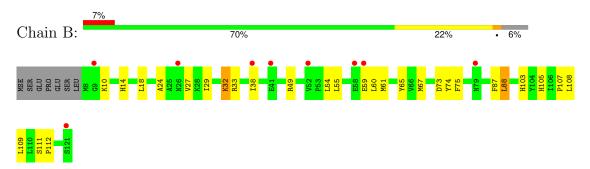
# 3 Residue-property plots (i)

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 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: Transthyretin-like protein pucM



• Molecule 1: Transthyretin-like protein pucM





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 2 2	Depositor
Cell constants	73.17Å 73.17Å 145.02Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 - 2.20	Depositor
rtesolution (A)	50.00 - 2.20	EDS
% Data completeness	78.5 (50.00-2.20)	Depositor
(in resolution range)	94.8 (50.00-2.20)	EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.99 (at 2.20Å)	Xtriage
Refinement program	CNS	Depositor
$R, R_{free}$	0.224 , $0.256$	Depositor
It, It free	0.238 , 0.268	DCC
$R_{free}$ test set	2007  reflections  (9.77%)	wwPDB-VP
Wilson B-factor $(\mathring{A}^2)$	39.8	Xtriage
Anisotropy	0.438	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.37, 32.4	EDS
L-test for twinning <sup>2</sup>	$ < L >=0.49, < L^2>=0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	1859	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.36% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical values of <|L|>,  $<L^2>$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 5 Model quality (i)

## 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL

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
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	A	0.36	0/910	0.62	0/1225
1	В	0.35	0/904	0.63	0/1216
All	All	0.36	0/1814	0.62	0/2441

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	892	0	874	24	0
1	В	888	0	878	25	0
2	A	6	0	8	1	0
2	В	6	0	8	0	0
3	A	34	0	0	0	0
3	В	33	0	0	0	0
All	All	1859	0	1768	48	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 13.

All (48) close contacts within the same asymmetric unit are listed below, sorted by their clash



### magnitude.

Atom-1	Atom-2	Interatomic	Clash
1 D 22 A DC 111112	1 D C1 MCD HD9	distance (Å)	overlap (Å)
1:B:33:ARG:HH12	1:B:61:MSE:HE3	1.29	0.94
1:B:33:ARG:NH1	1:B:61:MSE:HE3	2.01	0.74
1:A:14[B]:HIS:HB2	1:A:49:ARG:HG2	1.74	0.69
1:A:85:GLN:HE21	1:A:85:GLN:HA	1.58	0.67
1:B:32:LYS:HB3	1:B:38:ILE:HA	1.76	0.66
1:B:33:ARG:HH12	1:B:61:MSE:CE	2.07	0.64
1:A:33:ARG:HG3	1:A:39:MSE:CE	2.31	0.61
1:B:10:LYS:HG2	1:B:103:HIS:ND1	2.16	0.60
1:A:33:ARG:NH1	1:A:39:MSE:HE1	2.16	0.60
1:A:14[B]:HIS:CE1	1:A:107:PRO:HB3	2.38	0.59
1:A:64:GLU:OE1	1:A:94:ARG:HD3	2.05	0.57
1:A:42:VAL:HG12	1:A:43:TYR:N	2.21	0.56
1:B:105:HIS:O	1:B:107:PRO:HD3	2.07	0.54
1:A:42:VAL:HG11	1:A:50:VAL:HG11	1.91	0.52
1:B:33:ARG:HH22	1:B:61:MSE:HE3	1.75	0.51
1:B:27:VAL:HG22	1:B:74:TYR:CD1	2.46	0.50
1:B:14:HIS:HD2	1:B:49:ARG:CZ	2.25	0.50
1:A:39:MSE:HG3	1:A:40:LYS:HG3	1.93	0.50
1:A:14[B]:HIS:NE2	1:A:16:LEU:HD21	2.27	0.48
1:B:87:PHE:CD2	1:B:88:LEU:HD13	2.48	0.48
1:A:33:ARG:HG3	1:A:39:MSE:HE1	1.95	0.48
1:B:18:LEU:HD21	1:B:109:LEU:HB3	1.95	0.48
1:A:17:ASP:OD1	1:A:19:THR:HB	2.14	0.48
1:A:79:ASN:CB	1:A:82:ALA:HB2	2.44	0.47
1:A:50:VAL:HG23	1:A:52:VAL:O	2.15	0.47
1:A:12:THR:OG1	1:A:105[A]:HIS:ND1	2.46	0.46
1:B:54:LEU:HD11	1:B:67:MSE:HE1	1.97	0.46
1:B:87:PHE:CE2	1:B:88:LEU:HD13	2.52	0.44
1:B:18:LEU:HD12	1:B:111:SER:HB2	2.00	0.44
1:B:55:LEU:CD2	1:B:59:GLU:HG2	2.48	0.43
1:A:26:ASN:HD22	1:A:26:ASN:HA	1.61	0.43
1:B:54:LEU:CD1	1:B:67:MSE:HE1	2.49	0.43
1:A:42:VAL:CG1	1:A:43:TYR:N	2.82	0.43
1:B:55:LEU:HD22	1:B:59:GLU:HG2	1.99	0.43
1:B:33:ARG:NH2	1:B:61:MSE:HE3	2.33	0.42
1:A:62:SER:HA	1:A:97:LEU:O	2.19	0.42
1:A:14[B]:HIS:CE1	2:A:200:GOL:O2	2.71	0.42
1:B:60:LEU:C	1:B:60:LEU:HD23	2.39	0.42
1:A:28:LYS:HE2	1:B:73:ASP:OD1	2.20	0.42
1:A:109:LEU:HB2	1:A:116:GLN:HB3	2.01	0.42
1:B:55:LEU:HD11	1:B:65:TYR:CE2	2.54	0.41

Continued on next page...



I 'omtamalod	tmom	mmonia	maaa
Continued	110111	DIEUIUU	DUJUE
00,000,000	.,	p	p = 9 =

Atom-1	Atom-2	Interatomic	Clash
		distance (Å)	overlap (Å)
1:A:42:VAL:HG11	1:A:50:VAL:CG1	2.50	0.41
1:B:33:ARG:HH22	1:B:61:MSE:CE	2.33	0.41
1:B:24:ALA:HB1	1:B:27:VAL:HG21	2.03	0.40
1:A:85:GLN:HA	1:A:85:GLN:NE2	2.31	0.40
1:B:29:ILE:HG21	1:B:67:MSE:CE	2.51	0.40
1:B:75:PHE:CG	1:B:112:PRO:HG3	2.56	0.40
1:A:107:PRO:HB2	1:A:118:TYR:CE1	2.56	0.40

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	Percei	ntiles
1	A	110/121 (91%)	107 (97%)	3 (3%)	0	100	100
1	В	112/121 (93%)	107 (96%)	5 (4%)	0	100	100
All	All	$222/242 \ (92\%)$	214 (96%)	8 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	95/96 (99%)	90 (95%)	5 (5%)	19 24		

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	В	95/96 (99%)	92 (97%)	3 (3%)	34 45
All	All	190/192 (99%)	182 (96%)	8 (4%)	25 33

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	26	ASN
1	A	31	LEU
1	A	79	ASN
1	A	85	GLN
1	A	119	ARG
1	В	32	LYS
1	В	88	LEU
1	В	108	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	ASN
1	A	85	GLN
1	A	96	GLN
1	В	14	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 oligosaccharides in this entry.

## 5.6 Ligand geometry (i)

2 ligands are modelled in this entry.



In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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| > 2 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	В	ond leng	$\operatorname{gths}$	В	ond ang	gles
IVIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GOL	В	201	-	5,5,5	0.29	0	5,5,5	0.29	0
2	GOL	A	200	-	5,5,5	0.29	0	5,5,5	0.29	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	В	201	-	-	0/4/4/4	-
2	GOL	A	200	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	200	GOL	O1-C1-C2-O2
2	A	200	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	$\operatorname{Res}$	Type	Clashes	Symm-Clashes
2	A	200	GOL	1	0

## 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,  $95^{th}$  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 $>$	$\#\mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q < 0.9
1	A	108/121 (89%)	0.74	15 (13%) 7 6	18, 39, 67, 76	2 (1%)
1	В	109/121 (90%)	0.82	9 (8%) 19 16	27, 44, 62, 76	0
All	All	217/242 (89%)	0.78	24 (11%) 12 9	18, 42, 65, 76	2 (0%)

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	121	SER	5.7
1	A	83	ALA	5.0
1	В	26	ASN	4.8
1	A	84	ASP	4.5
1	A	121	SER	4.2
1	В	58	GLU	3.7
1	A	82	ALA	3.6
1	A	79	ASN	3.3
1	A	105[A]	HIS	2.9
1	В	41	GLU	2.7
1	A	78	LYS	2.7
1	В	59	GLU	2.6
1	A	85	GLN	2.6
1	В	52	VAL	2.5
1	В	79	ASN	2.5
1	A	10	LYS	2.3
1	A	42	VAL	2.2
1	A	58	GLU	2.2
1	В	9	GLY	2.2
1	A	47	ASP	2.1
1	В	38	ILE	2.1
1	A	33	ARG	2.1
1	A	51	ASP	2.0
1	A	98	ALA	2.0



## 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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
2	GOL	В	201	6/6	0.33	0.74	49,50,51,51	6
2	GOL	A	200	6/6	0.80	0.24	49,50,52,52	6

### 6.5 Other polymers (i)

There are no such residues in this entry.

