A challenging EXAFS analysis problem

Bruce Ravel

Synchrotron Science Group National Institute of Standards and Technology & Beamline for Materials Measurements National Synchrotron Light Source II

ASEAN Workshop on X-ray Absorption Spectroscopy Synchrotron Light Research Institute

June 2-4, 2014

Á

Metal sensors	Experime	nt	DNA	Model building	The fit	Post mortem
			Сор	yright		
This document	is copyright © 2010	-2015	Bruce Ravel.			
			C			
This work is lic http://creat: 94305, USA.	ensed under the Cr ivecommons.org/1	eative (icense	Commons Attribution-ShareA es/by-sa/3.0/ or send a le	like License. To view a copy of t tter to Creative Commons, 559 N	his license, visit Jathan Abbott Way, Stanford	1, California
	You are free:	0000	to Share — to copy, distrib to Remix — to adapt the w to make commercial use of	oute, and transmit the work ork the work		
Under the follo	wing conditions:	•	Attribution – You must attri any way that suggests that Share Alike – If you alter, only under the same, simila	bute the work in the manner spe they endorse you or your use o transform, or build upon this wou or or a compatible license.	cified by the author or licen: f the work). rk, you may distribute the re	sor (but not in sulting work
With the und	erstanidng that:	0	Waiver – Any of the above Public Domain – Where th that status is in no way aff	conditions can be waived if you e work or any of its elements is i ected by the license.	get permission from the cop n the public domain under a	yright holder. Ipplicable law,
		٩	Other Rights – In no way a Vour fair dealing The author's mora Rights other perso publicity or privac	are any of the following rights af or fair use rights, or other applic I rights; ons may have either in the work ry rights.	fected by the license: able copyright exceptions ar itself or in how the work is	nd limitations; used, such as
		٩	Notice – For any reuse or	distribution, you must make clea	r to others the license terms	of this work.
This is a humar	n-readable summary	of the	Legal Code (the full license).		

Transport of metal contaminants in the environment

There are numerous natural and man-made point sources of toxic metals which find their way into water systems used for human and agricultural applications.



The safe use of water requires monitoring and eventual remediation of bioavailable metal species.

image from 'a http://lightsources.org, Credit: Argonne National Laboratory

Real-time, field-ready sensors

Sophisticated laboratory and synchrotron methods exist to detect and speciate water contaminants at very low concentrations. The real-world task of environmental monitoring requires a fast, flexible, sensitive, selective method of detecting contaminants *in the field*.

Fast Obtain results while still in the field

Flexible Easy to carry and easy to use in the field

Sensitive Detect contaminant concentrations below regulated human health hazard levels

Selective Respond to the target metal but not to other metals

Real-time, field-ready sensors

Sophisticated laboratory and synchrotron methods exist to detect and speciate water contaminants at very low concentrations. The real-world task of environmental monitoring requires a fast, flexible, sensitive, selective method of detecting contaminants *in the field*.



We want Spock's tricorder!

Fast Obtain results while still in the field

- Flexible Easy to carry and easy to use in the field
- Sensitive Detect contaminant concentrations below regulated human health hazard levels
- Selective Respond to the target metal but not to other metals

Catalytic DNA-based sensors



The sensor has a receptor, a cleavage site, and paired fluorophore and quencher.

J. Liu, et al. A catalytic beacon sensor for uranium with parts-per-trillion sensitivity and millionfold selectivity PNAS, 104:7 (2007) 2056-2061 To DOI: 10.1073/pnas.0607875104

Building a sensor device

These DNA sensors can be incorporated into a hand-held device. Water is dropped onto an array of sensors and read using photodiodes.







Wells containing selective DNAzyme sensors

DNA-based Hg sensor

U.S. EPA limit on Hg in water is 10 nM (2 ppb)

The DNA-based sensor for Hg has a detection limit of 2.4 nM

Questions:

- How and where does the metal bind?
- Under what conditions does the metal remain bound to the DNA?
- How many binding sites are there on a sensor?
- Do different metals behave differently?
- Can DNAzymes be designed more rationally?
- And, of course, what can XAS tell us about any of these questions (keeping in mind the very local nature of the XAS measurement)?

J. Liu and Y. Lu. Rational Design of "Turn-On" Allosteric DNAzyme Catalytic Beacons for Aqueous Mercury Ions with Ultrahigh Sensitivity and Selectivity, Angew. Chemie, 46:40 (2007) 7587–7590 * DOI: 10.1002/anie.200702006

A challenging EXAFS analysis problem

XAS measurements

Solutions:

- 50 mM cacodylic acid as a buffer
- 100 mM NaClO₄ to maintain pH=6.10
- glycerol to promote glassification upon freezing

Samples:

Control 15 mM Hg Sample 3 mM Hg with 3 mM DNA Sample with excess Hg 6 mM Hg with 3 mM DNA

Measure EXAFS at 10 K

Experiment		

Cryostat



Displex cryostat at APS 20BM.

- He exchange gas
- 10 mm wide opening for beam
- $\sim 12 \text{ mm}$ wide inner shroud
- Fluorescence measured through hole on side with a Ge detector
- At that time, 20BM did *not* have a focusing mirror

Here is the fluorescence spectrum:

🔁 MCA Display v6a			
Live Time	1.26+4-		ROI Events/Live Time
Real Time	1.1E+4-		Hg 151
85.68 sec	1.0E+4-	1 1	AS 1340
Spectrum Events	9.0E+3-		Fe ka
1326096	8.0E+3-		
Events/Live Time	7.0E+3-		
15813.56	6.0E+3-		
Main Peak	5.0E+3-		Rh 1994
Energy 60.0 ev	4.0E+3-		TVBa
chan. # 3	3.06+3-		Ba/Or B
Element 7 🔻 << >>	2.06+3-	11	0u 30
Detector DUP 12 Element	1.0E+3-		total 77776
	0.00+0-, (*,)		
Save Data	100 2500 5000 750	0 10000 12500 15000 17500	20460
	Cursor 0 [9911.3]0000.0	Plot Full Spectrum	Div A
	Farser I from hyp man	VS Energy	
Start/Stop/Erase/Read	rt Acquiring Erase Data Real	d Data Collecter Set up Ri	II's Set up Sums EXIT

The Hg L α peak is the tiny thing near the green line.

The neighboring peak is vastly larger!

Here is the fluorescence spectrum:



The Hg L α peak is the tiny thing near the green line.

The neighboring peak is vastly larger!

What's cacodylic acid?

MCA Display v6a			_ 🗆 ×
Live Time	1.26+4-	1	ROE Events/Live Time
Real Time	1.1E++-		Hg 151
85.68 sec	1.0E+4-	Λ.	As 1340
Spectrum Events	9.0E+3-		Fe ka
1326096	8.0E+3-	1	<u>00</u>
Events/Live Time	7.0E+3-		
15813.56	6.0E+3-		
Main Peak	5.0E+3-		20 10 10 10 10 10 10 10 10 10 10 10 10 10
Energy 60.0 eV	4.0E+3-		TURN
Chan. # 3	3.0E+3-		Ballor B
Element 7 🔻 << >>	2.06+3-	لم)	0 0
Detector DXP 12 Element	1.06+3-	11	total 7776
	0.0E+0-	10000 10000 10000 20400	
Save Data	Cursor 0 19911.30000.00 2	ot Full Spectrum	•
Start/Stop/Enzo/Read act on all elements	t Acquiring Erase Data Read Data	Calibrate Set up ROI's Set u	p Sums EXIT

Here is the fluorescence spectrum:

The Hg L α peak is the tiny thing near the green line.

The neighboring peak is vastly larger!



The big peak is As $K\alpha$ (~10.5 keV), our Hg L α (~10 keV) peak is on its shoulder.



The samples were packaged back at the University of Illinois and were about 15 mm by 3 mm.

We had to put the samples in the cryostat upright and slit the beam down to ~ 1 mm.



The samples were packaged back at the University of Illinois and were about 15 mm by 3 mm.

We had to put the samples in the cryostat upright and slit the beam down to ~ 1 mm.

Plan ahead!

Forgetting about the details leads to much worse data!

Our main sample



This poor data is due to low concentration, small beam, and large background from the As.

We measured 42 scans, taking about 22 hours.



Sample and control



Chemistry has certainly happened.

The control is clearly Hg in some kind of aqueous form.

The sample with DNA is clearly different from the control.

Is all Hg taken up by the DNA?

To answer this, we measured a sample with excess Hg.



Let's go do some linear combination fitting. (Note the isosbestic points.)

First question

Is all Hg taken up by the DNA?

To answer this, we measured a sample with excess Hg.



Let's go do some linear combination fitting. (Note the isosbestic points.)



Yes, all the Hg is taken up by the DNA. 47(1)% sample + 53(1)% control



2D and 3D representations

The 2D figures on the previous page were generated from the canonical SMILES strings:

 Adenisine
 c1=Nc2=c(c(=N1)N)N=CN2C3C(c(c(03)COP(=0)(0)0)0)0

 Thymidine
 cc1=cN(c(=0)Nc1=0)c2Cc(c(02)COP(=0)(0)0)0)0

 Guanosine
 c1=Nc2=c(N1C3C(c(c(03)COP(=0)(0)0)0)Nc(=Nc2=0)N

 Cytidine
 c1=cN(c(=0)N=C1N)c2c(c(c(02)COP(=0)(0)0)0)0

Neat! But we need 3D structures to run FEFF...

Structure from PubChem

Conclusion N Additional base with the main of the m		Thymidine Monophospha	te - PubChem - Mozilla Firefox	~ *
Image: Contract of the stage of the st	🔢 Google Calendar 🛛 🗙 🧘 A	SEAN Workshop on X 🗴 🖸 bruceravel/demeter 🛛 🗙	🖬 Rational Design of "Tur 🗙 🔘 Thymidine Monophosp 🛪	•
Example Events	🔶 🕲 pubchem.ncbi.nlm. nih.gov /summar	ry/summary.cgi?cid=9700&loc=ec_rcs#	🚭 🕶 😋 🔯 🕶 mal of synchrotron radiation 🚭	▶ ↓ 合 ☆ 自 ◎ ▼ 魚 目
 ▶ Control Company ▶ Control C	S NCBI			
 ■ Image: A state of the state of	PubCompound PubCherr	Compound	Search	Help
Thymolifie Monophosphate 20 Statistics Backwares Krimphils acts Muhyli ABP, CHIP. Thymidre 5 physphers. Thymidre 5 physpheric acts As Muhyli ABP, CHIP. Thymidre 5	Compound Summary for: CID 9700		응 값 응 책임 70% 80%	» Links and Related Information
Date of Contents \$ Brow subcontext fills statistication Rainade Effects as Truth Landaue 20 Sinclus 30 Continues Continues Continues Statistication Rainade Effects as Truth Landaue 10 Sinclus 10 Sinclus Continues Continues Remediate Effects as Truth Landaue Implicit Mitigatistication Remediations and Physical Populations Classification Implicit Mitigatistication Remediations Implicit Mitigatistication Remediations Implicit Mitigatistication Remediations Expand all sub-sections Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Expand all sub-sections Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Linear Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Linear Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Linear Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Remediation Remediation Linear Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Remediation Remediation Linear Implicit Mitigatistication Remediation Implicit Mitigatistication Remediation Remediation Remediation Remediation <	Thymidine Monophosphate Also known as: 5'Thymidylic acid, thym Motecular Formula: C10H15N2O8P M 5-Thymidylic acid. A thymine nucleotide o	nidylic acid, 5-Methyl-dUMP, dTMP, Thymidine 5'-phosphate, Thymio olecular Weight: 322.209462 InCAIKey: 6YO2YWVXFINGLUX containing one phosphate group esterfied to the decovritose	20 SDF Stypesy dime 5'sphosphoric acid, Deoxy TMP 12/20RFOSA-N 006by. From: MeSH 30 SDF: Save	
Biological Transformation BioActivity Data Links Commutal and Physical Physicing Image: Compound of the	Table of Contents & Show subcontent Identification Related Records Biomedical Effects and Toxicity Learnary Patents Reconcention International and Bothuman	tiles 20 Structure 30 Conformer	b	Properties Compound ID: 9700 Molecular Weight: 322 208462 (gmol) Molecular Formula: C ₁₀ H ₁₅ N ₂ O ₆ P XLog ³³ - 2.8 H-Bond Donor: 4 H-Bond Acceptor: 8
H=0 H=0 A A Identification H=0 A A A A A A A A A A A A A	Biological Text Results Biological Text Results Classification Chemical and Physical Properties Expand all sub-sections			BioActivity Data Links This Compound with Similar Compounds with Similar Conformers
Identification		H-0 10 10 10 10 10 10 10 10 10 10 10 10 10		Related Compounds Same, Connectity (14) Same, Stereochemistry (2) Same, Isotopes (13) Similar Compounds (1048) Similar Conformers (977) Veer
	Identification			

http://pubchem.ncbi.nlm.nih.gov/

Cartesian coordinates: 3D SDF file

9700

-OEChem-05141416293D

36 37	0	1	0	0	0	0 0	999	V20	00										
-3.55	15	-1.8	5175	;	0.	1599	Ρ	0	0	0	0	0	0	0	0	0	0	0	0
-0.43	89	1.3	3396	5	1.	0202	0	0	0	0	0	0	0	0	0	0	0	0	0
-0.91	01	4.	1569)	-0.	0812	0	0	0	0	0	0	0	0	0	0	0	0	0
-2.75	52	-0.	1874	ł	0.	6247	0	0	0	0	0	0	0	0	0	0	0	0	0
3.61	73	1.	7470)	0.	3907	0	0	0	0	0	0	0	0	0	0	0	0	0
3.83	78	-2.8	8022	2	-0.	2452	0	0	0	0	0	0	0	0	0	0	0	0	0
-2.54	75	-2.1	2163	3	-0.	8977	0	0	0	0	0	0	0	0	0	0	0	0	0
-4.72	67	-0.5	9241		-0.	7790	0	0	0	0	0	0	0	0	0	0	0	0	0
-4.01	97	-2.4	4002	2	1.	2798	0	0	0	0	0	0	0	0	0	0	0	0	0
1.61	13	0.1	5684	ł	0.	1973	Ν	0	0	0	0	0	0	0	0	0	0	0	0
3.71	27	-0.	5224	ł	0.	0726	Ν	0	0	0	0	0	0	0	0	0	0	0	0
-1.01	01	2.1	8736	5	-0.	6948	С	0	0	2	0	0	0	0	0	0	0	0	0
-1.56	99	1.3	8660)	0.	2995	С	0	0	1	0	0	0	0	0	0	0	0	0
0.37	33	2.3	3378	3	-0.	9829	С	0	0	0	0	0	0	0	0	0	0	0	0
0.77	01	1.	7196	5	0.	3478	С	0	0	1	0	0	0	0	0	0	0	0	0
-2.27	96	0.0	6993	3	-0.	3750	С	0	0	0	0	0	0	0	0	0	0	0	0
1.01	12	-0.	6708	3	0.	0146	С	0	0	0	0	0	0	0	0	0	0	0	0
3.01	76	0.0	6816	5	0.	2323	С	0	0	0	0	0	0	0	0	0	0	0	0
1.67	92	-1.8	8209	•	-0.	1381	С	0	0	0	0	0	0	0	0	0	0	0	0
3.16	56	-1.	7831		-0.	1119	С	0	0	0	0	0	0	0	0	0	0	0	0
1.01	30	-3.	1449	•	-0.	3336	С	0	0	0	0	0	0	0	0	0	0	0	0
-1.62	78	2.	9841		-1.	5911	Η	0	0	0	0	0	0	0	0	0	0	0	0
-2.23	03	2.3	3332	2	1.	0386	Н	0	0	0	0	0	0	0	0	0	0	0	0

(+ several more hydrogen atoms + bonding information)

Here is the "SDF" file for thymidine monophosphate from PubChem.

Along with lots of stuff not relevant to the EXAFS analysis, we find the Cartesian coordinates of all the atoms in thymidine monophosphate!

SDF = Structure data file

Cartesian coordinates: Feff input file

TITLE Hg decorating thymidine monophosphate

HOLE	4	1.0	* H _i	g L3	edge	(12284	eV),	S0^2
*	mph	ase,m	path,m	feff,	mchi			
CONTROL	1	1	1		1			
PRINT	1	0	0		0			
RMAX	6.0)						

POTENTIALS

ipot element Hg 0 50 1 8 0 2 7 Ν з 6 С 4 15 Ρ 5 1 н

ATOMS

*	x	У	z	ipo
	-3.5515	-1.5175	0.1599	4
	-0.4389	1.3396	1.0202	1
	-0.9101	4.1569	-0.0812	1
	-2.7552	-0.1874	0.6247	1
	3.6173	1.7470	0.3907	1
	3.8378	-2.8022	-0.2452	1
	-2.5475	-2.2163	-0.8977	1
	-4.7267	-0.9241	-0.7790	1
	-4.0197	-2.4002	1.2798	1
	1.6113	0.5684	0.1973	2
	3.7127	-0.5224	0.0726	2
	-1.0101	2.8736	-0.6948	3
	-1.5699	1.8660	0.2995	3
	0.3733	2.3378	-0.9829	3
*	(and so	on)		

- Do some cutting and pasting
- Add some boilerplate for the header
- Make a sensible POTENTIALS list

What about the Hg atom?

What is the likely location of the Hg atom?

- Thymine forms its hydrogen bond with adenisine via the N atom
- The engineered DNA sensor is known to have a T-T mismatch
- Earlier NMR work was interpreted at having the Hg bridging the T-T mismatch.

That said, I don't know much about this chemistry.



Y. Miyake, et al., Mercury^{II}-Mediated Formation of Thymine-Hg^{II}-Thymine Base Pairs in DNA Duplexes. J. Am. Chem. Soc. (2006) v.128, 2172-2173 ***a** DOI: 10.1021/ja056354d



challenging EXAFS analysis problem





Not a great fit, but it tells us that the Hg atom is about 2.05 Å away from it's neighbor. Using the known nucleotide structures, I wrote a small program to solve some trigonometry:

The Hg atom is ...

- … 2.05 Å away from its neighbor
- In the same plane as the neighboring atoms
- In equidistant from the second neighbors (6- and 5-member ring options)
- In collinear with the 1st and 2nd neighbors (monodentate option)

Finally, write out 'feff.inp' files with Hg as the absorber.

5-member ring option: coordinates

TITI	LE Hg	decor	rating	thymi	dine m	onoph	lospha	ate	
HOLE	Ξ	4	1.0	* Hg	L3 ed	ge (1	2284	eV),	S0^2
CONT	FROL	1	1	1	1				
PRIN	T	1	0	0	0				
RMAX	C C	6.0							
POTE	ENTIAL	.S							
*	ipot	Z	eleme	ent					
	0	50	Hg						
	1	8	0						
	2	7	N						
	3	6	С						
	4	15	Р						
ATON	1S								
*	x		У	z	ip	ot			
	0.49	977	0.6	53093	2.	85314	L 0	Hg	0.00000
	-3.71	800	-2.0	00000	-1.	24900) 1	0	6.44507
	-3.91	000	-1.5	59800	0.	29000) 4	Р	5.56632
	-5.35	700	-1.5	58600	0.	60000) 1	0	6.65531
	-3.02	000	-2.4	14600	1.	11000) 1	0	4.98947
	-3.35	200	-0.3	12200	0.	45900) 1	0	4.59727
	-1.95	500	0.0	01100	0.	30500) 3	С	3.59210
	-1.48	700	1.4	10500	0.	63700) 3	С	3.07534
	-0.11	000	1.3	31500	1.	03100) 1	0	2.05000
	-1.52	300	2.3	38600	-0.	51700) 3	С	4.30462
	-1.77	700	3.6	59900	-0.	00600) 1	0	4.77194
	-0.15	400	2.3	19400	-1.	16100) 3	С	4.35705
	0.73	400	1.7	75700	0.	00100) 3	С	3.07532
	1.68	600	0.6	52300	-0.	15600) 2	N	3.23452
	1.54	600	-0.3	35900	-1.	10700) 3	С	4.21393
	0.64	900	-0.3	39000	-1.	93000) 1	0	4.89315
	2.52	400	-1.3	32100	-1.	06300) 2	N	4.82117
	3.58	700	-1.4	10200	-0.	18100) 3	С	4.78223
	4.40	700	-2.3	31400	-0.	22400) 1	0	5.77995
	3.65	500	-0.3	33500	0.	78500) 3	С	3.89431
	4.86	400	-0.3	10800	1.	62800) 3	С	4.59276
	2.71	300	0.8	59300	0.	75000) 3	С	3.05336



5-member ring option: paths

22				Arte	mis (Feff] Atom	s and Pett			
Ren	ame 🧯	Discard	💕	Fef	f in D	emeter	🔒 Feff doc			
\$	🐢 Fel	r 🔯 P	aths	20	Pat	h-like	Console			
R. Save	[,	lot paths		k ×(k)		 R ×(R)	Re(x(R))	l Inf	(17 ×(R)]	Rank
ame o	f this Fel	'f calculat	ion:	off						
* This * The • Clus	paths. central ter siz aths we	dat file atom is e = 5.00 re found	was wi denote A, con within	nitt nitai ni S.	en b y th ning 000	y Demeter is token: 21 atoms A	0.9.20 0			
Forw	and sca	ttering :	utoff	201	~~					
# Forw	and sca	ttering (cutoff	20.	~~					-
For Disc	ring Path	ttering (L Scat	Larie		eh.	Bank	Lin	ari Tyne	
For For	ring Path Deger	ttering (15 1 Reff 2 0396	Scat	terir	ig pa	ch	Rank	Leg	gs Type	ttering
catte	and sca ning Path Deger 1.000 3.000	ttering of 15 15 1 Reff 2.0396 3.0580	Scat	terir () ()	ig pa	ch	Rank 100.00 90.99	Leg 2 2	gs Type single sca single sca	ittering ittering
C000 0001 0002	and sca ning Path Deger 1.000 3.000 1.000	ttering (Scat	terir Q Q		ch	Rank 100.00 90.99 28.47	Leg 2 2 2	gs Type single sca single sca single sca	ittering ittering
COOD 0001 0002 0003	and sca ning Path Deger 1.000 3.000 1.000 4.000	ttering (Reff 2.0396 3.0680 3.2345 3.2670	Scat	terir @ @ C		¢h	Rank 100.00 90.99 28.47 51.62	Leg 2 2 2 3	gs Type single sca single sca single sca obtuse tr	ittering ittering ittering iangle
CO00 CO00 CO01 CO02 CO03 CO04	and sca ning Path Deger 1.000 3.000 1.000 4.000 2.000	ttering (Utoff Scat C C C N C O C O C O C O C O C O C O C O C O C O	terir @ @ C		ch (P	Rank 100.00 90.99 28.47 51.62 16.35	Le; 2 2 3 4	gs Type single sca single sca single sca obtuse tr dog-leg	ittering ittering ittering iangle
Forw 5catte 0000 0001 0002 0003 0004 0005	and sca nno. fu Deger 1.000 3.000 1.000 4.000 2.000 1.000	ttering (Reff 2.0396 3.0680 3.2345 3.2670 3.4587 3.5921	© C	terir @ @ C C		¢h	Rank 100.00 90.99 28.47 51.62 16.35 19.42	Leg 2 2 3 4 2	gs Type single sca single sca obcuse tr dog-leg single sca	ittering ittering ittering iangle ittering
Forw Diate Scatte 0000 0001 0002 0003 0004 0005 0006	and sca nno. fu Deger 1.000 3.000 1.000 4.000 2.000 1.000 2.000	ttering (© C © C © C © C © C © C	20. terir @ @ C C N		ch @	Rank 100.00 90.99 28.47 51.62 16.35 19.42 6.38	Le; 2 2 3 4 2 3	gs Type Single sca single sca obtuse tr dog-leg single sca other dor	ittering ittering ittering iangle ittering ible scat
Fore Dial Scatte 0000 0001 0002 0003 0004 0005 0006 0007	and sca nno. fu Deger 1.000 3.000 1.000 4.000 2.000 4.000 4.000 4.000	ttering (Reff 2.0396 3.0680 3.2345 3.2670 3.4587 3.5921 3.7677 3.8041	© C © C © C © C © C © C © C © C © C © C	terir @ @ C C @ N C		eh @	Rank 100.00 90.99 28.47 51.62 16.35 19.42 6.38 14.82	Leg 2 2 3 4 2 3 3 3 3 3	gs Type Single sca single sca obtuse tr dog-leg single sca other dor other dor	ittering ittering iangle ittering ible scat
Fore codd 0001 0002 0003 0004 0005 0006 0007 0008	and sca ning Path Deger 1.000 3.000 1.000 2.000 1.000 2.000 4.000 1.000	ttering (Reff 2.0306 3.0680 3.2345 3.2670 3.4587 3.5921 3.7677 3.0641 3.8943	© C © C © C © C © C © C © C © C © C © C	terir 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 pa 0 0 0 0 0 0	Ø	Rank 100.00 90.99 28.47 51.62 16.35 19.42 6.38 14.82 15.29	Le; 2 2 3 4 2 3 3 3 3 2	gs Type single sca single sca obtuse tr dog-leg single sca other dor other dor single sca	ittering ittering iangle ittering ible scat ible scat ittering
Forw 5catte 0000 0001 0002 0003 0004 0005 0006 0007 0008 0009	and sca neo_fu ing Path Deger 1.000 3.000 1.000 2.000 1.000 2.000 1.000 2.000 1.000 2.000	ttering (Reff 2.0306 3.0680 3.2345 3.2670 3.4587 3.5921 3.7677 3.8641 3.8943 4.0024	© C © C © C © C © C © C © C © C © C	20. terir @ @ C C @ N C @ N C @ O		¢	Rank 100.60 90.99 28.47 51.62 16.35 19.42 6.38 14.82 15.29 5.98	Les 2 2 3 4 2 3 3 3 3 2 3 3	gs Type single sca single sca obtuse tr dog-leg single sca other dor other dor single sca other dor other dor	ittering ittering iangle ittering ible scat ible scat ittering ible scat
Forv Could Cou	and sca ring Path Deger 1.000 3.000 1.000 4.000 2.000 1.000 2.000 1.000 2.000 1.000 2.000 1.000	ttering 4 1 Reff 2.0396 3.0680 3.2345 3.2670 3.4587 3.5921 3.5921 3.5923 3.5943 4.0024 4.0792	Utoff Scat © C © N © C © O © C © C © C © C © C © C	20. terir @ @ C C C @ N C @ O		¢ Ø	Rank 100.60 90.99 28.47 51.62 16.35 19.42 6.38 14.82 15.29 5.98 6.62	Ley 2 2 3 4 2 3 3 2 3 3 4 3 4	gs Type single sca single sca obcuse tr dog-leg single sca other dor other dor single sca other dor single sca other dor single sca	Ittering Ittering Iangle Ittering Ible scat Ittering Ible scat
Forv Diate Concern Con	and sca ring Path Deger 1.000 3.000 1.000 4.000 2.000 1.000 2.000 1.000 2.000 1.000 2.000 1.000 2.000	ttering (1 Reff 2.03% 3.0680 3.2345 3.2670 3.4587 3.5921 3.7677 3.8641 3.8943 4.0024 4.0792 4.0875	Scat C	20. terir @ @ C C @ O C C @ O C		¢	Rank 100.00 90.99 28.47 51.62 16.35 19.42 6.38 14.82 15.29 5.98 6.62 5.04	Leg 2 2 3 4 2 3 4 2 3 3 2 3 4 3 4 3	gs Type single sca single sca obcuse tr dog-leg single sca other dor other dor single sca other dor rattle other dor	Ittering Ittering Iangle Intering Ible scat Ittering Ible scat
* Forward Control Cont	and sca ring Path Deger 1.000 3.000 1.000 2.000 1.000 2.000 1.000 2.000 1.000 2.000 2.000 2.000 2.000	ttering of Reff 2.0396 3.0680 3.2345 3.2670 3.4587 3.5921 3.7677 3.8641 3.8943 4.0024 4.0074 4.0075 4.1352	Scat C	20, terir @ @ C C @ N C @ O C C C		¢	Rank 100.90 90.99 28.47 51.62 16.35 19.42 6.38 14.82 15.29 5.98 6.62 5.98 6.62 5.04 9.04	Les 2 2 3 4 2 3 3 2 3 3 2 3 4 3 3 4 3 3 3 3	gs Type single sca single sca obtuse tr dog-leg single sca other dou other dou single sca other dou other dou other dou	Ittering Ittering iangle Ittering Ible scat Ible scat Ible scat Ible scat Ible scat
# Forw 5catte 0000 0001 0002 0003 0004 0005 0006 0007 0008 0009 0011 0012 0013 0014 0012 0013 0014	and sca ring Path Deger 1.000 3.000 1.000 2.000 1.000 2.000 1.000 2.000 1.000 2.000 2.000 2.000 2.000 2.000	ttering a Reff 3.0580 3.2345 3.2670 3.4587 3.5921 3.7677 3.8541 3.8943 4.0024 4.0792 4.0875 4.1352 4.2593	Comparison C	20. terir 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		¢	Rank 100.60 90.99 28.47 51.62 19.42 6.38 14.82 19.29 5.98 6.62 5.04 9.04 23.25	Les 2 2 3 4 2 3 3 4 2 3 4 3 2 3 4 3 3 2 2 3 4 2 3 2 3	gs Type single sca single sca single sca obtuse tr dog-leg single sca other dor single sca other dor rattle other dor rattle other dor single sca other dor sol obtuse tr single sca obtuse tr single sca obtu	Ittering Ittering Ittering Iangle Ittering Ible scat Ible scat Ible scat Ible scat Ible scat

Run FEFF, drag-n-drop first 6 paths, transfer them to the plotting list, plot in R:



This looks sort of promising ... or does it?

A challenging EXAFS analysis problem

5-member ring option: VPath

We fit a sum of paths to the data, so let's examine the sum of these paths. In ARTEMIS, this is called a "VPath."



Not so promising, after all.

A challenging EXAFS analysis probler

Monodentate option: coordinates

TITI	LΕ	Hg	decor	rating	; thymi	dine	e mono	ophos	pha	ate		
HOLE	Ξ		4	1.0	* Hg	L3	edge	(122	84	eV),	S0^2	
CONT	rro	L	1	1	1		1					
PRIN	TΙ		1	0	0		0					
RMAX	(6.0									
POTE	ENT	IAI	.s									
*	i	pot	; Z	elem	lent							
		0	50	Hg								
		1	8	0								
		2	7	N								
		3	6	С								
		4	15	Р								
ATON	1S											
*	x			У	2		ipot					
	5	.74	1339	-3.	80032		-0.294	108	0	Hg	0.	00000
	-3	.71	.800	-2.	00000		-1.249	900	1	0	9.	67837
	-3	.91	.000	-1.	59800		0.290	000	4	Ρ	9.	91863
	-5	.35	5700	-1.	58600		0.600	000	1	0	11.	35435
	-3	.02	2000	-2.	44600		1.110	000	1	0	8.	97789
	-3	.35	200	-0.	12200		0.459	900	1	0	9.	83988
	-1	.95	500	0.	01100		0.305	500	3	С	8.	61105
	-1	.48	3700	1.	40500		0.63	700	3	С	8.	95772
	-0	.11	000	1.	31500		1.03	100	1	0	7.	88572
	-1	.52	2300	2.	38600		-0.51	700	3	С	9.	54572
	-1	.77	700	з.	69900		-0.006	500	1	0	10.	62446
	-0	.15	6400	2.	19400		-1.16	100	3	С	8.	45356
	0	.73	3400	1.	75700		0.00	100	3	С	7.	48765
	1	. 68	3600	0.	62300		-0.156	500	2	N	6.	00394
	1	.54	600	-0.	35900		-1.10	700	3	С	5.	48832
	0	.64	900	-0.	39000		-1.930	000	1	0	6.	34502
	2	. 52	2400	-1.	32100		-1.063	300	2	N	4.	13555
	3	. 58	3700	-1.	40200		-0.18	100	3	С	з.	22719
	4	.40	700	-2.	31400		-0.224	100	1	0	2.	05000
	3	.65	500	-0.	33500		0.78	500	3	С	4.	18739
	4	.86	6400	-0.	10800		1.628	300	3	С	4.	25452
	2	.71	300	0.	59300		0.750	000	3	С	5.	43826



Monodentate option: VPath

Same exercise – run feff, drag-n-drop the first few paths, make a VPath, plot with the data.



Better than the 5-member ring option, but still not so great.

A challenging EXAFS analysis problem

6-member ring option: coordinates

TITLE Hg decorating thymidine monophosphate													
HOLE	Ξ		4	1.0	* H	Ig	L3	edge	(12	2284	eV),	S0^2	
CONT	[RO	L	1	1	1			1					
PRINT 1			1	0	()		0					
RMAX	(6.0										
POTE	ENT	IAI	.s										
*	i	pot	; Z	elen	lent								
		0	50	Hg									
		1	8	0									
		2	7	N									
		3	6	С									
		4	15	Р									
ATON	1S												
*	x			У		z		ipot					
	2	.40	463	-2.	80748	3		-2.458	560	0	Hg	0	.00000
	-3	.71	.800	-2.	00000)		-1.249	900	1	0	6	.29242
	-3	.91	.000	-1.	59800)		0.290	000	4	Ρ	6	.99112
	-5	. 35	5700	-1.	58600)		0.600	000	1	0	8	.43040
	-3	. 02	2000	-2.	44600)		1.110	000	1	0	6	.50160
	-3	. 35	200	-0.	12200)		0.459	900	1	0	6	.98896
	-1	. 95	500	0.	01100)		0.305	500	3	С	5	.87972
	-1	.48	3700	1.	40500)		0.63	700	3	С	6	.51567
	-0	. 11	000	1.	31500)		1.03	100	1	0	5	.95606
	-1	. 52	2300	2.	38600)		-0.51	700	3	С	6	.79387
	-1	.77	700	з.	69900)		-0.006	600	1	0	8	.11301
	-0	. 15	6400	2.	19400)		-1.16	100	3	С	5	.76519
	0	.73	3400	1.	75700)		0.00	100	3	С	5	.44613
	1	. 68	3600	0.	62300)		-0.156	600	2	N	4	. 19199
	1	. 54	600	-0.	35900)		-1.10	700	3	С	2	.92421
	0	. 64	900	-0.	39000)		-1.930	000	1	0	3	.03360
	2	. 52	2400	-1.	32100)		-1.063	300	2	N	2	.05000
	3	. 58	3700	-1.	40200)		-0.18	100	3	С	2	.92356
	4	.40	700	-2.	31400)		-0.224	400	1	0	3	.03859
	3	. 65	500	-0.	33500)		0.78	500	3	С	4	.26357
	4	. 86	6400	-0.	10800)		1.628	B00	3	С	5	.47827
	2	.71	300	0.	59300)		0.750	000	3	С	4	.68340



6-member ring option: VPath

Again – run FEFF, drag-n-drop the first few paths, make a VPath, plot with the data.



I actually like this one quite a bit! The amplitude is off by about a factor of 2, but the phase is quite close.

Metal sensors Experiment DNA Model building The fit Post mortem

Number of independent pointsk-range: $2 Å^{-1}$ to $8.8 Å^{-1}$ R-range: 1 Å to 2.8 Å $N_{idp} = 2\Delta k \Delta R / \pi \approx 7.8$

0	E_0 and amp are variables $\ldots \ldots \ldots$	(1,2)
2	Hg-N distance and σ^2 are variables \ldots	(3,4)
3	Hg-O distance and σ^2 are variables \ldots	(5,6)
4	Assume that the ring is completely rigid , this allows us to approximate contributions of various single and multiple scattering paths without	the

introducing any more variables.

Metal sensors Experiment DNA Model building The fit Post mortem Trigonometry



Here's a formula for a triangle in a plane:

$$D(Hg - C) = \frac{a - b}{\cos(\theta)} \cos(\varphi/2)$$
$$\tan(\theta) = \frac{a + b}{a - b} \tan(\varphi/2)$$

$$arphi$$
 =116.25°
b =1.378 Å

a and $\sigma^2_{Hq\cdot N}$ are variables of the fit.

Assuming the ring is **rigid**, then we approximate σ^2_{Hg} . (and others) by scaling geometrically from σ^2_{Hg} .N

				The fit	
		Paths	5		
Path 1	(SS)	Path 2 (SS)		Path 3 (SS)	
$0 \stackrel{C}{=} \stackrel{N}{\underset{H_{1}}{\sim}} N$	\times ×1 $\sim^{c} \approx_{0}$ are variables	$0 = \sum_{ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\$	×2 o	$\Delta R_3 \text{ and } \sigma_3^2 \text{ are variables} \times 2$	
Path 4	(MS)	Path 5 (MS)		Path 6 (MS)	
	×4 < ^c ≈₀	$\begin{array}{c} C & N \\ & \\ 0 \\ \\ 0 \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\$	×2 °	° N ×4 ° N C N C N C N C N C N C N C N C N C N	
ΔR_4 compute	d from paths 1 and 2	ΔR_5 computed from pat	h 1	ΔR_6 computed from paths 1 and	d 3
$\sigma_4^2 := \sigma_2^2$		$\sigma_5^2 := \sigma_2^2$		$\sigma_6^2 := \sigma_1^2 + \sigma_3^2$	

Metal sensors	Experiment	DNA	Model building	The fit	Post mortem
		Fittin	g result		
			amp	1.86 ± 0.44	
4	3mM 3mM		Eo	1.41 ± 1.91	

Λ

Λ



	1 06 1 0 44
amp	1.80 ± 0.44
E_0	1.41 ± 1.91
$\Delta R(N)$	0.006 ± 0.028
$\Delta R(O)$	-0.058 ± 0.063
$\sigma^2(N)$	0.0046 ± 0.0045
$\sigma^2(O)$	0.0096 ± 0.0081

Why is amp near 2?

The Hg atom bridges 2 thymines. Our FEFF model had Hg bound to 1 thymine. So S_0^2 is really 0.93(44)!

Wieldt Sellsuis		woder buttuting	The m	
	llncor	taintiac		
	Uncer	lalilles		

- The data are short i.e. little information content and noisy
- The uncertainties are all quite large, although the best fit values all make sense
- S_0^2 came out right, although with large uncertainty
- The σ^2 approximations are sensible, but certainly not correct
- The assumption that the ring is rigid is sensible, but certainly not correct
- The assumption that the Hg atom sits in the plane of the ring is sensible, but certainly not correct

Our data are $\ensuremath{\textit{consistent}}$ with the Hg atom bound to the N atom in the 6-member nitrogenous base

- The As in the cacodylic acid hurt. Use a different buffer.
- The sample geometry hurt. Use better packaging or a focusing mirror.

Those two things could have increased efficiency by about an order of magnitude. Another couple inverse Ångstroms would have made a huge difference!

What have we learned?

- The science question required interpretation of both XANES and EXAFS
- Quick first shell fit to approximate the first shell distance
- Made input for FEFF from published structural data and a sensible guess for the location of the Hg atom
- Tried several possible coordination geometries, but only pursued the one that looked promising
- Dealt with limited information by applying interesting constraints
- We didn't exactly *solve* the structure, but we demonstrated that the EXAFS data are consistent with the assumption from NMR