

wwPDB Validation Report

PDB ID: NEW1
RCSB ID: RCSB123456
TITLE: Structure of interesting protein
AUTHOR(S): J.Crystal

Summary of PDB Entry Validation

The results of the validation of this PDB entry are shown below. Some major issues were raised during data processing which must be resolved prior to release of the file.

Geometry Validation

1. Atomic Clashes

The atom pairs listed below display physically unrealistic interatomic distances.

Chain	Atom	Res	Seq	Chain	Atom	Res	Seq	Symm_Code	Distance
A	CD	LYS	220 -	W	O	HOH	448	(1, 5, 5, 5)	Dist = 0.57
A	NH1(A)	ARG	228 -	W	O	HOH	343	(1, 5, 5, 5)	Dist = 0.82
A	NH1(B)	ARG	228 -	W	O	HOH	337	(1, 5, 5, 5)	Dist = 0.89
A	OE1	GLU	409 -	W	O	HOH	355	(1, 5, 5, 5)	Dist = 0.91
A	CZ (A)	ARG	228 -	W	O	HOH	343	(1, 5, 5, 5)	Dist = 0.93

2. Peptide Linkage

The C-N bonds listed below lie outside of the accepted range for the peptide bond (1.30-1.45). The main chain geometry needs to be corrected.

Residue A SER 183 and Residue A HIS 184 are not properly linked:
C-N bond distance is 0.89 A.
Residue A HIS 184 and Residue A TRP 185 are not properly linked:
C-N bond distance is 0.89 A.

3. Covalent Geometry

The following bonds deviate significantly from standard values.

Covalent bond lengths:

Deviation	Residue Name	Chain ID	Sequence Number	AT1	-	AT2	Bond Distance	Dictionary Value	Standard Deviation
1.482	GLU	A	149	CD	-	OE2	2.734	1.252	0.011
0.677	ARG	A	189	NE	-	CZ	2.003	1.326	0.013
0.993	LYS	A	190	CD	-	CE	2.501	1.508	0.025
-0.606	MET	A	238	SD	-	CE	1.168	1.774	0.056
0.640	GLN	A	408	CB	-	CG	2.161	1.521	0.027

Covalent bond angles:

Deviation	Residue Name	Chain ID	Sequence Number	AT1	-	AT2	-	AT3	Bond Angle	Dictionary Value	Standard Deviation
42.4	GLU	A	149	OE1	-	CD	-	OE2	165.7	123.3	1.2
-44.1	GLU	A	149	CG	-	CD	-	OE2	74.2	118.3	2.0
-64.1	HIS	A	184	CA	-	C	-	O	56.0	120.1	2.1
26.9	SER	A	183	CA	-	C	-	N	144.1	117.2	2.2
-46.5	SER	A	183	O	-	C	-	N	76.2	122.7	1.6
-27.7	GLN	A	187	CB	-	CG	-	CD	83.9	111.6	2.6
-41.1	GLN	A	408	CA	-	CB	-	CG	72.3	113.4	2.2

Sequence Validation

The following residues given in the deposited sequence are in conflict with the sequence given in coordinates. The coordinate sequence should match to the deposited sequence.

MISMATCH: ChainID=A ResNum=1 (? <----> GC1)

Biological Assembly

The biological assembly predicted by PISA is a dimer. This agrees with author's annotation.

Ligand Chemistry

Ligand chemistry has been checked against the Chemical Component Dictionary. The following is a summary.

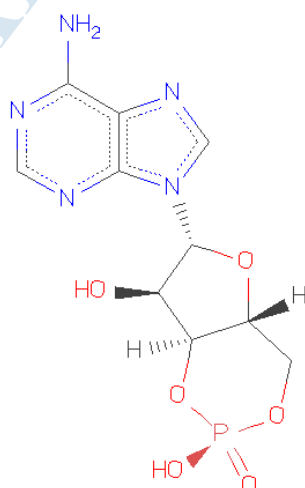
There are outstanding issues with following ligand(s) in the coordinates.

CRO A 100
GSH A 200
MH1 A 500

The geometry of the ligand extracted from deposited coordinates does not match the geometry of the corresponding ligand in the PDB chemical component dictionary.

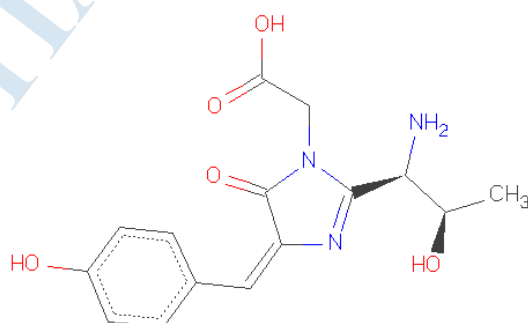
Identifier: CMP
Name: ADENOSINE-3',5'-CYCLIC-MONOPHOSPHATE
Synonym: CYCLIC AMP; CAMP
Formula: C10 H12 N5 O6 P

Type	Program	Version	Descriptor
SMILES	ACDLabs	10.04	<chem>O=P3(OCC4OC(n1c2ncnc(N)c2nc1)C(O)C4O3)O</chem>
InChI	InChI	1.02b	InChI=1/C10H12N5O6P/c11-8-5-9(13-2-12-8)15(3-14-5)10-6(16)7-4(20-10)1-19-22(17,18)21-7/h2-4,6-7,10,16H,1H2,(H,17,18)(H2,11,12,13)/t4-,6-,7-,10-/m1/s1/f/h17H,11H2
InChIKey	InChI	1.02b	IVOMOUWHDPKRL-BJEHYBLCDS
SMILES CANONICAL	CACTVS	3.341	<chem>Nc1ncnc2n(cnc12)[C@@H]3O[C@@H]4CO[P@](O)(=O)O[C@H]4[C@H]3O</chem>
SMILES	CACTVS	3.341	<chem>Nc1ncnc2n(cnc12)[CH]3O[CH]4CO[P](O)(=O)O[CH]4[CH]3O</chem>
SMILES CANONICAL	OpenEye OEToolkits	1.5.0	<chem>c1nc(c2c(n1)n(cn2)[C@H]3[C@@H]([C@H]4[C@H](O3)CO[P@](=O)(O4)O)O)N</chem>
SMILES	OpenEye OEToolkits	1.5.0	<chem>c1nc(c2c(n1)n(cn2)C3C(C4C(O3)COP(=O)(O4)O)O)N</chem>



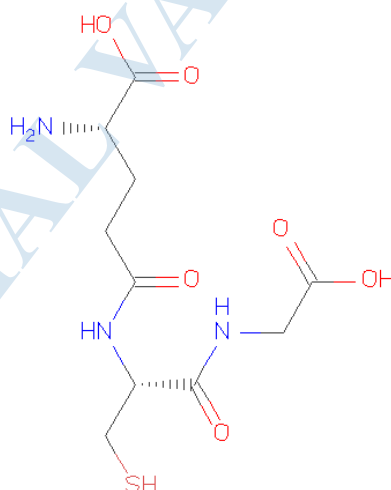
Identifier: CRO**Name:** {2-[(1R,2R)-1-amino-2-hydroxypropyl]-4-(4-hydroxybenzylidene)-5-oxo-4,5-dihydro-1H-imidazol-1-yl}acetic acid**Synonym:** PEPTIDE DERIVED CHROMOPHORE**Formula:** C15 H17 N3 O5

Type	Program	Version	Descriptor
SMILES	ACDLabs	10.04	<chem>O=C1C(N=C(N1CC(=O)O)C(N)C(O)C)=Cc2ccc(O)cc2</chem>
InChI	InChI	1.02b	InChI=1/C15H17N3O5/c1-8(19)13(16)14-17-11(15(23)18(14)7-12(21)22)6-9-2-4-10(20)5-3-9/h2-6,8,13,19-20H,7,16H2,1H3,(H,21,22)/b11-6-/t8-,13+/m1/s1/f/h21H
InChIKey	InChI	1.02b	UZCDFHUXSDKGEZ-QGLMYTBDI
SMILES CANONICAL	CACTVS	3.341	<chem>C[C@@H](O)[C@H](N)C1=NC(=C/c2ccc(O)cc2)C(=O)N1CC(O)=O</chem>
SMILES	CACTVS	3.341	<chem>C[CH](O)[CH](N)C1=NC(=Cc2ccc(O)cc2)C(=O)N1CC(O)=O</chem>
SMILES CANONICAL	OpenEye OEToolkits	1.5.0	<chem>C[C@H]([C@@H](C1=NC(=Cc2ccc(cc2)O)C(=O)N1CC(=O)O)N)O</chem>
SMILES	OpenEye OEToolkits	1.5.0	<chem>CC(C(C1=NC(=Cc2ccc(cc2)O)C(=O)N1CC(=O)O)N)O</chem>



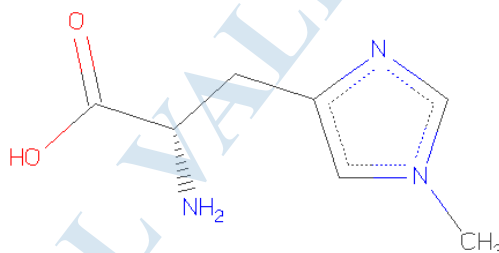
Identifier: GSH
Name: GLUTATHIONE
Formula: C10 H17 N3 O6 S

Type	Program	Version	Descriptor
SMILES	ACDLabs	10.04	<chem>O=C(NCC(=O)O)C(NC(=O)CCC(C(=O)O)N)CS</chem>
InChI	InChI	1.02b	InChI=1/C10H17N3O6S/c11-5(10(18)19)1-2-7(14)13-6(4-20)9(17)12-3-8(15)16/h5-6,20H,1-4,11H2,(H,12,17)(H,13,14)(H,15,16)(H,18,19)/t5-,6-/m0/s1/f/h12-13,15,18H
InChIKey	InChI	1.02b	RWSXRVCMGQZWBV-VSCBVDDUDG
SMILES CANONICAL	CACTVS	3.341	<chem>N[C@@H](CCC(=O)N[C@@H](CS)C(=O)NCC(O)=O)C(O)=O</chem>
SMILES	CACTVS	3.341	<chem>N[CH](CCC(=O)N[CH](CS)C(=O)NCC(O)=O)C(O)=O</chem>
SMILES CANONICAL	OpenEye OEToolkits	1.5.0	<chem>C(CC(=O)N[C@@H](CS)C(=O)NCC(=O)O)[C@@H](C(=O)O)N</chem>
SMILES	OpenEye OEToolkits	1.5.0	<chem>C(CC(=O)NC(CS)C(=O)NCC(=O)O)C(C(=O)O)N</chem>



Identifier: MH1
Name: 1-methyl-L-histidine
Formula: C7 H11 N3 O2

Type	Program	Version	Descriptor
SMILES	ACDLabs	10.04	<chem>O=C(O)C(N)Cc1ncn(c1)C</chem>
InChI	InChI	1.02b	InChI=1/C7H11N3O2/c1-10-3-5(9-4-10)2-6(8)7(11)12/h3-4,6H,2,8H2,1H3,(H,11,12)/t6-/m0/s1/f/h11H
InChIKey	InChI	1.02b	BRMWTNUJHUMWMS-LBISWHJPDF
SMILES CANONICAL	CACTVS	3.341	<chem>Cn1cnc(C[C@H](N)C(O)=O)c1</chem>
SMILES	CACTVS	3.341	<chem>Cn1cnc(C[CH](N)C(O)=O)c1</chem>
SMILES CANONICAL	OpenEye OEToolkits	1.5.0	<chem>Cn1cc(nc1)C[C@@H](C(=O)O)N</chem>
SMILES	OpenEye OEToolkits	1.5.0	<chem>Cn1cc(nc1)CC(C(=O)O)N</chem>



Summary of Structure Factor Validation

1. Structure Factor - high R-factors, low correlation

SFcheck shows high R-factors and low correlation coefficient.

2. Structure Factor - incomplete data

The deposited structure factor file is incomplete. The SF file should include h, k, l, F, SigmaF (and/or I and Signal) and test flags.

3. Structure Factor - inconsistency in number of reflections

The number of reflections reported for the refinement exceeds the number of reflections listed in the SF file.

Resolution	
High Resolution (Author reported)	1.85
High Resolution (Calculated by SFCHECK)	1.75
High Resolution (Calculated by REFMAC)	1.750
Low Resolution (Author reported)	43.44
Low Resolution (Calculated by SFCHECK)	43.52
Low Resolution (Calculated by REFMAC)	43.526

Crystal data	
Space group	P 21 21 2
Total number of reflections	43103
Number of reflections used	40931
Completeness of data	100.0

R-factors	
R-factor (Author reported)	0.150
R-factor (Calculated by SFCHECK)	0.196
R-factor (Calculated by REFMAC)	0.1819
Free R-factor (Author reported)	0.188
Free R-factor (Calculated by SFCHECK)	0.195
Free R-factor (Calculated by REFMAC)	0.1820

Structure quality	
Average Real space R-factor (Deviation) (Calculated by SFCHECK)	0.0808
Average Real space R-factor (Deviation) (Calculated by MAPMAN)	0.0940

Average Real-space correlation coefficient (Deviation) (Calculated by SFCHECK)	0.9834
Average Real-space correlation coefficient (Deviation) (Calculated by MAPMAN)	0.966
Average Occupancy-weighted avg temperature factor (Deviation)	29.42

Wilson statistics	
Wilson B-factor	31.64
Wilson Scale	0.12

Padilla-Yeates statistics	
Padilla-Yeates $\langle L \rangle$	0.499
Padilla-Yeates $\langle L^*L \rangle$	0.333