Amino Acid Sequence of the Ligand Binding Domain of the Aryl Hydrocarbon Receptor 1 (AHR1) Predicts Sensitivity of Wild Birds to Effects of Dioxin-like Compounds

Reza Farmahin[*, †] (mfarm070@uottawa.ca),
Gillian E. Manning[*, †] (gmann017@uottawa.ca),
Doug Crump[†] (Doug.Crump@ec.gc.ca),
Dongmei Wu[†] (dongmei_wu@hotmail.com),
Lukas J. Mundy[*,†] (lukasmundy@gmail.com),
Stephanie P. Jones[†] (Stephanie.Jones@ec.gc.ca),
Mark E. Hahn[‡] (mhahn@whoi.edu),
Sibel I. Karchner[‡] (skarchner@whoi.edu),
John P. Giesy[§, ¶, ||] (john.giesy@usask.ca),
Steven J. Bursian[|||] (bursian@msu.edu),
Matthew J. Zwiernik[|||] (zwiernik@msu.edu),
Timothy B. Fredricks[¶] (timothy.b.fredricks@monsanto.com) and
Sean W. Kennedy[*, †, 1] (sean.kennedy@ec.gc.ca)

[*] Centre for Advanced Research in Environmental Genomics, Department of Biology, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5
[†] Environment Canada, National Wildlife Research Centre, Ottawa, Ontario, Canada, K1A 0H3
[‡] Department of Biology, Woods Hole Oceanographic Institution, Woods Hole, MA, USA, 02543
[§] Department of Veterinary Biomedical Sciences and Toxicology Centre, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, S7N 5B3
[¶] Department of Zoology and Center for Integrative Toxicology, Michigan State University, East Lansing, Michigan, USA, 48824
[||] Department of Biology & Chemistry, City University of Hong Kong, Kowloon, Hong Kong, SAR, China
[|||] Department of Animal Science, Michigan State University, East Lansing, Michigan, USA, 48824

[1]To whom correspondence should be addressed:

Sean W. Kennedy, Environment Canada, National Wildlife Research Centre, 1125 Colonel By Drive, Raven Road, Ottawa, Ontario, K1A 0H3

sean.kennedy@ec.gc.ca Tel: 613-998-7384; Fax: 613-998-0458

Abstract. The sensitivity of avian species to the toxic effects of dioxin-like compounds (DLCs) varies up to 1000-fold among species and this variability has been associated with inter-species differences in aryl hydrocarbon receptor 1 ligand binding domain (AHR1 LBD) sequence. We previously showed that LD_{50} values, based on *in ovo* exposures to DLCs, were significantly correlated with *in vitro* EC_{50} values obtained with a luciferase reporter gene (LRG) assay that measures AHR1-mediated induction of cytochrome P4501A in COS-7 cells transfected with avian AHR1 constructs. Those findings suggest that the AHR1 LBD sequence and the LRG assay can be used to predict avian species sensitivity to DLCs. In the present study, the AHR1 LBD sequences of 86 avian species were studied and differences at amino acid sites 256, 257, 297, 324, 337 and 380 were identified. Site-directed mutagenesis, the LRG assay and homology modeling highlighted the importance of each amino acid site in AHR1 sensitivity to 2,3,8,8-tetrachlorodibenzo-*p*-dioxin and other DLCs. The results of the study revealed that: (1) only amino acids at sites 324 and 380 affect the sensitivity of AHR1 expression constructs of 86 avian species to DLCs and (2) *in vitro* luciferase activity in AHR1 constructs containing only the LBD of the species of interest is significantly correlated (r^2 = 0.93, p<0.0001) with *in ovo* toxicity data for those species. These results indicate promise for the use of AHR1 LBD amino acid sequences independently, or combined with the LRG assay, to predict avian species sensitivity to DLCs.

Keywords: dioxin, risk assessment, bird, Ah receptor, molecular toxicology

Introduction

Dioxins and dioxin-like compounds (DLCs) share structural similarities and common mechanisms of biochemical and toxic action, but differences in sensitivity to DLCs exist among species and strains of animals. For example, sensitivity to the biochemical and toxic effects of some DLCs can vary up to 1000-fold among avian species (Brunstrom, 1988; Head *et al.*, 2008; Head and Kennedy, 2010; Hoffman *et al.*, 1998). Based on the available evidence, the domestic chicken (*Gallus gallus domesticus*) is the most sensitive species to the toxic effects of DLCs. For risk assessment purposes, the toxic potencies of individual DLCs are compared to the potency of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), which is generally considered to be the most potent compound within this class of chemicals (Van den Berg *et al.*, 1998; Van den Berg *et al.*, 2006). However, recent studies have shown that, in some species, 2,3,4,7,8-pentachlorodibenzofuran (PeCDF) is more potent than TCDD (Cohen-Barnhouse *et al.*, 2011; Farmahin *et al.*, 2012; Herve *et al.*, 2010a; Herve *et al.*, 2010b).

Most, if not all, biochemical and toxic effects of DLCs are thought to be mediated by the aryl hydrocarbon receptor (AHR) (Denison *et al.*, 2011; Okey, 2007). While it has long been recognized that structural characteristics of the AHR contribute to differential sensitivity among mammalian species (Okey, 2007; Poland *et al.*, 1976), key characteristics of the AHR responsible for differential sensitivity among birds were not known until recently (Farmahin *et al.* 2012; Head *et al.* 2008; Karchner *et al.* 2006). Birds express at least two AHR paralogs

(AHR1 and AHR2) (Yasui *et al.*, 2004) and AHR1 is reported to be more transcriptionally active than AHR2 (Yasui *et al.* 2007). In addition, the basal level of expression of AHR1 is several-fold greater than that of AHR2 in liver of birds (Mol *et al.*, 2012; Yasui *et al.*, 2007).

The avian AHR1 contains three major domains - the DNA binding domain (DBD), the ligand binding domain (LBD) and the transactivation domain (TAD). By swapping the three domains of chicken and common tern (*Sterna hirundo*) AHR1 and making six chimeric constructs, Karchner *et al.* (1996) demonstrated that the LBD, and not the DBD or TAD, was responsible for the distinct functional properties of chicken and tern AHR1. This led to the hypothesis that the identity of amino acids within the LBD could determine the sensitivity of avian species to DLCs. The identities of amino acids at sites 324 and 380 within the LBD contributed to the differential sensitivity of chicken and tern AHR1 exposed to TCDD (Karchner *et al.*, 2006) and chicken, ring-necked pheasant (*Phasianus colchicus*) and Japanese quail (*Coturnix japonica*) exposed to TCDD, 2,3,7,8-tetrachlorodibenzofuran (TCDF) and PeCDF (Farmahin *et al.* 2012). It has been proposed that birds can be classified into three main groups, for risk assessment purposes, based on the identities of these two amino acids: high sensitivity (type 1; Ile324_Ser380), moderate sensitivity (type 2; Ile324_Val380) and low sensitivity (type 3; Val324_Ala380) (Farmahin *et al.,* 2012; Head *et al.,* 2008; Karchner *et al.* 2006; Manning *et al.* 2012).

The major goals of the present study were to: (1) sequence the AHR1 LBD of a relatively large number of birds to identify amino acid residues that differ

among species, (2) determine the influence of the identified amino acids, individually or in combination, on AHR1 activation by DLCs using a luciferase reporter gene assay (LRG) that measures transactivation of an aryl hydrocarbon response element (AHRE)-driven luciferase reporter gene, (3) determine if avian sensitivity to selected DLCs [TCDD, PeCDF and polychlorinated biphenyl 126 (PCB 126)] can be predicted from results of the LRG assay and/or knowledge of amino acids at sites 324 and 380 within the AHR1 and (4) attempt to understand why sites 324 and 380 play key roles in AHR1 activation by use of homology modelling studies.

Experimental

Sources of Avian Samples Used for AHR1 LBD Sequencing

Liver and blood were obtained from the Specimen Bank at the National Wildlife Research Centre (NWRC; Ottawa, ON), commercial suppliers in the Ottawa region or field-collected by our group in Michigan, USA (Table 1). All samples were collected, salvaged or donated in accordance with permit requirements. The samples obtained from Michigan were collected with either solvent-rinsed utensils (liver) or sterile syringes (blood) (Head *et al.*, 2010), preserved in RNAlater™ (Ambion, TX, USA) at -20ºC at Michigan State University (East Lansing, MI, USA) and shipped on wet or dry ice to the NWRC for RNA isolation and AHR1 LBD sequencing. Albatross, great cormorant and red jungle fowl AHR1 LBD sequences were obtained from GenBank.

AHR1 LBD Sequencing

RNA was isolated from liver (~ 3 mg) or blood (~ 300 µl) with TRIzol™ reagent (Invitrogen, Burlington, ON, Canada). Isolation of RNA, reverse transcription, polymerase chain reaction (PCR) and molecular cloning methods are described in detail elsewhere (Head *et al.*, 2008; Head *et al.*, 2010). Briefly, total RNA was reverse transcribed to cDNA and the AHR1 LBD was amplified by PCR and either (a) ligated into a vector, transformed into chemically-competent cells and purified (plasmid DNA) or, (b) separated on E-Gel 0.8% SYBR Safe™ pre-cast agarose gels using an E-Gel iBase system (Invitrogen). The latter procedure was used to reduce time and cost per sample (Gibson *et al.*, 2010). Purified plasmid DNA and E-Gel products were sequenced by use of an Applied Biosystems 3730 DNA Analyzer at the Ottawa Hospital Research Institute (OHRI, Ottawa, ON, Canada).

To ensure accuracy, most sequences of the LBD of AHR1 were obtained from two or more individuals per species and sequencing was conducted several times for each individual and on products from independent PCR reactions. Sequences of nucleotides were analyzed by use of Sequencher version 4.9 software (Gene Codes Corporation, Ann Arbor, MI, USA). Detailed analysis of each chromatogram ensured quality and accuracy of the sequence data. Amino acid sequences were translated from the consensus nucleotide alignments using Sequencher. The AHR1 LBD sequences corresponded to amino acid residues 235-402 in chicken.

159 Expression constructs and site-directed mutagenesis of chicken AHR1

160 Full-length chicken, ring-necked pheasant and Japanese quail AHR1
161 constructs were prepared (Farmahin *et al.*, 2012) and the chicken AHR1
162 construct was used for the preparation of twelve constructs that were mutated at
163 sites 256, 257, 297, 324, 337 or 380 by site-directed mutagenesis. The primers
164 and templates are provided in Table S1. All of the mutations were constructed
165 using PfuUltra High-Fidelity DNA polymerase (Stratagene, La Jolla, CA, USA).
166 PCR conditions were as follows: 95 ˚C 1min followed by 21 cycles of 95 ˚C
167 30sec, 55 ˚C 1min, and 68˚C 18min. PCR products were treated with *DpnI*
168 endonuclease (New England Biolabs, Beverly, MA) and transformed into MAX
169 Efficiency DH5α Competent Cells (Invitrogen). Full-length sequencing of each of
170 the 12 mutant constructs was conducted to ensure correct sequence identity.
171 Other constructs were cormorant ARNT1 (pcDNA-ccArnt1), reporter vector GL4-
172 ccCYP1A5 (both were kindly provided by Dr. Hisato Iwata, Ehime University,
173 Japan) (Lee *et al.*, 2009; Yasui *et al.*, 2007) and *Renilla* luciferase vector (phRL-
174 CMV, Promega).
175
176 *In vitro* transcription and translation (IVTT) of AHR1 mutants

177 TnT Quick-Coupled Reticulocyte Lysate Systems kits (Promega, Madison,
178 WI, USA) were used to synthesize AHR1 proteins for all mutants; proteins were
179 labelled with FluroTect Green$_{lys}$ tRNA. Fluorescently labelled products were
180 separated on 10% SDS polyacrylamide gels and scanned with a Typhoon 9210

Imager (Molecular Dynamics Inc., Sunnyvale, CA, USA; excitation 532 nm, emission 580 nm).

AHR1 protein expression in COS-7 cells

Western blot analysis was performed as described elsewhere (Farmahin *et al.*, 2012) with minor modifications. In brief, COS-7 cells transfected with AHR1 constructs were lysed and cell lysates were resolved by SDS-PAGE. Proteins were transferred to PVDF membranes (GE Healthcare Bio-Sciences, Baie d'Urfé, QC, Canada) and anti-V5-HRP antibody (Invitrogen) was used at 1:2000 dilution for detecting V5-AHR1. Anti-β-actin-peroxidase (Sigma-Aldrich, Oakville, ON, Canada) was used at 1:10000 as a loading control. The blots were then visualized by enhanced chemiluminescence using a Typhoon 9210 Imager (Molecular Dynamics Inc).

LRG assays

i. Preparation of TCDD, PeCDF and PCB 126 solutions

Serial dilutions of TCDD and PeCDF were prepared from dimethyl sulfoxide (DMSO) stock solutions and concentrations of the stock solutions were determined by isotope dilution following EPA method 1613 (U.S.EPA, 1994) by high-resolution gas chromatography high-resolution mass spectrometry as described in detail elsewhere (Herve *et al.*, 2010a). PCB 126 (AccuStandard, New Haven, CT, USA; lot # 061204MS-AC; 99.7% purity) was weighed on an analytical balance and dissolved in 4 ml of DMSO to obtain a stock solution with

a nominal concentration of 153 µg/mL. Serial dilutions of PCB 126 were subsequently prepared from this solution.

ii. Cell culture, transfection and luciferase assay

A detailed description of the LRG assay is provided elsewhere (Farmahin *et al.*, 2012). Briefly, COS-7 cells (provided by Dr. R. Haché, University of Ottawa, Ottawa, ON, Canada) were plated at a concentration of 10,000 cells/well in 96-well plates and transfected after 18 hours. The amounts of transfected expression vectors were 8 ng of mutated avian AHR1, 1.55 ng of pcDNA-ccArnt1, 5 ng of pGL4-ccCYP1A5 and 0.75 ng of phRL-CMV. The total amount of transfected DNA was kept constant at 50 ng by the addition of salmon sperm DNA (Invitrogen). Cells were dosed 5 hours after transfection with DMSO or DMSO solutions of TCDD, PeCDF or PCB 126 at 0.05% final DMSO concentration. Cells were removed from the incubator 18–20 hours after dosing to measure luciferase activity. Luminescence values are expressed as a ratio of firefly luciferase units to *Renilla* luciferase units.

iii. Concentration-response curves

Two independent studies (referred to as Studies 1 and 2 below) were conducted and for each of the studies, four concentration-response curves were obtained for each DLC and each AHR1 construct. The four concentration-response curves were derived from four replicate wells/plate for each concentration. GraphPad (GraphPad Prism 5.0 software, San Diego, CA, USA)

was used for curve-fitting and data were fit to a four parameter logistic model (Head and Kennedy, 2007). Statistical differences among the four EC_{50} values, derived from each of the two independent studies, were tested using one-way analysis of variance (ANOVA) for a representative DLC (TCDD).

Relative sensitivity and relative potency

A detailed description of the calculation of the relative sensitivity (ReS) of COS-7 cells to AHR1 activation by DLCs is provided elsewhere (Farmahin *et al.*, 2012). Briefly, EC_{20}, EC_{50} and EC_{80} values were determined for each concentration-response curve. ReS is defined as $EC_{20,\ 50\ or\ 80\ (compound\ A)}$ of chicken construct ÷ $EC_{20,\ 50\ or\ 80\ (compound\ A)}$ of construct *X*. The relative potency (ReP) of PeCDF or PCB 126 compared to TCDD for each AHR1 construct is defined as: $EC_{20,\ 50\ or\ 80}$ of TCDD determined for construct *X* ÷ $EC_{20,\ 50\ or\ 80}$ of PeCDF *or* PCB 126 determined for construct *X*.

Homology modeling

PSI-BLAST searches for chicken, ring-necked pheasant and Japanese quail AHR1 LBD sequences (residues 235-402) were performed against the Protein Data Bank (PDB) (Berman *et al.*, 2000). The sequences that produced the most significant alignments were identified and nuclear magnetic resonance (NMR) structures were obtained from the PDB. When different NMR structures were found in a PDB file, the most representative structure was determined by use of the NMRCLUST procedure from the On-Line Database Ensemble Representatives and Domains (OLDERADO) (Kelley and Sutcliffe, 1997) server.

The model for the avian AHR1 LBD was generated by Easy Modeller version 2.1 (Kuntal *et al.*, 2010). The protein structure analysis (ProSA) validation method (Wiederstein and Sippl, 2007) (web-based version) was used to assess the quality of the model and the PROCHECK program (Laskowski *et al.*, 1993) was used to assess the stereochemical quality of the models. Templates and structural motifs of models were analysed with the Promotif program (Hutchinson and Thornton, 1996). The amino acid sequence and structure of models were aligned with templates using the DALI server (Holm and Rosenstrom, 2010). The Multiple Alignment of Protein Structures (MultiProt) (Shatsky *et al.*, 2004) server was used for the structural alignment of avian AHR1 models. Molegro Virtual Docker (Thomsen and Christensen, 2006) and UCSF chimera (Pettersen *et al.*, 2004) was used for three dimensional visualization and imaging of AHR1 LBD structures. The computed Atlas of Surface Topography of Proteins (*CASTp)* server (Dundas *et al.*, 2006) was used to analyze the ligand binding cavity. The potential binding site was predicted using a grid-based cavity prediction algorithm by Molegro Virtual Docker (Thomsen and Christensen, 2006).

Results

Avian AHR1 LBD sequences

Sequences of the AHR1 LBD (sites 235-402; 168 amino acids) of 86 avian species were determined by our group or obtained from GenBank (black-footed albatross, great cormorant and red jungle fowl). The species and Genbank

accession numbers of the AHR1 LBD sequences are listed in Table 1. The nucleotide and amino acid identities of the AHR1 LBDs among species were greater than 91% and 96%, respectively. Species of birds were classified into three main types according to the identities of the amino acids at sites 324 and 380: a) type 1 [chicken-like species (Ile324_Ser380; 4 of 86)]; b) type 2 [pheasant-like species (Ile324_Ala380; 47 of 86)]; and c) type 3 [Japanese quail-like species (Val324_Ala380; 35 of 86)]. Four other amino acid sites within the AHR1 LBD (256, 257, 297 and 337) were variable among species and the identities of the amino acids at these sites were used to define 13 sub-types (Table 1). The full-length chicken AHR1 construct (sub-type 1A) was used to generate 12 mutant constructs representative of the other AHR1 sub-types. AHR1 protein expression levels in COS-7 cells transfected with chicken and 12 mutant constructs (1B, 1C, 2A, 2B, 2C, 2D, 2E, 2F, 2G, 3A, 3B, 3C) were similar (Figure 1, panel A). The mutant AHR1 proteins synthesized by IVTT migrated to similar positions on SDS polyacrylamide gels and were expressed to approximately the same degree (Figure S1).

Concentration-dependent effects of TCDD, PeCDF and PCB 126

i. Relative sensitivity (ReS) - inter-construct comparisons

TCDD, PeCDF and PCB 126 elicited concentration-dependent increases in luciferase expression in COS-7 cells containing full-length AHR1 constructs of chicken, pheasant, Japanese quail or mutant AHR1 constructs (Figure 1, panel B; Figures S2 and S3 [normalized luciferase ratios]; Figures S4 and S5 [raw

296 luciferase ratios]). $EC_{20,\ 50,\ and\ 80}$ values were determined from the fitted curves
297 (Table S2) and were used to calculate ReS_{20}, ReS_{50} and ReS_{80} values (Table
298 S3).
299 The rank order of sensitivity of AHR1 constructs to TCDD and PCB 126
300 was type 1 > type 2 > type 3 (Table 2 and Figure S6). For example, in cells
301 exposed to TCDD, ReS_{50s} ranged from 0.75 - 1.0 for the type 1 constructs, 0.061
302 - 0.14 for the type 2 constructs, and 0.0073 - 0.013 for the type 3 constructs
303 (Table 2). In cells exposed to PCB 126, ReS_{50s} were 1.0 - 1.5 for type 1
304 constructs, 0.032 - 0.038 for type 2 constructs and 0.0060 - 0.0091 for type 3
305 constructs. The rank order of ReS_{50} values of cells exposed to PeCDF was also
306 type 1 > type 2 > type 3, but the range was not as great as that observed in cells
307 treated with TCDD or PCB 126. One-way analysis of variance (ANOVA)
308 comparing the four EC_{50} values (derived from the four concentration-response
309 curves/study) obtained for all AHR1 types and sub-types exposed to TCDD
310 indicated that (a) there were significant differences between the main types and
311 (b) within each type, there were no statistical differences among the sub-types
312 (Figure 2).
313 *ii. Relative potency (ReP) - inter-compound comparisons*
314 PeCDF was approximately equipotent to TCDD in cells transfected with type
315 1 AHR1s (ReP_{50} range = 1.2 - 1.9), slightly more potent than TCDD in cells
316 transfected with type 2 AHR1s (ReP_{50} range = 2.3 - 6.8) and substantially more
317 potent than TCDD in cells transfected with type 3 AHR1s (ReP_{50} range = 10 – 21;
318 Table 2). PCB 126 was less potent than TCDD and PeCDF in cells transfected

with all sub-types of AHR1 constructs. For example, PCB 126 was 7- to 14-fold less potent than TCDD in cells transfected with type 1 AHR1, 24- to 58-fold less potent than TCDD in cells transfected with type 2 AHR1 and 11- to 28-fold less potent than TCDD in the cells transfected with type 3 AHR1 (Table 2). The rank order of DLC potency based on $ReP_{average}$ values (the mean value of $ReP_{20,\ 50\ and\ 80}$) of the three compounds studied was (a) PeCDF≈TCDD>PCB 126 for the type 1 constructs, (b) PeCDF≥TCDD>PCB 126 for the type 2 constructs and (c) PeCDF>TCDD>PCB 126 for the type 3 constructs (Table S4; Figure S7).

In vitro - In ovo comparisons

The LRG $EC_{20,\ 50\ and\ 80}$ values obtained in the present study for full-length and mutant AHR1s were compared to LD_{50} values obtained for all domestic and wild avian species that, to our knowledge, have been used for egg injection (*in ovo*) studies (Table S5). LD_{50} values, based on *in ovo* exposures, were significantly correlated ($r^2 = 0.95$, $p < 0.0001$) with *in vitro* EC_{50} values obtained with the LRG assay (Figure 3). The equation obtained from the linear regression was used to predict the sensitivity of type 1, 2 and 3 avian embryos to TCDD, PeCDF and PCB 126 (Table 3). Correlations between LD_{50} values and LRG $EC_{20,\ 50\ and\ 80}$ values are presented (Figure S8); the results indicate that $EC_{20,\ 50\ and\ 80}$ values are all significantly correlated with LD_{50} values. LD_{50} values were similarly correlated with EC_{50} values for both wild-type AHR1 constructs (r^2=0.91, p=0.001; Figure S9, panel A) and mutant AHR1 constructs (r^2=0.93, p<0.0001; Figure S9, panel B) in the LRG assay.

Homology modeling

Homology modeling of avian AHR1 LBD resulted in the models shown in Figures 4 and 5. PSI-BLAST searches revealed that of all of the proteins in the database, the ones with greatest sequence identity with chicken, ring-necked pheasant and Japanese quail AHR1 were (a) HIF-2α (PDB ID: 1P97) and (b) a docked complex containing HIF-2α and ARNT (PDB ID: 2A24). Pairwise sequence identities among avian AHR1 LBDs and 1P97 and 2A24 (Table S6) were similar to those reported for mammalian AHRs and HIF-2α (Pandini *et al.*, 2009). The NMR structures of 1P97 and 2A24 were selected as templates to construct three-dimensional structures of avian AHR1 LBD. These templates were also used by others to construct AHR LBD models for mouse (Miyagi *et al.*, 2012) and rat (Yoshikawa *et al.*, 2010). The region between amino acids 283 and 389 of avian AHR1 was used to develop two-template models for chicken, pheasant and quail AHR1 LBDs. The 1P97 and 2A24 PBD files each contained 20 different candidate structures. The most representative candidate structures determined by ORLANDO were structure 17 for 1P97 and structure 11 for 2A24.

The ProSA z-scores were -4, -3.5 and -3.9 for chicken, pheasant and quail, respectively (Figure S10, panel A). The z-score indicates overall model quality and measures the deviation of the total energy of the structure from an energy distribution derived from random conformations (Wiederstein and Sippl, 2007). The z-scores for the three avian AHR1 LBD structures are within the range of z-scores typically found for protein NMR structures of similar size. The

three avian models passed all criteria implemented by PROCHECK (which determines the stereo-chemical quality of models). Approximately 93%, 88% and 89% of amino acid residues in chicken, pheasant, and quail, respectively, reside in the “most favored” areas of the Ramachandran plots (90% for structures solved at a resolution of 2.0 Å), with only one residue (HIS325) in a “disallowed” region (Figure S10, panel B). The overall *G*-factors, which measure stereochemical quality, were -0.12, -0.15 and -0.19 (from -0.5 to 0.3 for structures solved at 1.5 Å resolution) for chicken, pheasant and quail, respectively.

The structures of chicken, pheasant and quail AHR1 LBD models were aligned with 1P97 and 2A24 by use of the DALI (Holm and Rosenstrom, 2010) server (Figure 4, panel A). Chicken, pheasant and quail models each contained five *β*-sheets and a central helix. The models of the AHR1 LBDs of quail, chicken and pheasant contained one, two or three short helices, respectively (Figure 4, panel A). The nomenclature (Gong *et al.*, 1998) for helices and beta sheets of the FixL protein PAS domain was used for avian AHR1 LBD models. The avian AHR1 LBD models were aligned using (1) the Multiprot server for visual comparison (Figure S11) and (2) the DALI server for obtaining average pairwise root-mean-square deviation (RMSD) values. The results showed high similarity between chicken, pheasant and quail with RMSD values ranging from 0.5 - 0.7 Å. RMSD is the measure of the average distance between the atoms of superimposed proteins and shows the similarity between two structures. For two perfectly identical structures, the RMSD value would be 0 Å; for two randomly chosen dissimilar proteins, the RMSD would likely be 10 Å or greater. The side

chains of amino acids at positions 324 and 380 in chicken, pheasant and quail models face within the binding pocket, but the side chains of amino acids at positions 297 and 337 face away from the binding pocket (Figure 4, panel B).

Potential binding site locations were identified by use of the Molegro cavity detection algorithm, and ligand-binding sites for chicken, pheasant and quail were located close to the centre of the side chains of amino acids 324 and 380 (Figure 5). The main cavities, identified by use of CASTp, were buried in the core of the avian models, and were delimited by the helices, β-sheets and their connecting loops. A list of amino acid residues with side chains that contribute to the internal cavity surface in the avian models was extracted from the CASTp output and compared (Figure 4, panel C). The cavity volumes were found to be 399, 501 and 542 $Å^3$ for chicken, pheasant and quail, respectively. *In silico* mutagenesis was performed and the chicken AHR1 LBD model was targeted at positions 324 and 380 to generate three mutant models (I324V, S380A and I324_S380A). CASTp analysis showed that mutation of amino acid 324 from Ile to Val (I324V) increased the cavity volume of the chicken AHR1 LBD model from 399 $Å^3$ to 485 $Å^3$. Our results also show that the amino acid at position 380 contributes to cavity volume. For example, mutation at position 380 (serine to alanine; S380A) caused expansion of the cavity volume of chicken from 399 $Å^3$ to 465 $Å^3$. When both positions 324 and 380 were changed (I324V_S380A), the cavity volume of mutant chicken AHR1 LBD was 528 $Å^3$ (Table S7).

<u>The six variable amino acids within AHR1 LBD</u>

Site 256 Site 256 was not within the templates (1P97 and 2A24) used to generate avian AHR1 LBD homology models. The amino acid residue at this site was either alanine (11 sub-types) or threonine (2G and 3C). The identity of the amino acid at site 256 does not appear to be associated with differential AHR1 transactivation after exposure of transfected cells to DLCs. In support of this conclusion, the only difference between the 2A and 2G AHR1 LBDs is the identity of the amino acid at site 256 (2A, alanine; 2G, threonine), and cells containing these constructs did not differ in sensitivity to AHR1 activation by TCDD (Figure 2). Similarly, the only difference between 3B and 3C is the amino acid at site 256 (3B, alanine; 3C, threonine), and cells containing these constructs did not differ in sensitivity to AHR1 activation by TCDD (Figure 2).

Site 257 The amino acid residue at site 257 was alanine (7 sub-types), threonine (5 sub-types) or proline (2D). Similar to site 256, site 257 was not within the templates used to generate models, nor did the identity of the amino acid at this site affect the sensitivity of AHR1 to transactivation in the LRG assay. For example, the only difference between 2B and 2E is the amino acid at site 257 (2B, alanine; 2E, threonine), and cells containing these constructs did not differ in sensitivity to AHR1 activation by TCDD (Figure 2). Similarly, 2D and 2F, which differed only at site 257 (2D, proline; 2F, threonine) did not differ in sensitivity to AHR1 activation by TCDD. Importantly, alteration of threonine-258 to alanine-258 in common tern AHR1 (site 258 in tern corresponds with site 257 in most other avian species) did not alter the binding affinity to AHR1 or AHR1-mediated reporter gene activity in COS-7 cells exposed to TCDD (Karchner *et al.*, 2006).

Site 297 The amino acid residue at site 297 was threonine in the three sub-types of type 1 species and also in the three sub-types of type 3 species. In type 2 species, site 297 was isoleucine (2A and 2G), threonine (4 sub-types) or valine (2C). The identity of the amino acid at site 297 did not affect the sensitivity of AHR1 to transactivation in the LRG assay. For example, the only difference between 2A and 2C is the identity of the amino acid at site 297 (2A, isoleucine; 2C, valine), and there was no statistically significant difference between concentration-response effects of TCDD in cells expressing these sub-types (Figure 2). Homology modeling showed that amino acid position 297 in chicken, pheasant and quail AHR1 is located within one of the beta sheets (Bβ, Figure 4, panel A) and the side chain of this amino acid in the three avian models is directed away from the ligand binding pocket (Figures 4 and 5) and does not contribute to the cavity surface (Figure 4, panel C).

Site 324 Site 324 is occupied with either isoleucine (type 1 and type 2 species) or valine (type 3 species). The results of the present study showed that the identity of the amino acid at site 324 has a key effect on AHR1 activation by DLCs; this is in agreement with the results of earlier studies (Farmahin *et al.*, 2012; Head *et al.*, 2008; Karchner *et al.*, 2006). For example, the only difference between 2F and 3B AHR1 LBDs is the identity of the amino acid at site 324 (2F, isoleucine; 3B, valine), and there was a statistically significant difference (9-fold; $p<0.001$) between the concentration-dependent effect of TCDD on AHR1 activation in cells expressing these sub-types (Figure 2). The study by Farmahin *et al.* found that alteration of valine in Japanese quail AHR1 to isoleucine (V324I)

was responsible for a 12–fold increase in AHR-dependent luciferase activity in cells exposed to TCDD. The isoleucine and valine side chains are both hydrophobic but the isoleucine side chain (-CH-(CH3, CH2-CH3)) is longer than the valine side chain (-CH-(CH3)2). Interestingly, a study that used site-directed mutagenesis to change isoleucine to alanine at site 319 in mouse AHR (mouse site 319 and avian site 324 are equivalent) resulted in complete loss of the ability of AHR to bind to TCDD and DNA (Pandini *et al.*, 2009). As such, one might predict that site-directed mutagenesis of valine to alanine (-CH3) in avian type 3 AHR1 LBD constructs might also result in complete loss of activity in the LRG assay.

Homology modeling showed that amino acid residue 324 in chicken, pheasant and quail is located in a loop between Eα and Fα (Figure 4, panel A) of the AHR1 LBD, and the side chain of this residue contributes to the cavity surface (Figure 4, panel C). *In silico* mutagenesis of this residue from Ile-324 to Val-324 in chicken AHR1 resulted in an increase in the cavity volume from 399 $Å^3$ to 485 $Å^3$. Taken together, these results suggest that Val-324 is responsible for reducing the sensitivity of type 3 AHR1, perhaps by increasing the binding cavity volume and weakening the ligand-receptor interaction. Previous studies have shown the importance of this position in mouse (Goryo *et al.*, 2007) and tern (Karchner *et al.*, 2006).

Site 337 The amino acid residue at site 337 was valine (9 sub-types) or isoleucine (4 sub-types), and based on the results of the present study, the identity of the amino acid at this site does not contribute to differences in

concentration-dependent effects of DLCs. For example, the only difference between 2E and 2F is the identity of the amino acid at this site (2E, isoleucine; 2F, valine), but there was no statistically significant difference in the concentration-dependent effects of TCDD on AHR1 activation for constructs containing either amino acid residue in the LRG assay (Figure 2). Similarly, the only difference between types 1A and 1B is the amino acid at site 337 (1A, valine; 1B, isoleucine), and cells containing these constructs did not differ in sensitivity to AHR1 activation by TCDD (Figure 2).

Homology modeling showed that Val-337 in chicken, pheasant and quail AHR1 LBD is located in the helical connector (Fα). The finding that the identity of the amino acid residue at site 337 does not affect sensitivity to AHR1 activation is not surprising because the amino acid side chain of this residue does not contribute to the cavity surface (Figure 4, panel C) and points towards the outside of the binding cavity (Figure 4, panel B).

Site 380 The amino acid residue at site 380 was either serine (type 1 species) or alanine (type 2 and type 3 species). The identity of the amino acid at site 380 has a significant effect on differential sensitivity of AHR1 activation by DLCs, in agreement with the results of earlier studies (Farmahin *et al.*, 2012; Head *et al.*, 2008; Karchner *et al.*, 2006). For example, the only difference between 1B and 2B is the identity of the amino acid at site 380 (1B, serine; 2B, alanine), and there was a statistically significant difference (13-fold; $p<0.001$) between the concentration-dependent effects of TCDD on AHR1 activation in cells expressing these sub-types (Table 2, Figure 2). The results of a previous

study (Farmahin *et al.*, 2012) showed that changing Ala-380 to Ser-380 in Japanese quail AHR1 increased sensitivity to AHR1 activation 25-fold and 3.5-fold with TCDD and PeCDF, respectively. It has also been shown that mutation of alanine to serine at this site in common tern AHR1 increased the binding affinity of TCDD to AHR1 and the ability to transactivate a luciferase reporter gene (Karchner *et al.*, 2006). The importance of this amino acid in other vertebrates has been shown in several other studies (Backlund and Ingelman-Sundberg, 2004; Ema *et al.*, 1994; Murray *et al.*, 2005; Pandini *et al.*, 2007; Pandini *et al.*, 2009; Poland *et al.*, 1994; Ramadoss and Perdew, 2004).

Homology modeling showed that the amino acid residue 380 in avian AHR1 forms part of the antiparallel β-sheet (I β) and the amino acid side chain at this site contributes to the binding cavity of the AHR1 (Figures 4, panel B and C). The reason that serine at position 380 enhances the AHR1 transactivation ability is possibly due to hydrogen-bonding interactions of the hydroxyl group of the serine side chain with the dioxygen bridge or chlorine atom of the ligand, which could stabilize the ligand-receptor interaction.

Discussion

In this study, amino acid sequences of the AHR1 LBD from 86 avian species were determined and compared. Twelve mutant AHR1 constructs were generated and the sensitivity of the constructs to AHR1 activation by TCDD, PeCDF and PCB 126 was determined by use of the LRG assay. The results

revealed that AHR1 activation is controlled by the identity of amino acids at sites 324 and 380 in the AHR1 LBD. A statistically significant correlation between *in vitro* AHR1 activation and *in ovo* toxicity of DLCs provided conclusive evidence that the identity of amino acids 324 and 380 in the AHR1 LBD can be used to predict the relative sensitivity of avian species to DLC toxicity (LD_{50s}).

Among avian species, the AHR1 LBD was first shown to impart distinct *in vitro* functional properties in chicken and common tern (Karchner *et al.*, 2006). Only two amino acids, located at sites 324 and 380 in chicken AHR1 LBD, are responsible for differences observed in TCDD binding affinity between tern and chicken AHR1. These two amino acids are also responsible for the distinct abilities of chicken, pheasant and quail AHR1 to transactivate a luciferase reporter gene in cells exposed to TCDD, PeCDF or TCDF (Farmahin *et al.*, 2012). In addition, the LBD amino acid sequence and *in vitro* function of chicken, ring-necked pheasant and Japanese quail AHR1 predict *in ovo* toxicity (LD_{50}) of these three species to TCDD, PeCDF and TCDF (Farmahin *et al*., 2012). To our knowledge, there are no studies which demonstrate that the identities of amino acids within the DBD or TAD of avian AHR1 cause differential sensitivity to DLCs. However, based on findings in mammals (Ishiniwa *et al.*, 2010; Minsavage *et al.*, 2004; Pohjanvirta *et al.*, 1998), it is conceivable that amino acid deletions in an important segment of the DBD or TAD of avian AHR1 or a mutation in a critical and conservative amino acid position in these domains could affect avian sensitivity to DLCs. The determination of such potential differences in the DBD and TAD of avian AHR1 was beyond the scope of this study.

549 The objective of the current study was to establish a method to predict *in*
550 *ovo* sensitivity to the toxic effects of DLCs for a range of birds from diverse taxa
551 and feeding guilds, based on their AHR1 LBD sequence. Here, we showed that
552 AHR1 LBDs from 86 avian species belong to one of 13 different sub-types, and
553 that, when these AHR1s are expressed in COS-7 cells, the sensitivity of each
554 AHR1 to activation by TCDD, PeCDF or PCB-126 is (1) determined by the
555 identity of amino acids 324 and 380 and (2) significantly correlated with the
556 sensitivity of each species to *in ovo* toxicity (LD_{50}). The current study confirmed
557 our previous finding (Farmahin *et al.*, 2012; Head *et al.*, 2008; Karchner *et al.*,
558 2006) that birds can be classified into three major types based on their sensitivity
559 to the toxic and biochemical effects of DLCs (chicken-like, pheasant-like and
560 quail-like), and that the assignment to these three classes can be predicted from
561 the identity of amino acids at sites 324 and 380 in the AHR1 LBD. The results
562 reported here also revealed that the other four variable amino acid sites within
563 the LBD (256, 257, 297 and 337) do not affect the ability of AHR1 to transactivate
564 the AHRE-driven reporter gene when exposed to DLCs. Among avian AHR1
565 constructs, the rank order in sensitivity to induce AHRE-driven reporter gene
566 activity was: type 1(1A, 1B and 1C) > type 2 (2A, 2B, 2C, 2D, 2E, 2F and 2G) >
567 type 3 (3A, 3B and 3C). These results are consistent with those of a similar study
568 conducted in our laboratory that employed the LRG assay to determine avian
569 species sensitivity to AHR1 activation by several PCB congeners (Manning *et al.*,
570 2012).

571 The ReP of PeCDF was very similar to the ReP of TCDD (i.e.,
572 approximately 1.0) for AHR1 activation in cells expressing various mutants of
573 type 1 AHR1s. However, PeCDF was more potent than TCDD in cells expressing
574 type 2 AHR1s (ReP_{50}=2.3 to 6.8) and type 3 AHR1s (ReP_{50}=10 to 28). This
575 observation is similar to previous results showing that PeCDF was more potent
576 than TCDD in cells containing full-length ring-necked pheasant (2G), Japanese
577 quail (3A) or common tern (3B) AHR1 constructs (Farmahin *et al.*, 2012). The
578 ReP values obtained from the LRG assay are consistent with ReP values that
579 have been calculated by other studies such as *in ovo* toxicity (Cohen-Barnhouse
580 *et al.*, 2011) and induction of ethoxyresorufin *O*-deethylase (EROD) and
581 expression of CYP1A mRNA in primary hepatocyte cultures of chicken,
582 pheasant, quail and herring gull (Herve *et al.*, 2010a; Herve *et al.*, 2010b). Taken
583 together, the data from all of these studies are important because the current
584 assumption in risk assessments is to consider PeCDF to be as potent as TCDD,
585 which is correct only for type 1 birds.
586 Alteration of amino acids at sites 324 and 380 resulted in larger
587 differences in sensitivity between AHR1 constructs exposed to TCDD, compared
588 to the differences in sensitivity observed between AHR1 constructs exposed to
589 PeCDF. These apparent ligand-dependent differences are perhaps due to the
590 larger molecular size of PeCDF. Type 2 and 3 AHR1 LBDs contain larger binding
591 cavity volumes than type 1 AHR1 LBDs. As such, PeCDF might conform better to
592 the binding cavity of type 2 and type 3 AHR1s than TCDD. In addition, PeCDF
593 contains an additional electronegative Cl group that might contribute to hydrogen

bonding with Val-324 and Ala-380. Detailed docking studies of TCDD, PeCDF and other AHR agonists with avian AHR1 LBDs might be able to determine why PeCDF is more potent than TCDD in type 2 and type 3 avian species.

A significant and positive correlation (r^2=0.77, p<0.005) was previously observed between LD_{50} values for DLCs in chicken, pheasant, and Japanese quail embryos, and EC_{50} values from the LRG assay (Farmahin *et al.*, 2012). In the current study, mutant AHR1s that contained LBD sequences matching all 13 sub-types of avian species were developed by use of site-directed mutagenesis. Determination of EC_{50} values in cells expressing mutant AHR1s allowed comparisons with *in ovo* results for other species. A significant, positive correlation (r^2=0.95, p<0.0001; Figure 3) was observed between the *in vitro* LRG assay and *in ovo* toxicity data. This strong correlation was used to predict the sensitivity of a larger group of birds to TCDD, PeCDF and PCB 126 (Table 3). These predicted LD_{50} values could be used to estimate the embryotoxic effects of DLCs for wild birds in site-specific risk assessments. Our findings also suggest that amino acids within the DBD and TAD of AHR1s have no effect on the sensitivity of birds to the DLCs studied because the linear regression equations between wild-type or mutant AHR1 LRG EC_{50} data and *in ovo* LD_{50s} had similar slopes and Y-intercepts. Use of the EC_{20}, EC_{50} or EC_{80} gives similar correlations with LD_{50} data, thus any of these endpoints could be used to predict the embryotoxic effects of DLCs (Figure S8). Taken together, the results of the present study have (1) confirmed and extended the results of Karchner *et al.* (2006) showing that differential sensitivity of chicken and common tern AHR1 to

activation by TCDD reside in the ligand binding domain, (2) confirmed and extended the results of Farmahin *et al.* (2012) that LD_{50s} can be predicted from the LRG assay in cells expressing wild-type AHR1 and (3) suggested that LD_{50s} can be predicted using the LRG assay for cells expressing mutant AHR1. The predictive relationship was valid for all avian species for which LRG assay and *in ovo* toxicity data are available.

In conclusion, the findings from this study suggest that the sequence of the AHR1 LBD can predict the sensitivity of all avian species to DLCs. This approach, unlike other *in vitro* methods such as induction of EROD activity and measurement of CYP1A mRNA expression (Head and Kennedy, 2007; Kennedy *et al.*, 1996), does not require the bird of interest to be euthanized, because the AHR1 LBD sequence can be determined by using a drop of blood (Head *et al.*, 2010). Sequencing of the AHR1 LBD is likely to be useful in identifying the most susceptible avian species in ecological risk assessments.

Supplementary Data Attached

Funding Information

This research was supported by unrestricted grants from the Dow Chemical Company and Georgia-Pacific LLC to the University of Ottawa, Environment Canada's STAGE program and, in part, by a Discovery Grant from the National Science and Engineering Research Council of Canada (Project # 326415-07). The authors wish to acknowledge the support of an instrumentation grant from

the Canada Foundation for Infrastructure. Professor Giesy was supported by the Canada Research Chair program and an at large Chair Professorship at the Department of Biology and Chemistry and State Key Laboratory in Marine Pollution, City University of Hong Kong, the Einstein Professor Program of the Chinese Academy of Sciences and the Visiting Professor Program of King Saud University. M. Hahn and S. Karchner were supported by NOAA Sea Grant (grant number NA06OAR4170021 (R/B-179)), and by the Walter A. and Hope Noyes Smith endowed chair.

References

Backlund, M. and Ingelman-Sundberg, M. (2004). Different structural requirements of the ligand binding domain of the aryl hydrocarbon receptor for high- and low-affinity ligand binding and receptor activation. *Mol. Pharmacol.* **65**, 416-425.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235-242.

Brunstrom, B. (1988). Sensitivity of embryos from duck, goose, herring gull, and various chicken breeds to 3,3',4,4'-tetrachlorobiphenyl. *Poult. Sci* **67**, 52-57.

Cohen-Barnhouse, A. M., Zwiernik, M. J., Link, J. E., Fitzgerald, S. D., Kennedy, S. W., Herve, J. C., Giesy, J. P., Wiseman, S., Yang, Y., Jones, P. D., Wan, Y., Collins, B., Newsted, J. L., Kay, D., and Bursian, S. J. (2011). Sensitivity of Japanese quail (*Coturnix japonica*), common pheasant (*Phasianus colchicus*) and white leghorn chicken (*Gallus gallus domesticus*) embryos to in ovo exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), 2,3,4,7,8-pentachlorodibenzofuran (PeCDF) and 2,3,7,8-tetrachlorodibenzofuran (TCDF). *Toxicol. Sci.* **119**, 93-103.

Denison, M. S., Soshilov, A. A., He, G., Degroot, D. E., and Zhao, B. (2011). Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicol. Sci.* **124**, 1-22.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **34**, W116-W118.

Ema, M., Ohe, N., Suzuki, M., Mimura, J., Sogawa, K., Ikawa, S., and Fujii-Kuriyama, Y. (1994). Dioxin binding activities of polymorphic forms of mouse and human arylhydrocarbon receptors. *J Biol. Chem* **269**, 27337-27343.

Farmahin, R., Wu, D., Crump, D., Herve, J. C., Jones, S. P., Hahn, M. E., Karchner, S. I., Giesy, J. P., Bursian, S. J., Zwiernik, M. J., and Kennedy, S. W. (2012). Sequence and in vitro function of chicken, ring-necked pheasant and Japanese quail AHR1 predict in vivo sensitivity to dioxins. *Environ. Sci. Technol.* **46**, 2967-2975.

Gibson, J. F., Kelso, S., and Skevington, J. H. (2010). Band-cutting no more: A method for the isolation and purification of target PCR bands from multiplex PCR products using new technology. *Molecular Phylogenetics and Evolution* **56**, 1126-1128.

Gong, W. M., Hao, B., Mansy, S. S., Gonzalez, G., Gilles-Gonzalez, M. A., and Chan, M. K. (1998). Structure of a biological oxygen sensor: A new mechanism for heme-driven signal transduction. *Proc. Nati. Acad. Sci. USA* **95**, 15177-15182.

Goryo, K., Suzuki, A., Del Carpio, C. A., Siizaki, K., Kuriyama, E., Mikami, Y., Kinoshita, K., Yasumoto, K., Rannug, A., Miyamoto, A., Fujii-Kuriyama, Y., and Sogawa, K. (2007). Identification of amino acid residues in the Ah receptor involved in ligand binding. *Biochem. Biophys. Res. Commun.* **354**, 396-402.

Head, J. A. (2006). Variation in the cytochrome P4501A responses to Dioxin-Like Compounds in avian species, Ph.D.Dissertation. *University of Ottawa*.

Head, J. A., Farmahin, R., Kehoe, A. S., O'Brien, J. M., Shutt, J. L., and Kennedy, S. W. (2010). Characterization of the avian aryl hydrocarbon receptor 1 from blood using non-lethal sampling methods. *Ecotoxicology.* **19**, 1560-1566.

Head, J. A., Hahn, M. E., and Kennedy, S. W. (2008). Key amino acids in the aryl hydrocarbon receptor predict dioxin sensitivity in avian species. *Environ. Sci. Technol.* **42**, 7535-7541.

Head, J. A. and Kennedy, S. W. (2007). Same-sample analysis of ethoxyresorufin-O-deethylase activity and cytochrome P4501A mRNA abundance in chicken embryo hepatocytes. *Anal. Biochem.* **360**, 294-302.

Head, J. A. and Kennedy, S. W. (2010). Correlation between an in vitro and an in vivo measure of dioxin sensitivity in birds. *Ecotoxicology* **19**, 377-382.

Herve, J. C., Crump, D., Jones, S. P., Mundy, L. J., Giesy, J. P., Zwiernik, M. J., Bursian, S. J., Jones, P. D., Wiseman, S. B., Wan, Y., and Kennedy, S. W. (2010a). Cytochrome P4501A induction by 2,3,7,8-tetrachlorodibenzo-p-dioxin and two chlorinated dibenzofurans in primary hepatocyte cultures of three avian species. *Toxicol. Sci.* **113**, 380-391.

Herve, J. C., Crump, D. L., McLaren, K. K., Giesy, J. P., Zwiernik, M. J., Bursian, S. J., and Kennedy, S. W. (2010b). 2,3,4,7,8-pentachlorodibenzofuran is a more potent cytochrome P4501A inducer than 2,3,7,8-tetrachlorodibenzo-p-dioxin in herring gull hepatocyte cultures. *Environ. Toxicol. Chem.* **29**, 2088-2095.

Hoffman, D. J., Melancon, M. J., Klein, P. N., Eisemann, J. D., and Spann, J. W. (1998). Comparative developmental toxicity of planar polychlorinated biphenyl congeners in chickens, American kestrels and common terns. *Environ. Toxicol. Chem.* **17**, 747-757.

Holm, L. and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545-W549.

738 Hutchinson, E. G. and Thornton, J. M. (1996). PROMOTIF--a program to identify
739 and analyze structural motifs in proteins. *Protein Sci.* **5**, 212-220.

740 Ishiniwa, H., Sogawa, K., Yasumoto, K., and Sekijima, T. (2010). Polymorphisms
741 and functional differences in aryl hydrocarbon receptors (AhR) in Japanese field
742 mice, Apodemus speciosus. *Environmental Toxicology and Pharmacology* **29**,
743 280-289.

744 Karchner, S. I., Franks, D. G., Kennedy, S. W., and Hahn, M. E. (2006). The
745 molecular basis for differential dioxin sensitivity in birds: role of the aryl
746 hydrocarbon receptor. *Proc. Natl. Acad. Sci U. S. A* **103**, 6252-6257.

747 Kelley, L. A. and Sutcliffe, M. J. (1997). OLDERADO: On-line database of
748 ensemble representatives and domains. *Protein Sci.* **6**, 2628-2630.

749 Kennedy, S. W., Lorenzen, A., Jones, S. P., Hahn, M. E., and Stegeman, J. J.
750 (1996). Cytochrome P4501A induction in avian hepatocyte cultures: a promising
751 approach for predicting the sensitivity of avian species to toxic effects of
752 halogenated aromatic hydrocarbons. *Toxicol Appl. Pharmacol.* **141**, 214-230.

753 Kuntal, B. K., Aparoy, P., and Reddanna, P. (2010). EasyModeller: A graphical
754 interface to MODELLER. *BMC. Res. Notes* **3**, 226-230.

755 Laskowski, R. A., MacArthur M W, Moss D, and Thornton J M (1993).
756 PROCHECK: a program to check the stereochemical quality of protein structures.
757 *J. Appl. Cryst.* **26**, 283-291.

758 Lee, J. S., Kim, E. Y., and Iwata, H. (2009). Dioxin activation of CYP1A5
759 promoter/enhancer regions from two avian species, common cormorant
760 (*Phalacrocorax carbo*) and chicken (*Gallus gallus*): association with aryl
761 hydrocarbon receptor 1 and 2 isoforms. *Toxicol. Appl. Pharmacol.* **234**, 1-13.

762 Manning, G. E., Farmahin, R., Crump, D., Jones, S. P., Klein, J., Konstantinov,
763 A., Potter, D., and Kennedy, S. W. (2012). A luciferase reporter gene assay and
764 avian AHR1 genotype predict the embryolethality of polychlorinated biphenyls in
765 avian species. *Toxicol Appl. Pharmacol. (accepted)*.

766 Minsavage, G. D., Park, S. K., and Gasiewicz, T. A. (2004). The aryl hydrocarbon
767 receptor (AhR) tyrosine 9, a residue that is essential for AhR DNA binding
768 activity, is not a phosphoresidue but augments AhR phosphorylation. *Journal of*
769 *Biological Chemistry* **279**, 20582-20593.

770 Miyagi, S., Sawamura, S., Yoshikawa, E., Dedachi, K., Itoh, S., Ishihara-Sugano,
771 M., and Kurita, N. (2012). Ab initio fragment molecular orbital calculations on
772 specific interactions between aryl hydrocarbon receptor and dioxin. *Int. J.*
773 *Quantum Chem.* **112**, 289-299.

774 Mol, T. L., Kim, E. Y., Ishibashi, H., and Iwata, H. (2012). In vitro transactivation
775 potencies of black-footed albatross (Phoebastria nigripes) AHR1 and AHR2 by
776 dioxins to predict CYP1A expression in the wild population. *Environ. Sci.*
777 *Technol.* **46**, 525-533.

778 Murray, I. A., Reen, R. K., Leathery, N., Ramadoss, P., Bonati, L., Gonzalez, F.
779 J., Peters, J. M., and Perdew, G. H. (2005). Evidence that ligand binding is a key
780 determinant of Ah receptor-mediated transcriptional activity. *Arch. Biochem.*
781 *Biophys.* **442**, 59-71.

782 Okey, A. B. (2007). An aryl hydrocarbon receptor odyssey to the shores of
783 toxicology: the Deichmann Lecture, International Congress of Toxicology-XI.
784 *Toxicol. Sci* **98**, 5-38.

785 Pandini, A., Denison, M. S., Song, Y., Soshilov, A. A., and Bonati, L. (2007).
786 Structural and functional characterization of the aryl hydrocarbon receptor ligand
787 binding domain by homology modeling and mutational analysis. *Biochemistry.* **46**,
788 696-708.

789 Pandini, A., Soshilov, A. A., Song, Y., Zhao, J., Bonati, L., and Denison, M. S.
790 (2009). Detection of the TCDD binding-fingerprint within the Ah receptor ligand
791 binding domain by structurally driven mutagenesis and functional analysis.
792 *Biochemistry.* **48**, 5972-5983.

793 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M.,
794 Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera--a visualization system for
795 exploratory research and analysis. *J Comput. Chem* **25**, 1605-1612.

796 Pohjanvirta, R., Wong, J. M., Li, W., Harper, P. A., Tuomisto, J., and Okey, A. B.
797 (1998). Point mutation in intron sequence causes altered carboxyl-terminal
798 structure in the aryl hydrocarbon receptor of the most 2,3,7,8-tetrachlorodibenzo-
799 p-dioxin-resistant rat strain. *Mol. Pharmacol.* **54**, 86-93.

800 Poland, A., Glover, E., and Kende, A. S. (1976). Stereospecific, high affinity
801 binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin by hepatic cytosol. Evidence that
802 the binding species is receptor for induction of aryl hydrocarbon hydroxylase. *J*
803 *Biol. Chem* **251**, 4936-4946.

804 Poland, A., Palen, D., and Glover, E. (1994). Analysis of the four alleles of the
805 murine aryl hydrocarbon receptor. *Mol. Pharmacol.* **46**, 915-921.

806 Ramadoss, P. and Perdew, G. H. (2004). Use of 2-azido-3-[125I]iodo-7,8-
807 dibromodibenzo-p-dioxin as a probe to determine the relative ligand affinity of
808 human versus mouse aryl hydrocarbon receptor in cultured cells. *Mol.*
809 *Pharmacol.* **66**, 129-136.

810 Shatsky, M., Nussinov, R., and Wolfson, H. J. (2004). A method for simultaneous
811 alignment of multiple protein structures. *Proteins* **56**, 143-156.

812 Thomsen, R. and Christensen, M. H. (2006). MolDock: A new technique for high-
813 accuracy molecular docking. *Journal of Medicinal Chemistry* **49**, 3315-3321.

814 U.S.EPA (1994). Method 1613: tetra- through octa-chlorinated dioxins and furans
815 by isotope dilution HRGC/HRMS. *U S Environmental Protection Agency Office of*
816 *Water*.

817 Van den Berg, M., Birnbaum, L., Bosveld, A. T., Brunstrom, B., Cook, P., Feeley,
818 M., Giesy, J. P., Hanberg, A., Hasegawa, R., Kennedy, S. W., Kubiak, T., Larsen,
819 J. C., van Leeuwen, F. X., Liem, A. K., Nolt, C., Peterson, R. E., Poellinger, L.,
820 Safe, S., Schrenk, D., Tillitt, D., Tysklind, M., Younes, M., Waern, F., and
821 Zacharewski, T. (1998). Toxic equivalency factors (TEFs) for PCBs, PCDDs,
822 PCDFs for humans and wildlife. *Environ. Health Perspect.* **106**, 775-792.

823 Van den Berg, M., Birnbaum, L. S., Denison, M., De Vito, M., Farland, W.,
824 Feeley, M., Fiedler, H., Hakansson, H., Hanberg, A., Haws, L., Rose, M., Safe,
825 S., Schrenk, D., Tohyama, C., Tritscher, A., Tuomisto, J., Tysklind, M., Walker,
826 N., and Peterson, R. E. (2006). The 2005 World Health Organization reevaluation
827 of human and mammalian toxic equivalency factors for dioxins and dioxin-like
828 compounds. *Toxicol Sci* **93**, 223-241.

829 Wiederstein, M. and Sippl, M. J. (2007). ProSA-web: interactive web service for
830 the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids*
831 *Res.* **35**, W407-W410.

832 Yasui, T., Kim, E. Y., Iwata, H., Franks, D. G., Karchner, S. I., Hahn, M. E., and
833 Tanabe, S. (2007). Functional characterization and evolutionary history of two
834 aryl hydrocarbon receptor isoforms (AhR1 and AhR2) from avian species.
835 *Toxicol. Sci* **99**, 101-117.

836 Yasui, T., Kim, E. Y., Iwata, H., and Tanabe, S. (2004). Identification of aryl
837 hydrocarbon receptor 2 in aquatic birds; cDNA cloning of AHR1 and AHR2 and
838 characteristics of their amino acid sequences. *Mar. Environ Res.* **58**, 113-118.

839 Yoshikawa, E., Miyagi, S., Dedachi, K., Ishihara-Sugano, M., Itoh, S., and Kurita,
840 N. (2010). Specific interactions between aryl hydrocarbon receptor and dioxin
841 congeners: ab initio fragment molecular orbital calculations. *J. Mol. Graph.*
842 *Model.* **29**, 197-205.
843
844

845
846
847

848

Table 1. Main types (**type 1**: Ile324_Ser380; **type 2**: Ile324_Ala380; **type 3**: Val324_Ala380) and sub-types (letter designations) of AHR1 LBDs determined in 86 species of birds. The identities of amino acids at sites 256, 257, 297, 324, 337 and 380 are indicated; sites 324 and 380 are in bold. The AHR1 LBD sequences were determined by our group or obtained from GenBank (black-footed albatross, great cormorant and red jungle fowl). The red jungle fowl and the domestic chicken are considered to be the same species. AHR1 LBD sequences were determined from liver or blood samples obtained from the National Wildlife Research Centre, Ottawa, ON (underlined), local suppliers (domestic chicken, Japanese quail, ring-necked pheasant, turkey and emu), the Tittabawassee River basin, Michigan, USA (bold) or approximately 100 km west of the Tittabawassee River basin (bold italic). Previously reported AHR LBD sequences were from:*(Head, 2006; Head *et al.*, 2008) **(Karchner *et al.*, 2006).

sub-type	aa identity 256	257	297	324	337	380	name	scientific name	accession #
1A	A	A	T	I	V	S	domestic chicken	*Gallus gallus domesticus*	NM204118
1A							red jungle fowl	*Gallus gallus*	NC006089
1B	A	A	T	I	I	S	**European starling**	*Sturnus vulgaris*	FJ376538
1B							**ruby-throated hummingbird**	*Archilochus colubris*	FJ376541
1C	A	T	T	I	I	S	**gray catbird**	*Dumetella carolinensis*	FJ376509
2A	A	A	I	I	V	A	ruffed grouse*	*Bonasa umbellus*	FJ376525
2A							turkey*	*Meleagris gallopavo*	EU660874
2A							willow ptarmigan*	*Lagopus lagopus*	FJ376532
2B	A	A	T	I	I	A	***American redstart***	*Setophaga ruticilla*	JQ814789
2B							**Baltimore oriole**	*Icterus galbula*	FJ376487
2B							**black-and-white warbler**	*Mniotilta varia*	FJ376491
2B							**black-capped chickadee**	*Poecile atricapilla*	FJ376492
2B							**brown-headed cowbird**	*Molothrus ater*	FJ376495
2B							**cedar waxwing**	*Bombycilla cedrorum*	FJ376497
2B							**chipping sparrow**	*Spizella passerina*	FJ376498
2B							**common grackle**	*Quiscalus quiscula*	FJ376501
2B							**common yellowthroat**	*Geothlypis trichas*	FJ376503
2B							**indigo bunting**	*Passerina cyanea*	FJ376513
2B							**Northern cardinal**	*Cardinalis cardinalis*	FJ376516
2B							**ovenbird**	*Seiurus aurocapilla*	FJ376518
2B							**red-winged blackbird**	*Agelaius phoeniceus*	FJ376521
2B							**rose-breasted grosbeak**	*Pheucticus ludovicianus*	FJ376524
2B							***song sparrow***	*Melospiza melodia*	JQ824841
2B							**swamp sparrow**	*Melospiza georgiana*	FJ376528
2B							**tufted titmouse**	*Baeolophus bicolor*	FJ376536
2B							***white-throated sparrow***	*Zonotrichia albicollis*	JQ814794
2C	A	A	V	I	V	A	rock ptarmigan*	*Lagopus muta*	FJ376523
2D	A	P	T	I	V	A	**American woodcock**	*Scolopax minor*	JF969754
2D							**spotted sandpiper**	*Actitis macularius*	FJ376527
2E	A	T	T	I	I	A	***American crow***	*Corvus brachyrhynchos*	JQ814788
2E							**American goldfinch**	*Carduelis tristis*	FJ376484
2E							**American robin**	*Turdus migratorius*	FJ376485
2E							**bank swallow**	*Riparia riparia*	FJ376488
2E							**barn swallow**	*Hirundo rustica*	FJ376534
2E							**blue jay**	*Cyanocitta cristata*	FJ376493
2E							**cliff swallow**	*Petrochelidon pyrrhonota*	FJ376499
2E							**Eastern bluebird**	*Sialia sialis*	EU660870
2E							***hermit thrush***	*Catharus guttatus*	JQ814792
2E							**house finch**	*Carpodacus mexicanus*	FJ376510
2E							**house sparrow**	*Passer domesticus*	FJ376511
2E							**house wren**	*Troglodytes aedon*	FJ376512
2E							***Northern raven***	*Corvus corax*	JQ969022
2E							**red-eyed vireo**	*Vireo olivaceus*	FJ376519
2E							**tree swallow***	*Tachycineta bicolor*	FJ376530

sub-type	aa identity 256	257	297	324	337	380	name	scientific name	accession #
2E							***veery***	*Catharus fuscescens*	JQ814793
2E							**white-breasted nuthatch**	*Sitta carolinensis*	FJ376531
2F	A	T	T	I	V	A	black-footed albatross	*Phoebastria nigripes*	AB106109
2F							***brown thrasher***	*Toxostoma rufum*	JQ814790
2F							emu	*Dromaius novaehollandiae*	JF950300
2F							**mourning dove**	*Zenaida macroura*	FJ376515
2G	T	A	I	I	V	A	**bobwhite quail**	*Colinus virginianus*	FJ376494
2G							ring-necked pheasant*	*Phasianus colchicus*	EU660873
3A	A	A	T	V	V	A	great blue heron*	*Ardea herodias*	FJ376506
3A							Japanese quail*	*Coturnix japonica*	EU660871
3B	A	T	T	V	V	A	American kestrel*	*Falco sparverius*	EU660867
3B							arctic tern	*Sterna paradisaea*	HQ317441
3B							bald eagle*	*Haliaeetus leucocephalus*	FJ376486
3B							**barred owl**	*Strix varia*	FJ376489
3B							**belted kingfisher**	*Megaceryle alcyon*	FJ376490
3B							**common flicker**	*Colaptes auratus*	FJ376500
3B							**common loon**	*Gavia immer*	FJ376502
3B							common tern**	*Sterna hirundo*	AF192503
3B							***cooper's hawk***	*Accipiter cooperii*	JQ814791
3B							double-crested cormorant*	*Phalacrocorax auritus*	EU660869
3B							**downy woodpecker**	*Picoides pubescens*	FJ376504
3B							**Eastern kingbird**	*Tyrannus tyrannus*	FJ376505
3B							great cormorant	*Phalacrocorax carbo*	AB109545
3B							**great horned owl**	*Bubo virginianus*	FJ376507
3B							herring gull*	*Larus argentatus*	DQ371287
3B							ivory gull	*Pagophila eburnea*	FJ376540
3B							**killdeer**	*Charadrius vociferus*	FJ376514
3B							osprey*	*Pandion haliaetus*	FJ376517
3B							**red-tailed hawk**	*Buteo jamaicensis*	FJ376520
3B							ring-billed gull*	*Larus delawarensis*	FJ376522
3B							**sandhill crane**	*Grus canadensis*	FJ376535
3B							***Saw-whet owl***	*Aegolius acadicus*	JQ969021
3B							**screech owl**	*Megascops asio*	FJ376526
3B							***Sharp-shinned hawk***	*Accipiter striatus*	JQ969020
3B							thick-billed murre*	*Uria lomvia*	FJ376529
3B							**turkey vulture**	*Cathartes aura*	FJ376537
3C	T	T	T	V	V	A	brant goose	*Branta bernicla*	FJ376539
3C							Canada goose	*Branta canadensis*	FJ376496
3C							common eider*	*Somateria mollissima*	EU660868
3C							greater scaup	*Aythya marila*	FJ376508
3C							**mallard***	*Anas platyrhynchos*	EU660872
3C							wood duck*	*Aix sponsa*	EU660875
3C							**wood thrush**	*Hylocichla mustelina*	FJ376533

Table 2. EC_{50}, ReS_{50} and ReP_{50} values calculated from the concentration-response curves obtained after exposure of COS-7 cells transfected with avian AHR1 constructs to TCDD, PeCDF or PCB 126. Two separate studies were performed and in each study four replicate wells/DLC concentration were included. The four concentration response curves and resulting EC50 values were calculated based on the four replicate wells. The mean EC_{50} values were derived from the four concentration-response curves/study. ReS_{50s} and ReP_{50s} were calculated using the mean EC_{50} values from Study 1 and Study 2. ReS is defined as $EC_{50\ (compound\ A)}$ of chicken construct ÷ $EC_{50\ (compound\ A)}$ of construct *X*. ReP is defined as EC_{50} of TCDD determined in construct *X* ÷ EC_{50} of PeCDF or PCB 126 determined for construct *X*. The data for wild-type chicken, pheasant and quail AHR1 constructs exposed to TCDD and PeCDF are from Farmahin *et al.* (2012).

chemical	AHR1 construct	$EC_{50\ (nM)}$ Study 1	$EC_{50\ (nM)}$ Study 2	ReS_{50}	ReP_{50}
TCDD	Chicken	0.22	0.21	1.0	1.0
	1B	0.33	0.22	0.77	1.0
	1C	0.23	0.34	0.75	1.0
	Pheasant	1.7	1.4	0.14	1.0
	2A	2.3	2.0	0.10	1.0
	2B	3.4	3.6	0.061	1.0
	2C	3.3	3.8	0.061	1.0
	2D	2.8	3.2	0.072	1.0
	2E	2.5	4.0	0.066	1.0
	2F	1.6	2.0	0.12	1.0
	2G	2.4	2.1	0.10	1.0
	Quail	25	17	0.010	1.0
	3A	35	24	0.0073	1.0
	3B	15	18	0.013	1.0
	3C	19	17	0.012	1.0
PeCDF	Chicken	0.19	0.17	1.0	1.2
	1B	0.19	0.11	1.2	1.9
	1C	0.18	0.18	0.99	1.6
	Pheasant	0.33	0.42	0.48	4.1
	2A	0.43	0.29	0.50	6.0
	2B	0.65	0.39	0.35	6.8
	2C	0.72	0.74	0.25	4.8
	2D	0.89	0.76	0.22	3.6
	2E	0.55	0.71	0.29	5.2
	2F	0.48	1.1	0.23	2.3
	2G	0.58	0.44	0.35	4.4
	Quail	0.70	1.4	0.17	20
	3A	1.4	1.4	0.13	21
	3B	1.4	1.9	0.11	10
	3C	0.88	1.8	0.14	14
PCB 126	Chicken	2.9	3.0	1.0	0.072
	1B	1.4	2.6	1.5	0.14
	Pheasant	95	91	0.032	0.017
	2D	70	90	0.037	0.038
	2E	82	75	0.038	0.042
	Quail	288	410	0.0085	0.060
	3A	243	409	0.0091	0.090
	3B	620	347	0.006	0.035

886

887 **Table 3.** Predicted LD_{50s} (95% confidence intervals) of TCDD, PeCDF and
888 PCB 126 for the 3 main avian types calculated from the regression line shown in
889 Figure 3.
890
891

compound	*avian type*	*predicted LD_{50} (pmol/g egg)*	*predicted LD_{50} (ng/g egg)*
TCDD	Type 1	0.78 (0.40 - 1.4)	0.26 (0.13 - 0.45)
	Type 2	5.0 (2.2 - 10)	1.7 (0.73 - 3.4)
	Type 3	27 (11 - 71)	9 (3.8 - 24)
PeCDF	Type 1	0.56 (0.29 - 1.0)	0.19 (0.10 - 0.34)
	Type 2	1.5 (0.64 - 2.8)	0.51 (0.22 - 1.0)
	Type 3	2.9 (1.7 - 5.2)	1.0 (0.58 - 1.8)
PCB 126	Type 1	4.7 (2.6 - 8.7)	1.5 (0.86 - 2.8)
	Type 2	80 (33 - 203)	26 (11 – 66)
	Type 3	273 (90 - 911)	89 (29 - 297)

892
893
894
895
896
897
898
899
900
901
902

FIGURE LEGENDS

Figure 1. (A) Western blot analysis showing the expression of wild-type (chicken) and mutant (1B, 1C, 2A, 2B, 2C, 2D, 2E, 2F, 2G, 3A, 3B, 3C) avian AHR1 protein in COS-7 cells (upper panel). β-actin (lower panel) was used as an internal control. Non-transfected COS-7 cells (NT) were used as a negative control. Forty-eight hours after transfection, cell lysates were separated by SDS-PAGE, transferred to PVDF membranes and probed with anti-V5-HRP antibody for detection of AHR1 expression. The same blot was stripped and re-probed for β-actin using anti-β-actin-peroxidase antibody. (B) Representative example (study 1) of the concentration-dependent effects of TCDD, PeCDF and PCB 126 on aryl hydrocarbon response element (AHRE)-driven luciferase reporter gene activity in COS-7 cells transfected with either full-length AHR1 constructs of chicken, ring-necked pheasant, Japanese quail or mutant constructs of chicken AHR1 (designated 1B, 1C, etc; see Table 1 for details). Cells were exposed to DMSO or serial dilutions of TCDD, PeCDF or PCB 126 for 18-20h, luciferase ratios (the ratio of firefly luminescence units to *Renilla* luminescence units) were determined, and data were normalized to the maximal response. Individual data points represent the mean ratio derived from four individual wells/concentration and bars represent SE. Each curve represents the average of four curves. The dashed vertical lines within each panel indicate EC_{50} values for the wild-type AHR1 constructs (chicken, ring-necked pheasant and Japanese quail).

Figure 2. Comparison of EC_{50} values (mean ± SD), derived from four concentration-response curves/study, obtained in COS-7 cells transfected with either (a) full-length AHR1 constructs of chicken, ring-necked pheasant,

927 Japanese quail or (b) mutant AHR1 constructs and exposed to TCDD for 18-20
928 h. Letters indicate significant differences (one-way ANOVA) among AHR1
929 constructs (n=4; $p < 0.001$). Three general classes of avian species (chicken-like,
930 pheasant-like and quail-like) are indicated in white, light grey or dark grey,
931 respectively.

932 **Figure 3.** Correlation between LD_{50} data for TCDD, PeCDF and PCB 126
933 obtained from egg injection studies (sources for LD_{50} data are indicated in Table
934 S5) and EC_{50} data from the luciferase reporter gene (LRG) assay. Closed
935 symbols represent EC_{50} data for full-length (wild-type) AHR1 constructs for
936 chicken (C), ring-necked pheasant (P) and Japanese quail (Q). Open symbols
937 represent EC_{50} data for mutant constructs of chicken AHR1 (cormorant, tern,
938 kestrel, and bluebird). For example, the open symbol for bluebird represents the
939 EC_{50} for construct **chicken DBD_2E LBD_chicken TAD** (DBD, DNA-binding
940 domain; LBD, ligand-binding domain; TAD, transactivation domain). For pheasant
941 and quail, there are both closed and open symbols; open symbols represent the
942 chicken AHR1 mutant that includes the pheasant or quail LBD sequence. The
943 dotted lines represent the 95% confidence intervals.

944

945 **Figure 4.** **(A)** Alignments of chicken, ring-necked pheasant and Japanese
946 quail AHR1 LBD sequences with HIF-2α secondary structure templates 1P97A
947 and 2A24 that were obtained using the DALI server. *Sub-section 1*: The amino
948 acid sequence alignments of 1P97A (structure 17), 2A24 (structure 11), chicken,
949 pheasant and quail. *Sub-section 2:* The secondary structure assignments

obtained by use of the Define Secondary Structure of Proteins (DSSP) algorithm (H/h: helix, E/e: strand, L/l: coil). *Sub-section 3:* Illustration of the secondary structure labelled with conventional PAS domain structure nomenclature (Gong *et al.*, 1998). **(B)** Cartoon representations of chicken, ring-necked pheasant and Japanese quail structural models of the AHR1 LBD. The amino acids at positions 297, 324, 337 and 380 are indicated, and those at 324 and 380 point into the cavity, while amino acids 297 and 337 point away from the cavity. The volumes of the main cavities in the Connolly's molecular surface calculated by CASTp are indicated for the three avian species. **(C)** Identification of amino acids with side chains that contribute to the AHR1 ligand binding cavity. The amino acid residues that contribute to the internal cavities (highlighted in green) of chicken, pheasant, and quail AHR1 LBD models were identified by use of CASTp. Amino acids at position 297, 324, 337 and 380 are shown within black boxes.

Figure 5. Cartoon representations of structural models of chicken and three mutant (I324V, S380A and I324V_S380A) AHR1 LBDs. The location of the potential binding site in the chicken AHR1 LBD model was determined by use of the Molegro software. Volumes of the main cavities in the Connolly's molecular surface, calculated by CASTp, are indicated. The four key amino acid locations are shown; the amino acids at positions 324 and 380 point inward and those at positions 297 and 337 point outward.