SAMPLE POTENCY ESTIMATION REVISITED: REASONS AND OPTIONS FOR EMPLOYING A LINEARIZED POTENCY ESTIMATION MODEL

Baston, D.S.[1], Elskens, M.[1], Denison, M.S.[2], Goeyens, L., and Baeyens, W.[1]

1. Department of Analytical and Environmental Chemistry, Vrije Universiteit Brussels, Brussels Belgium.
2. Department of Environmental Toxicology, Meyer Hall, University of California, Davis, CA 95616.

Abstract
Exposure of humans and wildlife to HAH compounds through food and feed is a concern in recent years with the European Union and the USA implementing legislation to monitor exposure. These widespread persistent environmental contaminants have varied species and tissue specific biological effects including but not limited to: birth defects, tumor promotion, lethality, endocrine disruption, and other effects. Biological samples, in recent years, are exhibiting low level concentrations of HAH compounds pushing the lower limits of detection and quantification. One of the tools used for screening of biological samples is the CALUX bioassay, we have generated Mid and Low-Range linearized models from sigmoid dose curves providing an increased lower range for detection of HAHs. The EC_{50} values of these models correlate well with the historically used sigmoidal models with well structured dose response curves.

Introduction
Halogenated aromatic hydrocarbons (HAHs) are, at some level, a part of every day life, they are ubiquitous in the environment[1]. It is important to understand the level of exposure and consequently the risk experienced by the presence of these compounds due to their potential for significant biological effects: eg. endocrine disruption, immunotoxicity, induction of numerous enzymes, birth defects, as well and others[2]. Identification and quantitation of halogenated aromatic hydrocarbons (HAHs), such as polychlorinated dioxins (PCDDs), polychlorinated furans (PCDFs), and biphenyls has typically been performed by high resolution instrumental methods (i.e. GC/HRMS), the long term gold standard, however biological relevance is difficult to establish. Multiple inexpensive and rapid bioanalytical methods have been developed in the last decade[1], these methods allow interpretation of biological response, expressed as Bioanalytical Equivalents (BEQs), to quantities of HAHs and other HAH-like compounds extracted from various matrices, including but not limited to, soil, sediment, fly ash, food and feed, and biological samples including blood, milk, and fat. In contrast to environmental samples, biological samples appear to be decreasing in concentration as a result of legislation and regulatory actions controlling major HAH sources[3].

CALUX, a well defined detection system used for screening purposes, is an AhR-based cell bioassay system capable of responding to individual compounds, such as TCDD, PCB 77, other HAHs, and, beta-Naphthaflavone as well as complex mixtures extracted from various environmental and biological matrices. The potency of individual compounds as well as complex mixtures is based upon induction of luciferase reporter gene expression compared to expression of luciferase from TCDD standards, the most potent AhR agonist. Typical environmental samples yield complete or full receptor-dependent sigmoidal dose curves with upper and lower plateaus and a linear mid-section, however biological samples typically have low concentrations, not conducive to a complete dose curve and more often relegated to analysis of a single point making the potency difficult to estimate with confidence at the lower end of the sigmoidal standard curve. Resultant Bioanalytical Equivalents (BEQs), estimated from the TCDD EC_{50} and a sample potency from the CALUX bioassay are compared with analytical Toxicity Equivalents (TEQs) derived from Toxic Equivalency Factors (TEFs) established by the WHO in 2006[4].

Implementing the use of screening tools for HAHs in biological matrices at low levels is significantly important given the various levels of environmental contamination potentially leading to low level contamination of food, the source for 90% of the general population intake[5], ultimately leads to exposure of humans and wildlife. Screening of biological samples is usually performed with a single point as opposed to the usual multipoint environmental sample extract dose curve. Above and beyond the biological difficulties/implications with estimating potencies at low induction levels, sometimes there may not be any alternatives to determining the

potency of a sample from a single point induction can be difficult at low induction levels. We present here two different linearized models, a low-range linear model and a mid-range linear model, both based upon the sigmoidal dose response curve designed with the purpose of estimating potencies for low HAH concentrations in biological samples, in contrast to estimation models for environmental samples[6]. Both curves are capable of estimating and EC50, the low-range model is capable of estimating in the range of 6% - 60% where the mid-range model utilizes a range of 25% - 90%. Both models are derived from one sigmiodal dose response curve, current results from tests with standard curves and individual EC_{50} data indicates an excellent correlation between EC_{50} potency estimations from the two linear models as well as correlating well with EC_{50} potency estimations from nonlinear regression models and the sigmoidal model and from Sigma Plot. The ease of use and presentation of multiple correlative potencies is an important step towards providing more accurate data when screening biological samples for compliance with current EU regulations[6].

Methods

Cell Culture, Chemical Treatment and CALUX Analysis. Standard curves were generated using two different stably transfected CALUX cell lines, H1L6.1c3 and the recently developed G3 cell line were grown as previously described[9]. The inducing potency of each sample was evaluated in a range finding study, followed by a ten-point dose curve.[8-10] Cells in a 96-well microplate format were incubated with various standard solution concentrations for 24 hours, treatment solution was discarded appropriately, the cells were rinsed, lysed and luciferase activity determined using a Promega Glomax luminometer using 50□l of lysis buffer, 50 □l of luciferin with a 5.6 second lag time after luciferin addition and prior to integration. Integration time for the H1L6.1c3 cell line was set at 5.0 sec and integration time for the H1L7.5c1 cell line was 3.0 sec.

Statistical Evaluation of Standard Curves

Triplicate concentrations were used for each of a ten point standard curve. Multiple standard curves were diluted in DMSO from one series of 26 standard dilutions. The solutions were applied to multiple plates containing both the H1L6.1C3 cell line and the new G3 Cell line (H1L7.5). In all 32 standard curves were generated, the linearized model and the Sigma Plot version of the 4 Parameter Hill plot were used to estimate EC50 values, 4 in all for each curve.

Estimation of EC50 values for TCDD standard curves and sample extract dose curves was achieved using two different mathematical models: 1) the model 4-parameter Hill plot from the Sigma plot program offered by Systat; 2) a linear transformation function established optimizing the R^2 value for two different functions: a mid-range (25% to 90%) and low-range function (6% to 60%). These functions were developed from established statistical protocols.

Results and Discussion

Potency estimation of low concentration samples is confounded by two different, independent mathematical issues when using nonlinear regression models: 1) statistical error at low concentrations prevents precise estimation of potency based upon response from the data and 2) the uncertainty changes from an additive error to a multiplicative error when back calculating from potency to response, this rearrangement can be observed when transforming Equation A to Equation B. The Mid-range and Low-range linear models, based upon Equation C, a transformation of the error observed with each data point in the sigmoid plot, increase the precision, by decreasing data point error, of the potency estimation at low concentrations by minimizing the error associated with the data points.

To provide potency values from single point analysis of low concentration samples two linear models were created, a Mid-Range model capable of estimating potencies between 25% and 90% and a Low-Range model capable of estimating potencies for single point samples between 6% and 60%. Using a ten point TCDD standard curve we established the criteria for these models where the confirmatory data is the generation of correlative EC_{50} values from non-linear regression models and the two new linear regression models.

The process of providing accurate potency estimations at low concentrations and hence low induction using the Mid-range and Low-range is dependent upon the correlation of each of the models EC_{50} values. Review of data

used to properly estimate EC50 values with reasonable agreement there is needed a minimum of ten points to create a standard curve which satisfies enough criteria in all of the models to establish agreement between EC50 values. Specifically, the concentrations of TCDD standards are best distributed in a 2:1:4:1:2 manner where there are two concentrations which define both the upper and lower plateaus, single points falling in the transition region and four points within the middle section of the curve. In relation to the range of values for each linear model the four points in the middle of the sigmoid curve are four points which are common to both the Mid-range and Low-range linear models, the range of percent induction can be defined by the upper boundary of the Low-range linear model and lower boundary of the Mid-Range linear model, 25% to 60% of maximal induction. Furthermore, each of the individual points fall within the transition portion between the mid section and each of the upper and lower plateaus; between 6% and 25% for the lower transitional point and between 60% and 90% for the upper transitional point. The two sets of points defining the upper and lower plateaus are below 6% and above 90%. Data from two of the curves is presented in Table 1 with the resultant EC_{50} data provided in Table 2.

The mathematical process in fitting the Mid- and Low-range linear curves can be described as seen in the set of figures A1 – A4 and B1 – B4 where Figures A1 – A4 are for data that work with the two models and figures B1 – B4 are for data that does not work with the two models. Figures A1 and B1 the overall uncertainty in the experimental RLU values has a positive slope or as the experimental RLU values increase so does the percentage induction. Figures A2 and B2 illustrate the nonlinear regression fit using the 4 parameter Hill function as described by equation A above. Figures A3 and B3 are the linear transformation of the nonlinear data from Figures A2 and B2 (the equation in fact encompasses both the Mid- and Low-range models). Figures A4 and B4 show the residual plot expressed as a z-score (standard residuals). If the model correctly evaluates the data it is expected that $|z| > 2$ in less than 5% of the cases and $|z| > 3$ in less than 0.3% of the cases. The latter indicates unacceptably poor performance in terms or accuracy while for a satisfactory model performance $|z| \leq 2$ is required.

It was clear that the increased response of the H1L7.5c1 cell line rendered insufficient data points at low concentrations of the TCDD standards used for the H1L6.1c3 cell line. Studies are currently underway to evaluate the linear models using a greater number of lower concentration TCDD standards.

References:

1. Simonich SL, Hites RA. Science 1995;269:1851.
2. Denison MS, Zhao B, Baston DS, Clark GC, Murata H, Han, D-H, (2004) Talanta 63:1123.
3. Bayens W, Verstraete F, Goyens L, (2004) Talanta 63:1095
4. Van den Berg M, Birnbaum LS, Denison M, De Vito M, Farland W, Feeley M, Fiedler H, Hakansson H, Hanberg A, Haws L, Rose M, Safe S, Schrenk D, Tohyama C, Tritscher A, Tuomisto J, Tysklind M, Walker N, Peterson RE, Toxicol Sci 2006;93:223
5. Baeyens W, Leermakers M, Elskens M, Van Larebeke N, De Bont R, Vanderperren H, Fontaine A, Degroodt J.-M, Goeyens L, Windal I, (2007) Arch. Environ. Contam. Toxicol. 52:563
6. Windal I, Van Wouwe N, Eppe G, Xhrouet C, Debacher V, Bayens W, De Paul E, and Goeyens L, (2005) Environ Sci Technol 39:1741
7. Brown DJ, Orelien J., Gordon JD, Chu AC, Chu MD, Murata HJ, Kayama F, Denison MS, Clark GC (2007) Environ Toxicol Chem 41:4354.
8. Rogers JM, Denison MS (2000) In Vitro Molec Toxicol 13:67.
9. Ziccardi MH, Gardner IA, Denison MS (2000) Toxicol Sci 54:183.
10. Chen Y-H, Tukey RH (1996) J Biol Chem 271:26261.

Equation A.
$$y = y_0 + \frac{m \cdot (\ln x)^n}{k^n + (\ln x)^n} + \varepsilon_y$$

Basic Sigmoid model employed to estimate potency values (ε_y is the residual (error) term).

Equation B. $$x = \exp\left(\left(\frac{(y - y_0)\cdot k^n}{m - y + y_0}\right)^{1/n}\right)\cdot \varepsilon_x^{'}$$

Restructured Sigmoid equation for potency estimation of a sample as a function of response either RLU values or percent maximal induction.

Equation C. $$y^* = y^p = s\cdot x + i + \varepsilon_{y^*}$$

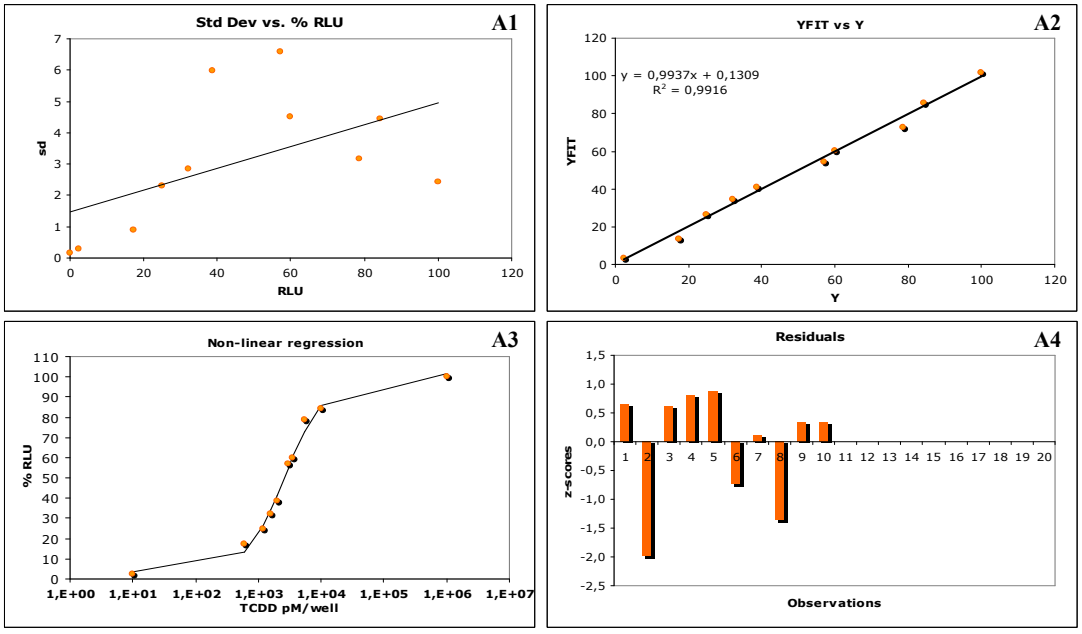
Linearizing equation for the Mid-range and Low-Range sections of the complete sigmoidal TCDD Dose curve, the linearized equations to be used for estimating potency of low concentration samples or single point analysis when sample concentrations are thought to be low.

A1-A4			B1-B4		
pM	Avg % RLU	Std Dev	pM	Avg % RLU	Std Dev
1000000	100	2,4	37117,3	100	3,7
9928, 9	84,2	4,4	7843,8	85,4	5,6
5436,7	78,5	3,2	4838,6	70,1	2,6
3535,2	59,9	4,5	3535,2	56,5	0,2
2951,9	57,0	6,6	2951,9	53,4	1,5
1966,0	38,7	5,9	2259,8	40,9	1,9
1572,8	32,2	2,8	1179,6	26,6	0,9
1179,6	24,9	2,3	783,3	20,6	0,8
587,4	17,3	0,9	97,7	3,9	0,2
9,7	2,4	0,3	9,7	2,0	0,1

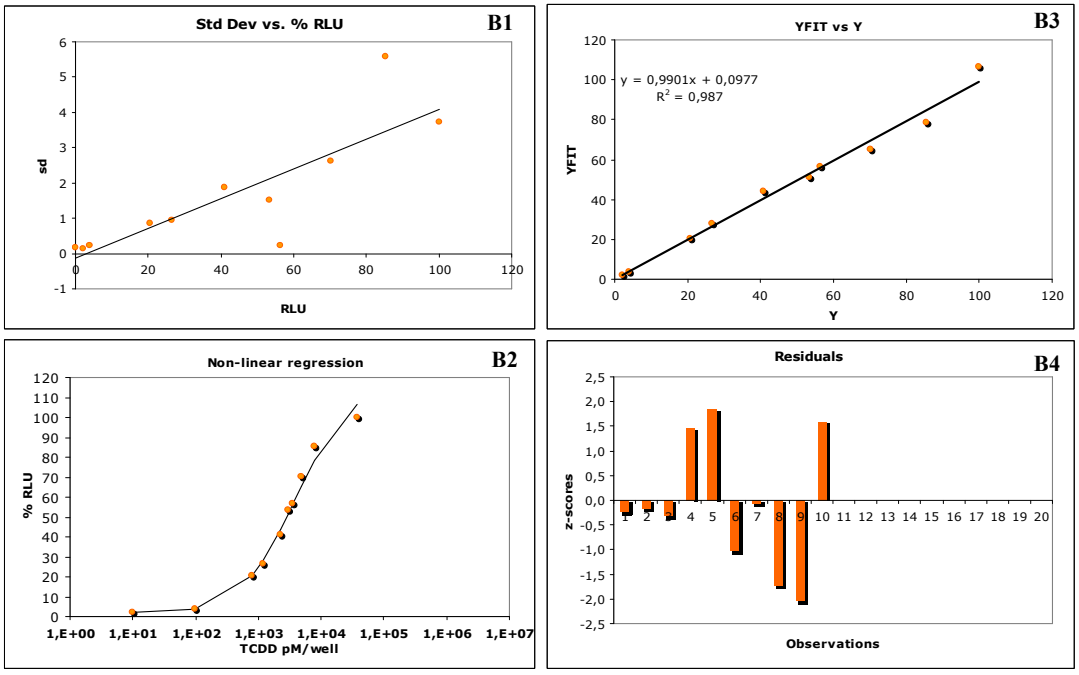
Table 1. Subset of 32 different TCDD Standard Curve data. Provided for illustration of standard concentrations used to establish response within the boundaries as described in the text.

	NLR		Mid-Range		Low-Range		Sigma Plot	
	EC50	**Std Err**	**EC50**	**Std Err**	**EC50**	**Std Err**	**EC50**	**Std Err**
A1-A4	2798,2	266.9	2749,2	162.2	2895,6	185.3	2700,4	213,0
B1-B4	4299,2	653.4	3850,8	92.4	3671,7	51.4	3054,1	218,0

Table 2. Estimated values potency values generated from the two different nonlinear models and the Mid- and Low-Range linearized models exhibiting the potency data.



Figures A1-A4: Statistical representation of bioassay data used to generate the Mid- and Low-Range potency estimation models with statistically correlative EC_{50} values.



Figures B1-B4: Statistical representation of bioassay data used to generate the Mid- and Low-Range potency estimation models generating non-correlative EC_{50} values.