



C571 Chemical Information Technology Exercise 4. Crystallography and Visualization

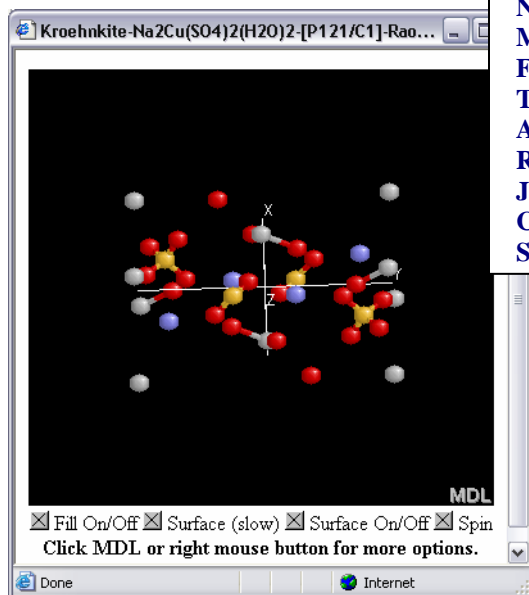
1. Locate the "demo" version of the ICSD (Inorganic Crystal Structure Database) at:

<http://www.fiz-informationsdienste.de/en/DB/icsd/produkte.html>

Download and look for several simple structures, e.g., copper sulfate, table salt, etc.

Copper Sulfate:

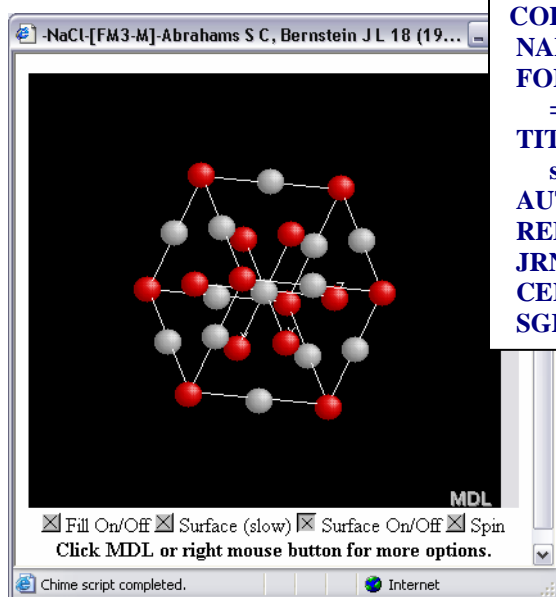
Reference 1



COL ICSD Collection Code 15434 (DATE=R961001/U 0 REL= 286/ 1)
NAME Disodium copper sulfate dihydrate
MINR Kroehnkite
FORM Na₂ Cu (S O₄)₂ (H₂ O)₂ = CU1 H4 NA2 O10 S2
TITL Die Verfeinerung der Kristallstruktur von Kroehnkit, Na₂ Cu (SO₄)₂ (H₂O)₂
AUT Rao R
REF ACCRA 14 (1961) P. 738-743
JRNL Acta Crystallographica (1,1948-23,1967)
CELL A=5.805 B=12.675 C=5.52 α =90.0 β =108.6 γ =90.0
SGR P 21/C (14)



Table salt (NaCl):

Reference 1



COL ICSD Collection Code 18189 (DATE=R961001/U 0 REL= 382/ 1)
NAME Sodium chloride
FORM Na Cl
= CL1 NA1
TITL Accuracy of an automatic diffractometer. measurement of the sodium chloride structure factors
AUT Abrahams S C, Bernstein J L
REF ACCRA 18 (1965) P. 926-932
JRNL Acta Crystallographica (1,1948-23,1967)
CELL A=5.62 B=5.62 C=5.62 α =90.0 β =90.0 γ =90.0
SGR F M 3 M (225)



2. Locate the PDB and look up snake toxins whose structures have been determined by X-Ray crystallography. Save one or more of the .pdb files.

 1B41	 Deposited: 05-Jan-1999 Exp. Method: X-ray	EXPLORE
Diffraction Resolution: 2.76 Å		

Title Human Acetylcholinesterase Complexed With Fasciculin-II, Glycosylated Protein

Classification Hydrolase/Toxin

Compound Mol_Id: 1; Molecule: Acetylcholinesterase; Chain: A; Fragment: Single Domain; Synonym: Huache H-Subunit; Ec: 3.1.1.7; Engineered: Yes; Mutation: Yes
Mol_Id: 2; Molecule: Fasciculin-2; Chain: B; Fragment: Single Domain; Synonym: Acetylcholinesterase Toxin F-Vii; Mutation: Yes

 1F8U	 Deposited: 05-Jul-2000 Exp. Method: X-ray	EXPLORE
Diffraction Resolution: 2.90 Å		

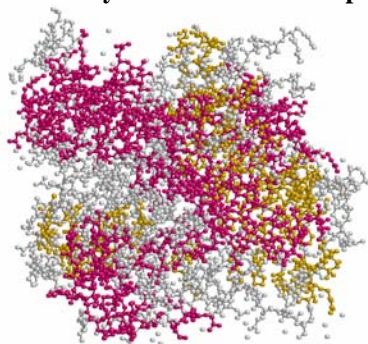
Title Crystal Structure Of Mutant E202Q Of Human Acetylcholinesterase Complexed With Green Mamba Venom Peptide Fasciculin-II

Classification Hydrolase/Toxin

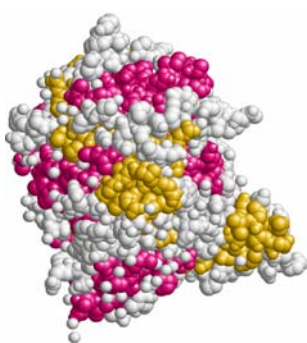
Compound Mol_Id: 1; Molecule: Acetylcholinesterase; Chain: A; Ec: 3.1.1.7; Engineered: Yes
Mol_Id: 2; Molecule: Fasciculin II; Chain: B

3. Download RasMol and use it to prepare figures of the snake toxins in a variety of formats (ball-and-stick, space filling, etc.). For sources of Rasmol, see: *Note example of wireframe not included (default view)* http://www.indiana.edu/~cheminfo/ca_mvts.html

A. Human Acetylcholinesterase Complexed With Fasciculin-II, Glycosylated Protein



Ball and Stick



Space Fill



Ribbons

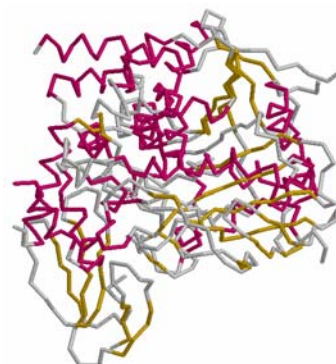
B. Crystal Structure Of Mutant E202Q Of Human Acetylcholinesterase Complexed With Green Mamba Venom Peptide Fasciculin-II



Cartoon



Strands



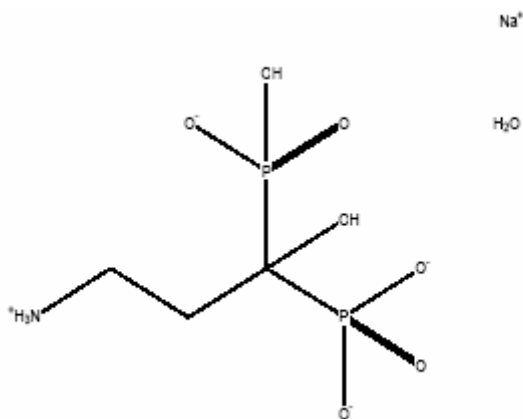
Backbone

4. Use the ConQuest version of the Cambridge Structural Database and do a few of the exercises found at: http://www.indiana.edu/~cheminfo/C472/CSD_exercises.html
HINT: The tutorials that are found under the "Help" drop-down menu of ConQuest correspond to the CSD exercises.
Notify Dr. Milosevich or Dr. Wiggins when you have completed this exercise, but no later than December 2.

See Attached.....

Cambridge Structural Database Tutorials from the CSD Workshop, New Orleans, 4/24/2003

1. Find the structure published by D. Vega, D. Fernandez and J. A. Ellena in *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, 58, m77, (2002).



2. Find the structure of 8-Methylpyridoxatin.

ADUWAX

Reference: M.M.Wagenaar, D.M.Gibson, J.Clardy (2002)
Organic Letters, 4, 671

Formula: C₁₆H₂₃N₁O₃C₁H₄O₁

Compound Name: 8-Methylpyridoxatin methanol solvate

Synonyms: Cordopyridone A methanol solvate

Space Group: P21 Cell: a 5.172(1) b 23.057(4) c 9.307(2)
Space Group No.: 4 (A, γ) α 90.00 β 95.73(1) γ 90.00

R-Factor (%): 3.81 Temperature(K): 173 Density(g/cm³): 1.176

IBOFUA01

Reference: M.M.Wagenaar, D.M.Gibson, J.Clardy (2002)
Organic Letters, 4, 671

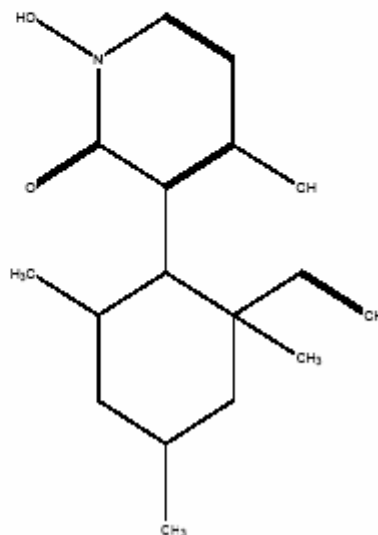
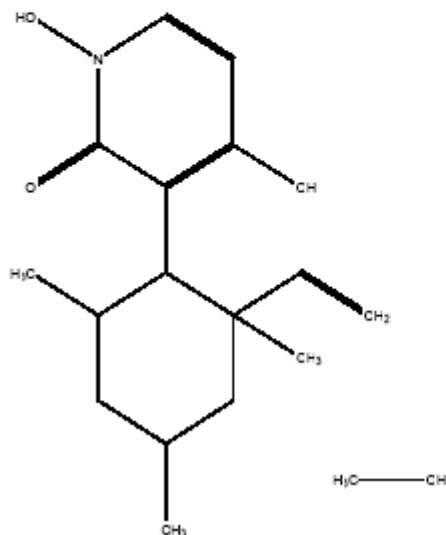
Formula: C₁₆H₂₃N₁O₃

Compound Name: 8-Methylpyridoxatin

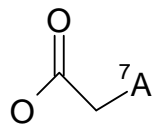
Synonyms: Cordopyridone B

Space Group: P21 Cell: a 13.055(1) b 9.798(1) c 13.308(1)
Space Group No.: 4 (A, γ) α 90.00 β 112.06(0) γ 90.00

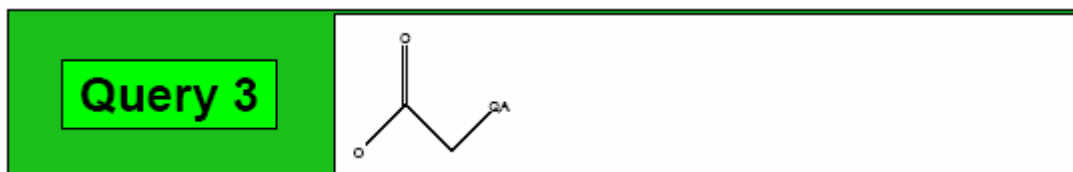
R-Factor (%): 3.43 Temperature(K): 173 Density(g/cm³): 1.164



3. Search for the following fragment:

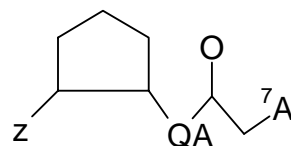


where 7A = F, Cl, Br, or I.



1303 Hits

4. Search for the following fragment:

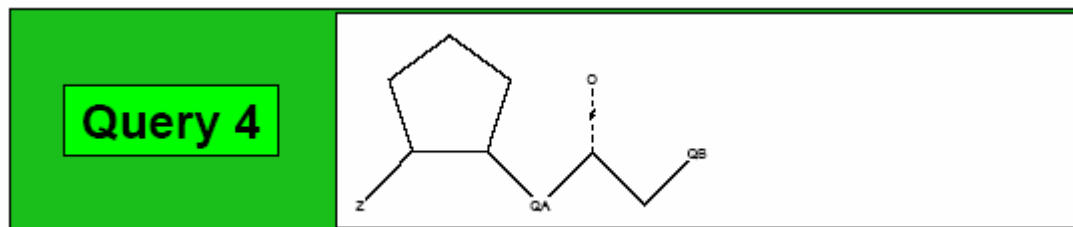


where 7A = F, Cl, Br, or I

QA = C, N, or O

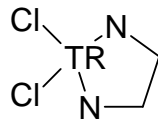
Z = any atom except H

Bond C....O can either be single or double.



174 Hits

5. Define and save important geometric parameters for the fragment:

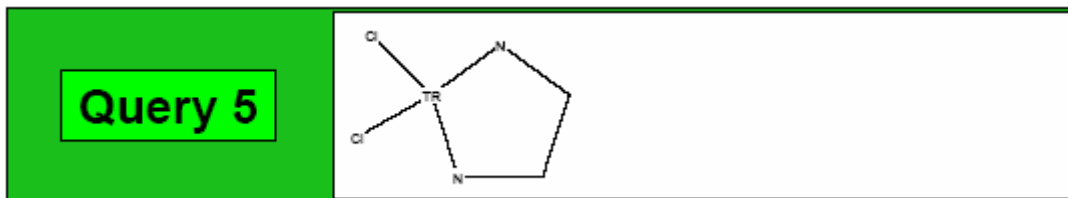


where TR = any transition metal and the coordination number of TR is exactly 4.
The parameters that will be saved are:

- TR-Cl, TR-N, C-C bond lengths
- Cl-TR-Cl and N-TR-N bond angles
- N-C-C-N torsion angle
- Element symbol of TR.

without parameters 110 hits

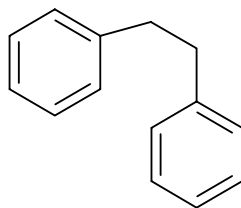
with froze at 2



Parameter Values for Query: 5

<i>Refcode</i>	<i>ANG1 (A)</i>	<i>ANG2 (A)</i>	<i>CC (D)</i>	<i>TOR1 (T)</i>	<i>TRCL1 (D)</i>	<i>TRCL2 (D)</i>	<i>TRN1 (D)</i>	<i>TRN2 (D)</i>
ACETIL	93.665	83.287	1.527	-50.506	2.307	2.321	2.02	2.036
ACHPTA	93.578	82.705	1.496	51.554	2.286	2.291	2.061	2.026
ACHPTA	91.488	84.801	1.517	-54.911	2.283	2.291	2.014	2.042
ACHPTA	91.397	82.283	1.513	-47.006	2.297	2.309	2.014	2.017

6. Search for the following fragment:

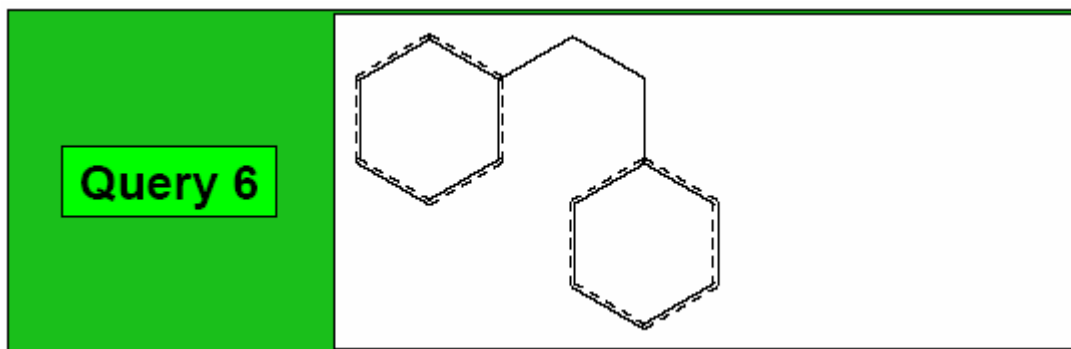


where the two central carbons are acyclic (not part of a ring) and are bonded to 1 or 2 hydrogen atoms. The two phenyl groups will be required to be cis (i.e., Cring-C-C-Cring torsion angle between -90 and $+90$ degrees). The following parameters will be saved:

- Distance between the centroids of the phenyl rings
- Angle between the planes of the phenyl rings.

without parameters 281

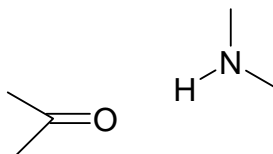
with parameters froze at 5



Parameter Values for Query: 6

Refcode	DIST1 (D)	TOR1 (T)
ACULIT	4.891	-71.843
AFADQ	4.721	61.297
BALRAH	4.386	-55.294
BALSIQ	4.333	-56.768
BALTOX	4.497	-55.394
BALTOX	4.292	-57.348

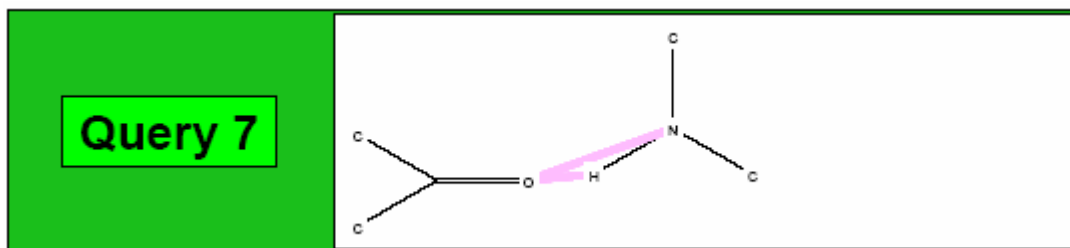
7. Search for intermolecular contact:



where N is bonded to exactly 3 atoms and there is an intermolecular contact < 2.1 angstroms between O and H. The following parameters will be saved:

- O...H, O...N nonbonded distances
- O...H-N and C-O...H angles
- O and H atom labels
- Orthogonal coordinates of the hit fragments.

without parameters 1768 (even no defined contact)
with froze at 3



Parameter Values for Query: 7

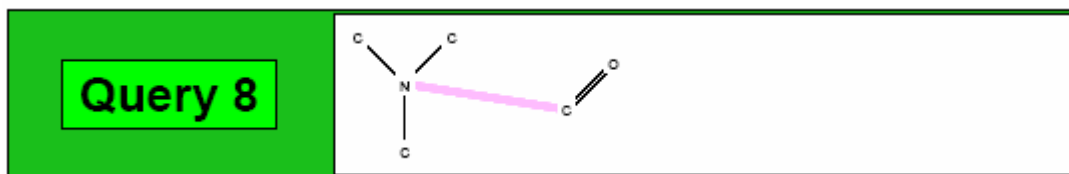
Refcode	ANG1 (A)	ANG2 (A)	DIST1 (D)	DIST2 (D)	LAB1 (L)	LAB2 (L)	No. of Coordinates (I)
ACAZFE	155.39	154.907	2.048	2.826	O1%	H1	26
ACMPIM10	176.816	143.751	2.067	2.871	O1%	H6	27
ACULEP	168.02	161.66	2.077	2.907	O2%	H4	104
ACULEP	170.93	140.499	1.931	2.881	O6	H3	104

8. Search for the following pharmacophore:



where the distance between the cationic nitrogen and the center of the carbonyl bond is in the range 5.0 - 5.8 angstroms. This is the distance range required between these two functional groups when acetylcholine and its analogs bind to the enzyme acetylcholinesterase.

without 2466
with froze at 8



Parameter Values for Query: 8

<i>Refcode</i>	<i>DIST1 (D)</i>
ABIRUY	5.314
ABIRUY	5.055
ABIRUY	5.15
ABISAF	5.298
ABISAF	5.058
ABISAF	5.171

9. Find di-molybdenum structures which are air-sensitive.

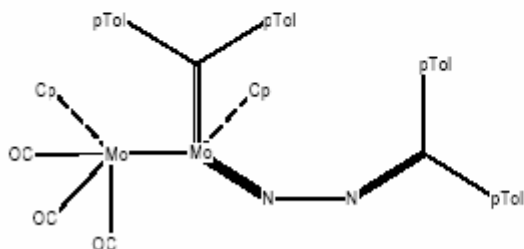
45 hits

Query 9	Text Search air-sensitive Text Search di-molybdenum
----------------	--

First 2 Hits

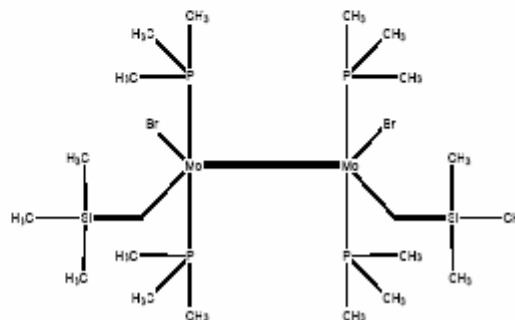
BEMXUM10

Reference: M.D. Curtis, L. Messerle (1987) *Organometallica*, 6, 1713
Formula: $C_{40} H_{38} Mo_2 N_2 O_5$
Compound Name: Tricarbonyl-bis(η^5 -cyclopentadienyl)-di-*p*-tolylmethylene-(*N*-(4-*p*-tolylmethylfrazino-*N'*)-di-molybdenum
Space Group: P-1 Cell: a 12.247(8) b 14.483(7) c 11.486(5)
Space Group No.: 2 α 113.22(3) β 95.85(4) γ 90.32(4)
R-Factor (%): 5.30 Temperature(K): 295 Density(g/cm³): 1.470



BOPXUZ10

Reference: K.J. Ahmed, M.H. Chisholm, J.C. Huffman (1985) *Organometallica*, 4, 1168
Formula: $C_{39} H_{58} Br_2 Mo_2 P_4 Si_2$
Compound Name: bis[(Trimethylsilyl)methylene]-tetra(iso(trimethylphosphine)-dibromo-di-molybdenum
Space Group: A2/a Cell: a 20.481(11) b 9.918(4) c 20.468(11)
Space Group No.: 15 α 90.00 β 120.67(2) γ 90.00
R-Factor (%): 2.30 Temperature(K): 108 Density(g/cm³): 1.540



10. Find structures that:

- contain 1 or 2 osmium atoms and > 6 oxygen atoms
- do not belong to space group $P2_1/c$
- were determined either by neutron diffraction or with a R-factor < 0.06.

74 hits

match

Query 10

do not match

Query 11

match at least one of

Query 12

Query 13

Query 10	Formula $Os_{1-2}O_{>6}$ Search on: each molecule in turn Allow other elements: Yes
Query 11	Space Group $P2_1/c$ (and equivalent space groups)
Query 12	Radiation Neutron
Query 13	R-factor < 0.06 fractional

First 2 hits

BALYIW

Reference: L. Buzio, M. Green, J.A.K. Howard, B. Hessner, J.C. Jeffrey, R.M. Miles, F.G.A. Stone, P. Woodward (1981) *Chem Commun*, 1101

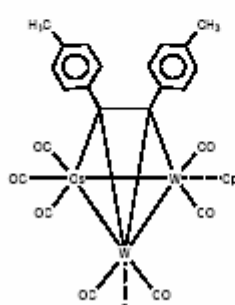
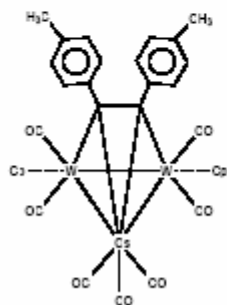
Formula: $0.5(C_{33}H_{24}O_7Os_1W_2)0.5(C_{33}H_{24}O_7Os_1W_2)$

Compound Name: $(\eta^3-O^1, O^1, \eta^2-W/Cs)(\eta^3-W/Cs)-1,2$ -bis(*p*-Tolyl-ethylene)-tricarbonyl-osmium-bis(dicarbonyl-(η^5 -cyclopentadienyl)-tungsten) $(\eta^3-O^1, O^1, \eta^2-W/Cs, W)-1,2$ -bis(*p*-tolyl-ethylene)-tricarbonyl-osmium-bis(dicarbonyl-(η^5 -cyclopentadienyl)-tungsten)

Space Group: P21 Cell: a 10.248(11) b 15.169(25) c 16.637(16)

Space Group No.: 4 (A_1) α 90.00 β 101.49(8) γ 90.00

R-Factor (%): 4.50 Temperature(K): 295 Density(g/cm^3): 2.385



BALYIW10

Reference: L. Buzio, M. Green, B. Hessner, J.A.K. Howard, J.C. Jeffrey, F.G.A. Stone (1983) *J. Chem. Soc., Dalton Trans*, 519

Formula: $0.5(C_{33}H_{24}O_7Os_1W_2)0.5(C_{33}H_{24}O_7Os_1W_2)$

Compound Name: $(\eta^3-O^1, O^1, \eta^2-W/Cs)(\eta^3-W/Cs)-1,2$ -bis(*p*-Tolyl-ethylene)-tricarbonyl-osmium-bis(dicarbonyl-(η^5 -cyclopentadienyl)-tungsten) $(\eta^3-O^1, O^1, \eta^2-W/Cs, W)-1,2$ -bis(*p*-tolyl-ethylene)-tricarbonyl-osmium-bis(dicarbonyl-(η^5 -cyclopentadienyl)-tungsten)

Space Group: P21 Cell: a 10.248(1) b 15.169(25) c 16.637(16)

Space Group No.: 4 (A_1) α 90.00 β 101.49(8) γ 90.00

R-Factor (%): 4.50 Temperature(K): 295 Density(g/cm^3): 2.385

