



LJMU Research Online

Mellor, CL, Tollefsen, K-E, LaLone, C, Cronin, MTD and Firman, JW

In silico identification of chemicals capable of binding to the ecdysone receptor

<http://researchonline.ljmu.ac.uk/id/eprint/12792/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Mellor, CL, Tollefsen, K-E, LaLone, C, Cronin, MTD and Firman, JW (2020) In silico identification of chemicals capable of binding to the ecdysone receptor. Environmental Toxicology and Chemistry, 39 (7). pp. 1438-1450. ISSN 0730-7268

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

In Silico Identification of Chemicals Capable of Binding to the Ecdysone Receptor

Claire L. Mellor,^{a,b} Knut Erik Tollefsen,^c Carlie LaLone,^d Mark T.D. Cronin,^a and James W. Firman^{a,*}

^aSchool of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Liverpool, United Kingdom

^bSchool of Forensic and Applied Sciences, University of Central Lancashire, Preston, Lancashire, United Kingdom

^cNorwegian Institute for Water Research, Oslo, Norway

^dGreat Lakes Toxicology and Ecology Division, Center for Computational Toxicology and Exposure, Office of Research and Development, US Environmental Protection Agency, Duluth, Minnesota, USA

Abstract: The process of molting, known alternatively as ecdysis, is a feature integral in the life cycles of species across the arthropod phylum. Regulation occurs as a function of the interaction of ecdysteroid hormones with the arthropod nuclear ecdysone receptor—a process preceding the triggering of a series of downstream events constituting an endocrine signaling pathway highly conserved throughout environmentally prevalent insect, crustacean, and myriapod organisms. Inappropriate ecdysone receptor binding and activation forms the essential molecular initiating event within possible adverse outcome pathways relating abnormal molting to mortality in arthropods. Definition of the characteristics of chemicals liable to stimulate such activity has the potential to be of great utility in mitigation of hazards posed toward vulnerable species. Thus the aim of the present study was to develop a series of rule-sets, derived from the key structural and physicochemical features associated with identified ecdysone receptor ligands, enabling construction of Konstanz Information Miner (KNIME) workflows permitting the flagging of compounds predisposed to binding at the site. Data describing the activities of 555 distinct chemicals were recovered from a variety of assays across 10 insect species, allowing for formulation of KNIME screens for potential binding activity at the molecular initiating event and adverse outcome level of biological organization. *Environ Toxicol Chem* 2020;39:1438–1450. © 2020 The Authors. *Environmental Toxicology and Chemistry* published by Wiley Periodicals LLC on behalf of SETAC.

Keywords: Computational toxicology; Endocrine-disrupting compounds; Invertebrate toxicology; Ecdysone receptor; Arthropod

INTRODUCTION

Endocrine systems are present in the great majority of animal phyla, playing integral roles in the mediation of essential developmental, reproductive, neurological, and immune functions. Although great diversity in form and complexity is found across species, common defining features may be discerned, including the secretion of hormones by glands into the circulatory system and the action of these signaling molecules at distant, dedicated receptor sites. Because of the physiological importance of endocrine pathways, disruption of function can

have substantial consequence to organisms (Tyler et al. 1998). Exogenous chemicals that interfere with endocrine physiology, and induce adverse effects, are defined as endocrine disruptors (Colborn et al. 1993). Typically, chemical stressors modulate pathways through agonism or antagonism by acting on receptors. Identification of endocrine-disrupting substances remains an ongoing enterprise, with continued expansion of the number of associated compounds (which currently exceed 8000) as further toxicological data are gathered (Bergman et al. 2013; Birnbaum 2013).

A report published in 2002 through the International Program on Chemical Safety (IPCS) sought to assess the existing state of knowledge of endocrine disruption. Considering a varied array of species, the IPCS study highlighted deficits in the availability of appropriate supporting evidence concerning the status of specific endocrine disruptors (Darmstra et al. 2002). Our understanding has since continued to develop, strengthened by an increased focus on mechanistic aspects underpinning pathway

This article contains online-only Supplemental Data.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

* Address correspondence to j.w.firman@jmu.ac.uk

Published online 26 April 2020 in Wiley Online Library

(wileyonlinelibrary.com).

DOI: 10.1002/etc.4733

activation (Diamanti-Kandarakis et al. 2009; Kortenkamp et al. 2011; Skakkebaek et al. 2011). An area of investigation that requires greater definition is endocrine disruption across invertebrate phyla. Within the arthropod phylum, which consists of a range of environmentally ubiquitous species from classes including Insecta, Crustacea, and Arachnida, a thorough understanding of the endocrine-disrupting potential relative to prevalent pesticide compounds is lacking. (Colborn et al. 1993; Stanley and Preetha 2016).

The adverse outcome pathway (AOP) framework can be useful in linking a given stressor to an adverse outcome through consideration of molecular initiating events and intermediate key events. Ecdysteroid signaling is of importance to arthropod growth and development and is considered an endocrine pathway integral to exoskeleton shedding in molting and metamorphosis (Yamanaka et al. 2013). Interaction of ecdysteroid hormones, such as ecdysone (Figure 1A) and 20-hydroxyecdysone (Figure 1B), with the ecdysone receptor (EcR) constitutes the main molecular initiating event in any AOP for this effect. The EcR exists as a complex composed of 2 subunits, these being the EcR and ultraspiracle proteins (Gunamalai et al. 2004; Hill et al. 2013; Sumiya et al. 2014). In common with other nuclear receptors, the EcR possesses a ligand-binding domain within which is incorporated a hydrophobic cleft formed from an arrangement of 12 α -helices. Its dimensions are such that it may accommodate binders displaying structural deviation from the classical steroidal template (Billas et al. 2003; Evenseth et al. 2019).

The AOPs associating perturbation of ecdysone signaling in arthropods with impaired molting culminating in death have been described (for a general outline, see Figure 2; Song et al. 2017b; Song and Tollefsen 2018). A variety of endogenous and xenobiotic substances are known to act through the EcR, but definitions of the structural and physicochemical characteristics required for binding remain incomplete. Modeling of receptor binding sites has provided insight into interactions underlying docking affinity—although this knowledge has not been applied in predictive toxicology (Kasuya et al. 2003; Zotti et al. 2012; Evenseth et al. 2019). It is clearly neither practical nor desirable to assess indiscriminately the binding activity of vast chemical libraries through existing *in vivo* or *in vitro* means, and hence a rationale is provided for the development of *in silico* techniques allowing strategies for identification of compounds likely to act at the site.

The aim of the present study was to construct a computational screen for the ability of compounds to bind to the EcR. Literature sources were examined for data concerning the identities of chemicals verified experimentally as acting at the site. Although the availability of these data necessarily led us to focus on insect species, it is intended that scope for applicability within crustacean, myriapod, and arachnid classes be demonstrated through means of sequence homology comparison. Both structural features and computed physicochemical characteristics were analyzed with the aim of developing sets of structural alerts and physicochemical property ranges enabling identification of molecules displaying high affinity for the EcR. In line with AOP methodology, a distinction was made between compounds shown to induce downstream effects characteristic of pathway activation (the adverse outcome) and compounds definitively acting at the receptor (the molecular initiating event). Following compilation, these sequences were implemented into freely available workflows for use in the open-access data analysis software Konstanz Information Miner (KNIME).

MATERIALS AND METHODS

Collection and consideration of data

Data concerning the identity and activity of potential EcR agonists were retrieved through both a literature search and the open-access ChEMBL database (Gaulton et al. 2011; European Bioinformatics Institute 2020). *In vivo* and *in vitro* experimental outcomes were considered suitable for inclusion. For the purposes of model building, assay systems were classified according to the biological level of organization (adopting the AOP, i.e., the molecular initiating event–key event–adverse outcome framework) at which the impact of EcR agonism on the test organism was characterized. Those systems describing effects on life-cycle progression of intact organisms (EcR agonist–associated molting abnormality) were deemed representative of the adverse outcome, whereas those conversely relating either to direct receptor binding (EcR activation) or to localized influence on isolated tissues (tissue phenotype or biochemical alteration) were adopted to describe the molecular initiating event. Two distinct models were

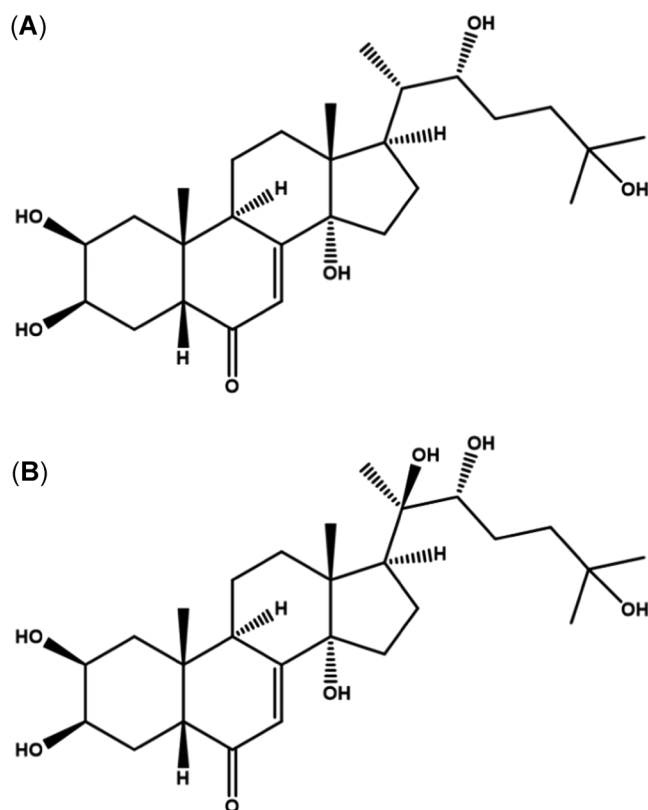


FIGURE 1: Chemical structures of the most prevalent ecdysteroids, ecdysone (A) and 20-hydroxyecdysone (B).

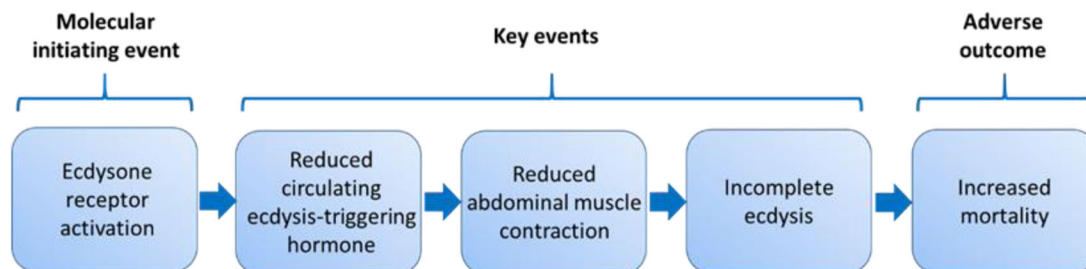


FIGURE 2: Schematic representation of the putative adverse outcome pathway (AOP) through which ecdysone receptor activation increases molting disruption and mortality in arthropods.

constructed from these data, one corresponding solely to the molecular initiating event level, the other solely to the adverse outcome. For a summary of relevant information concerning these assay systems, see Table 1.

Compounds exhibiting assay activity above defined thresholds were deemed “active”—indicating their greater likelihood of holding capacity to induce effects specifically mediated through EcR binding. Drawing on the distinction adopted by Dinan, these included all compounds possessing median effect concentration values less than 10^{-4} M (Dinan 2003). In studies employing alternative designs, such as that of Smith et al. (2003), an increase in activity relative to vehicle control was regarded as sufficient. Those falling outside these boundaries were, conversely, classified as “inactive.” It was judged that affinity was low enough that they might induce adverse effects through alternative pathways at the concentrations required for ecdysone

binding response, and hence may be considered as nonspecific binders to the receptor.

Under the “Targets” heading, a search for “ecdysone receptor” was performed in the ChEMBL₂₅ database (European Bioinformatics Institute 2020). Data describing activity at “single protein” targets were considered appropriate for inclusion.

Assessment of interspecies variability in ligand-binding domain structure

Homology in amino acid sequence corresponding to the ligand-binding domain of the EcR was determined between relevant species through use of the US Environmental Protection Agency (USEPA) Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool (Ver 4.0;

TABLE 1: Summary of study details and data distribution showing adverse outcome-level and molecular initiating event-level effects^a

	Species	Assay system	Tissue/life stage	Outcome	Entries		Source
					Act.	In.	
Adverse outcome	<i>Bombyx mori</i>	Whole organism	Larval phase	Molting abnormality	4	—	Rajathi et al. 2010
	<i>Calliphora erythrocephala</i>	Whole organism	Prepupal phase	Molting abnormality	32	29	Dinan 2003
	<i>Chilo suppressalis</i>	Whole organism	Larval phase	Molting abnormality	3	—	Dinan 2003
	<i>Musca domestica</i>	Whole organism	Larval phase	Molting abnormality	25	5	Dinan 2003
Molecular initiating event	<i>Aedes aegypti</i>	Transfected mammalian cell	3T3 cell line	EcR activation	34	2	Smith et al. 2003
	<i>Bombyx mori</i>	Arthropod cell	Bm5 cell line	EcR activation	32	1	Soin et al. 2010a
					70	12	Soin et al. 2010b
					19	—	Tice et al. 2003b
		Transfected mammalian cell	HEK-293 cell line	EcR activation	34	5	Tice et al. 2003a
	<i>Bovicola ovis</i>	Recombinant protein	Baculovirus vector	EcR activation	103	15	Birru et al. 2010
	<i>Calliphora erythrocephala</i>	Isolated tissue	Imaginal disc	Tissue form alteration	5	—	Dinan 2003
	<i>Chilo suppressalis</i>	Isolated tissue	Integument	Tissue form alteration	5	—	Dinan 2003
	<i>Chironomus tentans</i>	Arthropod cell	Epithelial cell line	EcR activation	3	—	Dinan 2003
	<i>Choristoneura fumiferana</i>	Transfected mammalian cell	HEK-293 cell line	EcR activation	18	—	Tice et al. 2003b
	<i>Drosophila melanogaster</i>	Arthropod cell	Bll cell line	EcR activation	145	10	Dinan 2003
			C18 + cell line	EcR activation	4	—	Dinan 2003
			Kc cell line	EcR activation	13	—	Dinan 2003
					15	4	Harada et al. 2009
				44	—	Soin et al. 2010b	
	Isolated tissue	Imaginal disc	Tissue form alteration	5	1	Dinan 2003	
<i>Lucilia cuprina</i>	Recombinant protein	Baculovirus vector	EcR activation	21	3	Birru et al. 2010	
<i>Spodoptera littoralis</i>	Arthropod cell	SL2 cell line	EcR activation	36	1	Soin et al. 2010a	

^aThe column “Entries” describes the quantity of individual compounds, both active (act.) and inactive (in.) present in each. EcR = ecdysone receptor.

LaLone et al. 2016). At evaluation level 1, sequences relating to the EcR in its entirety were sought (a list of the corresponding National Center for Biotechnology Information accession numbers is given in the Supplemental Data, Table S1). Once these sequences had been recovered, evaluation level 2 analysis was performed focused exclusively on the ligand-binding domain (NR_LBD_EcR). A comparison of level 2 homology determined the extent of variability in interspecies ligand receptibility, with values exceeding the susceptibility cut-offs defined by SeqAPASS considered representative of acceptable similarity (LaLone et al. 2016).

Identification and construction of key structural features and alerts

The molecular structure of each compound in the compiled binder list was visualized in the freely available Molecular Networks ChemoTyper software (Ver 1.0; Yang et al. 2015). Examination of the structures in the active set permitted the identification of key characteristics associated with receptor affinity. This process was informed through employment of chemical knowledge, combined with information gathered from existing literature sources. Features and alerts identified as being important for binding were coded into Simplified Molecular Input Line Entry System (SMILES) arbitrary target specification (SMARTS) strings, constructed so that they might enable retention of key spatial features, while simultaneously permitting degrees of deviation from known binder structure and excluding, as far as possible, those associated with inactivity and poor affinity in a procedure analogous to those previously reported (Steinmetz et al. 2015; Mellor et al. 2016). The strings compiled were subsequently applied in KNIME through use of the RDKit Substructure Filter node (Ver 2019.03.1; Landrum 2016). A distinction was drawn between those compounds possessing steroidal (ecdysteroids and derivatives) and nonsteroidal structure, with unique rule-sets in each workflow constructed for each.

Calculation of molecular physicochemical descriptors

Physicochemical properties for each binder were calculated in KNIME (Ver 3.5.3; Berthold et al. 2008) courtesy of the Chemistry Development Kit Molecular Properties node (Ver 1.5.12; Steinbeck et al. 2003). To facilitate processing, structures of compounds were encoded as SMILES strings (Anderson et al. 1987).

Integration of structural and physicochemical parameters in KNIME workflow

We incorporated the structural alert and physicochemical property-derived inclusion and exclusion criteria ascertained through methods described in the previous sections, *Identification and construction of key structural features and alerts* and *Calculation of molecular physicochemical descriptors*. Then workflows enabling prediction of potential EcR binding capacity across test compounds were constructed using the KNIME

software. In all, 2 distinct flows were constructed—one confined to compounds possessing molecular initiating event-level activity data and the other to compounds causing effects at the adverse outcome level. A sequence of filters, constituting a rule-set, was implemented through employment of appropriate nodes. Compounds were entered into the workflow as SMILES, and molecules matching SMARTS-encoded structural fragment markers (as well as those falling within the boundaries of the physicochemical property ranges) were classed as prospective EcR agonists. Those failing either count were deemed inactive.

RESULTS

Collection and profiling of data

Data availability and coverage. From sources detailed previously in the *Collection and consideration of data* section, a collection of 555 unique chemical entities with data relating activity either at, or through, the EcR was recovered. These data were present exclusively in insects (across a total of 10 species). Of these compounds, 470 were deemed active agonists (detailed in the Supplemental Data, Table S2): ecdysteroids and their derivatives accounted for 196 chemicals and nonsteroidal binders for the remaining 274. Classified according to experimental characterization, 434 compounds were identified as having molecular initiating event-level EcR binding potential, with 59 displaying adverse outcome effects. (A minority of 23 possessed data attesting to both characteristics.) The activities of 160 steroids were characterized by molecular initiating event-relevant assays, alongside all 274 of the nonsteroidal cohort. Adverse outcome-level data were present across 55 steroidal and 4 nonsteroidal binders.

Conversely, 85 compounds were categorized as inactive, 48 steroidal and 37 nonsteroidal (listed in the Supplemental Data, Table S3). Each of the latter was characterized solely at the molecular initiating event level. Of the steroidal set, 34 possessed only adverse outcome-applicable data and 14 had data only at the molecular initiating event level. No adverse outcome-level nonsteroidal inactives were identified.

Assessment of interspecies variability in ligand-binding domain structure

With data drawn from 10 distinct insect species, it was necessary to discern the extent to which the sequence homology of the relevant EcR binding domains remained conserved. It was furthermore desirable to assess how this similarity might extend to arthropods outside the hexapod class, including arachnids, crustaceans, and myriapods. Through use of the SeqAPASS tool (as described previously in *Assessment of interspecies variability in ligand-binding domain structure*), comparisons between the amino acid sequences present in the ligand-binding domains of the respective EcRs were performed. Data were available for 8 of 10 insect species, the exceptions being *Bovicola ovis* and *Musca domestica*. A predicted sequence present for *M. domestica* was adopted, whereas *B. ovis* was by necessity excluded. Representative arachnid (*Agelena silvatica*), crustacean (*Callinectes sapidus*), and myriapod (*Lithobius peregrinus*) species were included in addition.

TABLE 2: Interspecies sequence similarity of ecdysone receptor ligand binding domains (expressed as percentage), as determined through SeqAPASS^a

	A. aegypti	B. mori	C. erythrocephala^c	C. suppressalis	C. tentans	C. fumiferana	D. melanogaster	L. cuprina	M. domestica	S. littoralis	<i>A. silvatica</i>	<i>C. sapidus</i>	<i>L. peregrinus</i>
A. aegypti		72.77	88.95	75.45	79.08	72.77	86.98	88.95	88.48	74.35	58.01	56.51	54.06
B. mori	73.94		70.26	89.85	65.71	88.49	71.62	70.02	70.58	91.93	54.99	52.68	54.28
C. erythrocephala^b	89.23	69.6		71.89	77.04	70.86	96.6	99.6	98.81	71.97	57.72	55.74	55.42
C. suppressalis	74.75	90.15	70.99		66.77	92.26	70.84	70.99	71.62	88.27	55.28	53.32	53.63
C. tentans	81.62	67.99	79.71	70.28		67.91	78.84	79.71	78.84	69.97	54.99	51.66	52.06
C. fumiferana	72.48	88.91	70.36	92.77	65.25		71.07	70.36	70.68	86.56	54.24	51.81	53.14
D. melanogaster	87.39	71.05	96.75	71.85	76.76	71.69		96.35	97.62	71.37	57.49	54.88	54.88
L. cuprina	88.91	69.41	99.68	71.79	77.5	70.92	96.28		98.49	71.87	57.92	55.62	55.78
M. domestica^c	88.9	70.02	98.97	72.64	77.08	71.29	97.62	98.57		72.56	57.97	55.99	55.19
S. littoralis	75	90.84	72.37	89.89	67.83	87.66	71.66	72.21	72.85		56.61	55.57	55.49
<i>A. silvatica</i>	59.56	56.08	59.07	57.29	56.24	55.91	58.75	59.07	59.24	57.62		64.18	74.88
<i>C. sapidus</i>	56.44	53.18	55.46	53.83	51.3	53.02	54.57	55.46	55.63	55.87	64.27		65.25
<i>L. peregrinus</i>	57.51	57.01	58.77	57.6	54.41	56.76	58.1	58.77	58.44	58.52	77.58	67.42	

^aNoninsect species and their homology scores are in italics. Insect species and their homology scores are in bold. Coloration provides visual representation of extent of homology. For genus names, see Table 1.

^bAlternatively known as *Calliphora vicina*.

^cSequence predicted.

Table 2 shows that the general extent of homology in insects is high, far exceeding the susceptibility cut-off points determined using SeqAPASS (which had 16–27% homology; Supplemental Data, Table S1). Homology rates ranged from 65.25% (*Choristoneura fumiferana* vs *Chironomus tentans*) to 99.68% (*Lucilia cuprina* vs *Calliphora erythrocephala*). Thus it may be presumed with confidence that the binding domain of the receptor shares sufficient conservation across the sample of species to ensure that variability in receptiveness to ligands was minimal throughout. As expected, a reduced degree of sequence similarity was evident between insects and noninsects. Nevertheless, the extent of homology—encompassing values from 59.56% (*A. silvatica* vs *Aedes aegypti*) to 51.3% (*C. sapidus* vs *C. tentans*)—remained far above the calculated thresholds. Complete SeqAPASS reports, incorporating data related to all relevant species, may be found in the Supplemental Data, Table S4.

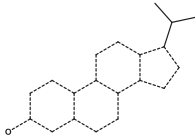
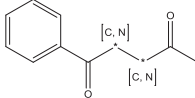
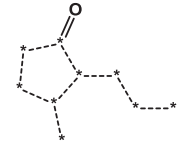

Construction of workflows

Derivation of structural alerts and accompanying physicochemical property ranges. Structural alerts, formulated and coded according to the protocols described previously in *Identification and construction of key structural features and alerts*, are described below and depicted in Table 3. Figure 3 shows representatives of the major classes of active binder recovered, from which the basis of the rule-sets was derived. Because of evidence indicating a potential for occupation of overlapping, yet distinct, sites on the receptor-binding domain, unique rules were developed accounting for both steroidal and nonsteroidal subtypes (Billas et al. 2003; Evenseth et al. 2019). Alerts were accompanied, where appropriate, by relevant physicochemical property ranges (calculated in KNIME according to protocols described previously in the section *Calculation of molecular physicochemical descriptors*), further refining the chemical space. These are displayed in Table 4.

Types of rule-sets: Steroidal. Compounds holding the characteristic tetracyclic core, complete with the presence of the C3 oxygen moiety and C17 branched carbon side chain (represented by turkesterone; Figure 3A), meet the primary criterion for inclusion in the steroidal rule-sets. These molecules embed within an L-shaped binding cleft, with the conservation of the tetracyclic hydrocarbon structure essential in the mediation of hydrophobic interactions anchoring molecules in a favorable position (Billas and Moras 2005). Omission of the C17 chain, positioned deepest within the pocket, eliminates the various van der Waals interactions that further assist in ligand–receptor complex stabilization. The requirement that the side chain occupy sufficient volume to facilitate these interactions is represented through incorporation of the atomic and bond contribution of van der Waals volume (VABC) volume descriptor—a minimum threshold of which must be met for “active” status to be returned. This scaffold further functions as a frame supporting the numerous polar groups engaging in hydrogen bonding with flanking amino acid residues—among which are the C2, C3, and C20 hydroxyls and the C6 and C14 carbonyl moieties (Billas and Moras 2005; Evenseth et al. 2019). Although it is the C3 hydroxyl that is apparently most essential in distinguishing active binders (hence its integration into the structural alert), the distribution of the further hydrogen-bond participants about the molecule is more flexible. Thus the requirement for these groups is represented through inclusion of a physicochemical parameter specifying the necessity for the sum of hydrogen-bond donors and acceptors as greater than, or equal to, 4.

Types of rule-sets: Nonsteroidal. Nonsteroidal entries must match (within the molecular initiating event pathway) representations of one of 3 defining scaffolds: either diacylhydrazine (e.g., tebufenozide; Figure 3B), methylene- γ -lactam (represented by Figure 3C), or substituted tetrahydroquinoline (represented by

TABLE 3: Visual representation of essential structural features, accompanied by SMILES arbitrary target specification (SMARTS) strings, as implemented within Konstanz Information Miner (KNIME) workflows^a

	Fragment	SMARTS string	Structural feature	MIE	AO
Steroidal	Tetracyclic steroid core	<chem>[#6]~1~[#6]([#6;R0]([#6])[#6])~[#6]~2~[#6]~[#6]~[#6]~3~[#6](~[#6]~[#6]~[#6]~4~[#6]~[#6](~O)~[#6]~[#6]~[#6]~3~4)~[#6]~2~[#6]~1</chem>		YES	YES
Nonsteroidal	Diacylhydrazine	<chem>c1cccc1~[#6;!R](~O)~[#6,#7]~[#6,#7]~[#6;!R](~O)~[#6]</chem>		YES	YES
	Methylene-γ-lactam	<chem>*([C;!R])~1~*~*~*(=O)~[*;R1]~1~*~*~*[*;R]</chem>		YES	NO
	Tetrahydro-quinoline	<chem>*~1~*~*(~*[*;R])~*~2~*~*~*~*~*~2~*~1~*[*;!R][*;R]</chem>		YES	NO

^aYES/NO means the presence or not of alerts within molecular initiating event (MIE) and adverse outcome (AO) screens.

Figure 3D). The binding site occupied by these molecules, although sharing a region of overlap with that of the steroids, is nevertheless distinct—possessing a pronounced V shape (Billas and Moras 2005; Evenseth et al. 2019). General to all classes is an upper molecular weight limit of 500, reflecting the compact, drug-like form of the studied compounds. The diacylhydrazine alert stipulates the inclusion of an acyl benzene unit—the ring occupying a position between adjacent methionine residues and the carbonyl acting as a hydrogen-bond acceptor (Billas and Moras 2005). A second carbonyl, serving a similar function, is separated from this courtesy of a linking unit consisting of 2 atoms limited to combinations of carbon or nitrogen, the latter of which is generally present in the form of a secondary amine free to participate in hydrogen-bond donation. A minimum degree of steric bulk, in the form, for example, of substitution at the aryl moiety, appears necessary to ensure that essential hydrophobic interactions are maintained. Thus a lower molecular weight bound of 290 is present solely in this group.

The methylene-γ-lactam unit is rendered as a 5-membered ring (atom identity unspecified) substituted with 3 essential units: the carbonyl, the methylene, and also a secondary cyclic unit. Although the former is unambiguously a hydrogen-bond acceptor, the latter 2 are implicated in hydrophobic interactions (Dinan et al. 2012). Because evidence suggests that the structure of the additional ring may vary with respect to size and composition, the alert is coded with generality in mind (Birru et al. 2010). A flexible unit linking the 2 cyclic fragments is common to all in this category—a factor represented by the requirement for the presence of at least one rotatable bond. Comparatively little is known

about the binding mode of the less characterized tetrahydroquinoline class (Dinan et al. 2012; Giaccoppo et al. 2017). Accordingly, it proved necessary to represent the structure as merely its core component—the 1,4-substituted, 10-membered fused ring, with atom identity and bond nature undefined. Excessive substituent bulk appears to be detrimental to activity, and thus a maximum VABC threshold is stipulated.

Integration into a sequential KNIME workflow. A combination of structural and physicochemical rule-sets was achieved within the framework of KNIME workflows, which are openly available for download (figshare 2020). Figure 4 outlines the general structure of the workflows, both for the molecular initiating event level and for the adverse outcome level, charting the passage of compounds (encoded as SMILES strings) through appropriate selection filters to final classification as either active or inactive.

Structures should be uploaded in the form of a .csv file composed of 3 columns, each with a header, one containing SMILES strings (headered “SMILES”), one a numerical identifier (“ID”), and one a compound name (“Name”). A sample entry file may be found, for illustrative and testing purposes, in the Supplemental Data, Table S5.

Assessment and validation of workflow performance

Internal validation. To assess the capacity of the workflows to successfully predict the status of known active and inactive

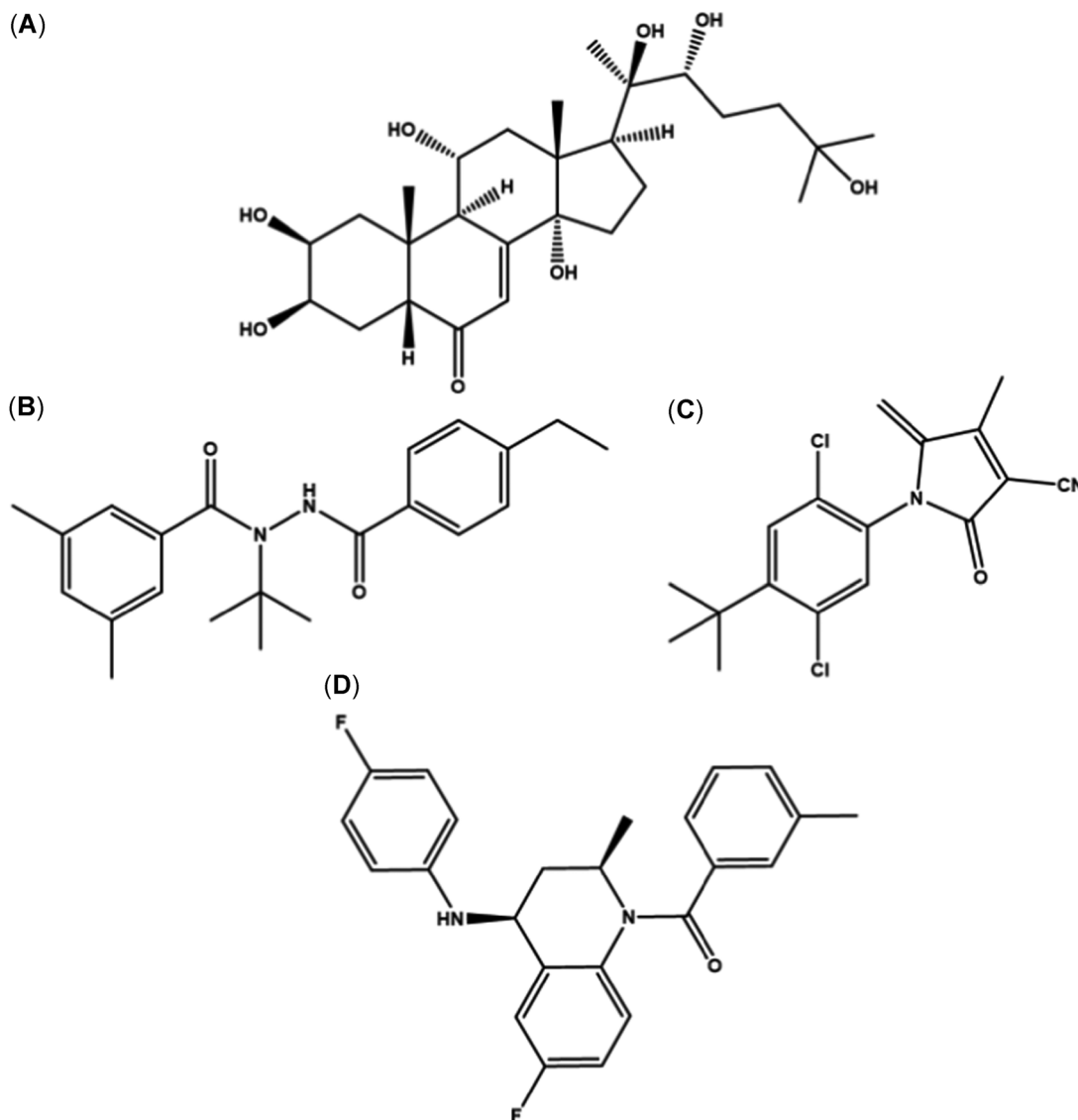


FIGURE 3: Structures of compounds representing major binder classes steroidal (A), diacylhydrazine (B), methylene- γ -lactam (C), and tetrahydroquinoline (D).

compounds, screening was performed on the complete inventory of 555 sourced chemicals. Outcomes are shown in Table 5, with performance metrics determined for both molecular initiating event- and adverse outcome-level rule-sets. The capacity to appropriately predict compounds as active binders

was uniformly high, with a minimum success rate of 94.5% (representing 53 correctly assigned from the 55 adverse outcome steroidal entries). Each excluded steroid lacked the stipulated 3-OH unit (Supplemental Data, Table S2: IDs 45, 47, 80, 95, and 195) whereas the single unmatched nonsteroidal compound appeared as a methylene- γ -lactam with an uncommon fused ring unit (ID 320).

Because of a general structural similarity between the active and nonactive compounds, a definitive exclusion of the latter was challenging. Across both molecular initiating event and adverse outcome cohorts, between 21.6% of the 44.1% of those compounds displaying inactivity at the receptor were, through screening, correctly excluded. Among steroidal entries, numerous inactives were correctly identified on account of their generally smaller VABC and reduced quantity of hydrogen-bonding participants. With regard to nonsteroidals, both tetrahydroquinoline and diacylhydrazine

TABLE 4: Physicochemical property ranges accompanying essential structural features included within workflows

Screen	Fragment	Descriptor	Rule
Steroidal		H-bonding participants	≥ 4
		VABC	≥ 436
Nonsteroidal	All classes	Molecular weight	≤ 500
	Diacylhydrazine	Molecular weight	≥ 290
	Methylene- γ -lactam	Rotatable bonds	≥ 1
	Tetrahydroquinoline	VABC	≤ 410

VABC = atomic and bond contribution of van der Waals volume.

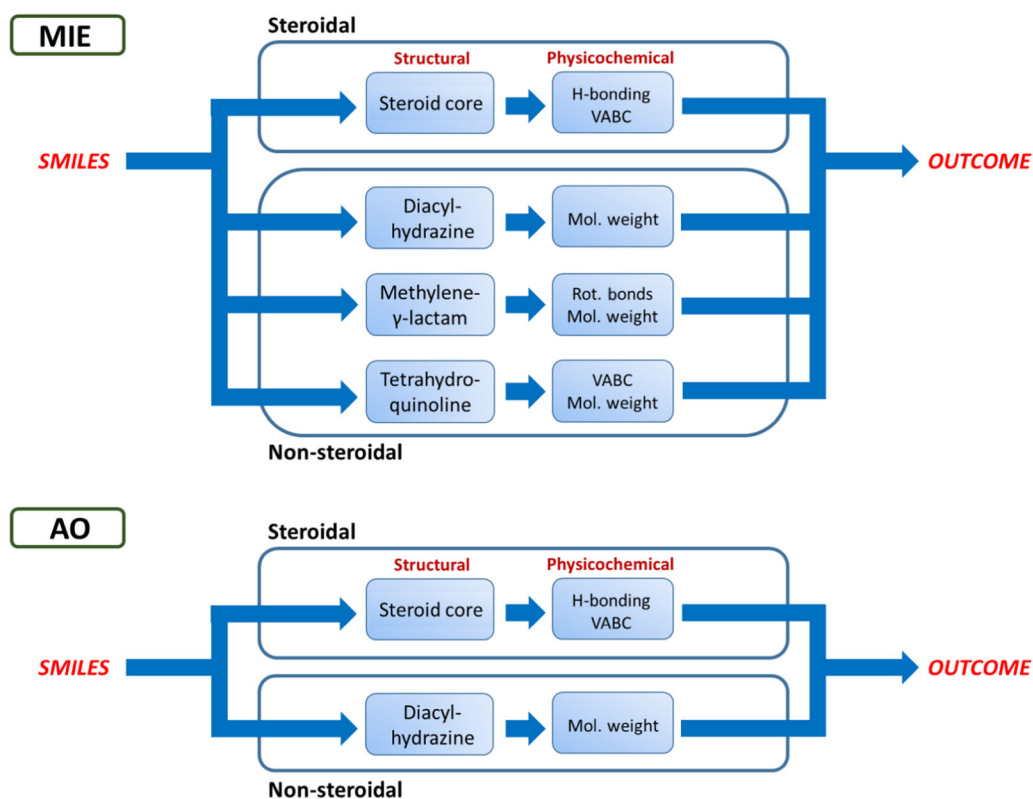


FIGURE 4: Integration of structural and physicochemical rule-sets in construction of molecular initiating event (MIE)-level and adverse outcome-level Konstanz Information Miner (KNIME) workflows. VABC = atomic and bond contribution of van der Waals volume; AO = adverse outcome.

classes were more reliably differentiated than were the methylene- γ -lactams.

Screening of external compound inventory. To investigate the performance of the rule-sets against a more general selection of chemicals, an inventory was screened of 8795 compounds sourced through the USEPA ToxCast initiative (incorporating a variety of pharmaceuticals, food additives, cosmetic ingredients, and synthetic precursors: listed in the Supplemental Data, Table S6; Richard et al. 2016). A list of matches is shown in the Supplemental Data, Table S7. Through the molecular initiating event-level screen, a collection of 34 potentially active binders was identified, 29 nonsteroidal in structure and 5 steroidal. Of this number, 16 were captured by adverse outcome rules.

Aside from a selection of pesticides—including the diacylhydrazines halofenozide, methoxyfenozide, and tebufenozide as present in the training set—the predicted binders are dominated by pharmaceuticals. Notably represented are several members belonging to the pyrazolone class of nonsteroidal anti-inflammatory drugs, including aminopyrine (Figure 5A), phenazone, and propyphenazone, each of which shares structural similarity with the methylene- γ -lactam core. This motif is further present as a fragment within the larger compounds including doxapam, el-trombopag, and a small number of azo dyes. Alongside the aforementioned pesticides, the diacylhydrazine alert was matched with members of the amphenicol antibiotic family, chloramphenicol (Figure 5B), florfenicol, and thiamphenicol. Vatalanib (Figure 5C), an inhibitor of the vascular endothelial

TABLE 5: Outcomes of internal validation exercise, describing the performance of both molecular initiating event (MIE) and adverse outcome (AO) workflows against appropriate training sets^a

Screen		Active			Inactive		
		Matched	Total	% Matched	Matched	Total	% Matched
MIE	Steroidal	158	160	98.8	10	14	71.4
	Nonsteroidal	273	274	99.6	29	37	78.4
AO	Steroidal	52	55	94.5	19	34	55.9
	Nonsteroidal	4	4	100.0	—	—	—

^aInclusion rates for known active and inactive compounds.

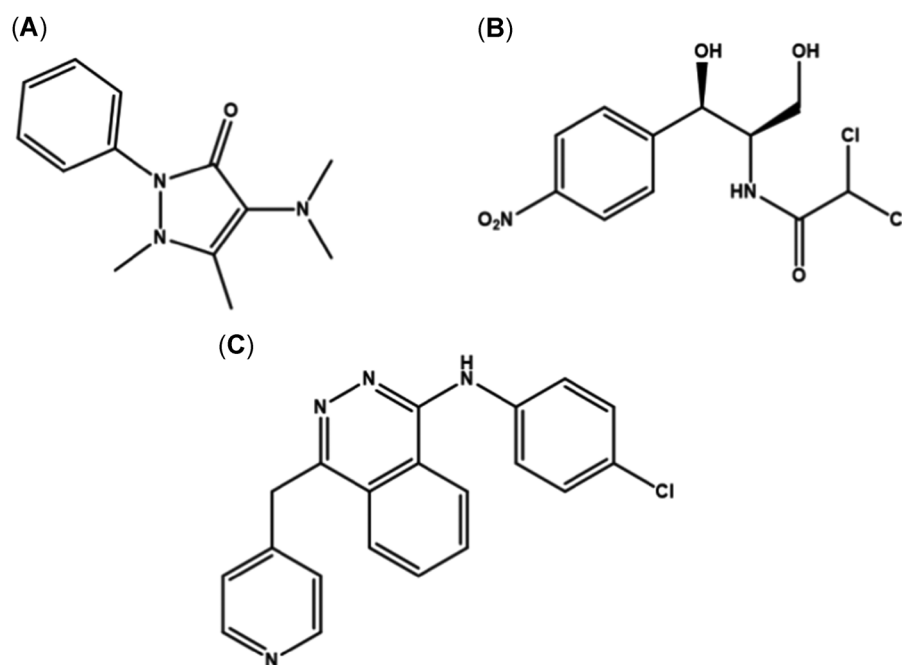


FIGURE 5: Structures of aminopyrine (A), chloramphenicol (B), and vatalanib (C).

growth factor receptor, was the sole recovered bearer of the tetrahydroquinoline moiety.

DISCUSSION

The present study accumulated existing data concerning the identity of known EcR binders, allowing delineation of shared structural and physicochemical characteristics. The data were adapted into a series of rules for implementation into a KNIME workflow, enabling detection of compounds likely to interfere with ecdysone signaling. It is intended that AOP-anchored, *in silico* techniques will, in time, support strategies for an integrated approach on testing and assessment (Tollefsen et al. 2014). Examination of literature sources produced an inventory composed of 555 distinct compounds (across 10 insect species) for which a capacity to act at the EcR had been experimentally assessed. Drawing from assay activity data, it was judged that a great majority of these compounds—470 in all—displayed potencies sufficient to indicate specific activity at the receptor, whereas only a comparatively small minority (the remaining 85) did not. Subdivision of these sets, both by the nature of available data with respect to adverse outcome pathway level and by key structural motif, led to the creation of the characteristic strands: molecular initiating event and adverse outcome, steroidal and nonsteroidal. The latter was necessitated by the apparent part-distinction of the site on the EcR ligand-binding domain occupied by each respective class (Billas et al. 2003).

Our approach was by necessity dictated through data availability, with a focus placed on one particularly prevalent class among the arthropod phylum—insects. Expansion of the taxonomic breadth of applicability to encompass other major

species classes would of course be beneficial to the utility of the model as a tool in an environmental risk assessment setting. Analysis by SeqAPASS indicated high conservation of ligand-binding domain sequence homology across the 10 insect species (6 Dipteran and 4 Lepidopteran) from which the data were drawn (LaLone et al. 2016). Importantly, this extended (with a minor reduction) to representative species sourced from alternative subphyla, such as crustaceans, myriapods, and arachnids. Such findings stand in support of evidence drawn from alternative studies, which suggest that EcR sequence homology conforms to a respectable degree across even substantially divergent species, albeit with minor amino acid variation at key positions holding association with variable sensitivity toward specific nonsteroidal pesticides (Nakagawa and Henrich 2009; Song et al. 2017b; Evenseth et al. 2019). Although experimental evidence from noninsect classes is scarce, a limited number of studies have reported the development of systems that may allow future widespread screening of EcR binding capacity in selected crustacean species, and hence providing routes through which these hypotheses might be tested (Yokota et al. 2011; De Wilde et al. 2013; Asada et al. 2014; Chan et al. 2019).

An example of an instance in which extension of domain may prove particularly useful is the impact of ecdysone signaling disruption on life-cycle progression in the crustacean *Daphnia magna*. Reproductive toxicity in this species, which forms the subject of Organisation for Economic Co-operation and Development (2012) test guideline 211, may be mediated through perturbations in the process of ecdysis (Rodriguez et al. 2007). Furthermore, EcR antagonism has been implicated in ecdysis-associated mortality, as described in an arthropod-specific AOP “Ecdysone receptor agonism leading to incomplete ecdysis associated mortality,” which has recently

been interrogated in detail in various *in silico*, *in vitro*, and *in vivo* assays (Fay et al. 2017; Song et al. 2017a, 2017b; Song and Tollefsen 2018).

It should be noted that the breadth of structural and physicochemical space present in the training set will of course dictate the specificity or generality of the alerts derived from it. The studies from which these data were drawn tended to display a focus on examining the influence on activity of small deviations about a central framework, be it the steroid, diacylhydrazine, methylene- γ -lactam, or tetrahydroquinoline cores. Thus the scope of domain is necessarily limited to derivations of each of these units, with extrapolation beyond informed by interpretation of key ligand-binding site interactions. Internal validation revealed that 68% of experimental nonactives were predicted to be active. Although this might appear somewhat conservative, it is important to reiterate that a distinction between features definitively separating apparent active and inactive compounds was not necessarily readily apparent, given the extent of the structural similarity

evident between them. Through detection of probable binders, the workflows may find utility in the guidance of future programs focused on experimental testing of chemicals liable to function as endocrine disruptors. After screening of a representative nonselective external compound inventory, it was revealed that approximately 0.4% of the 8795 entries were flagged as potentially active at the molecular initiating event level. With more extensive collections, such a proportion could represent identification of hundreds of chemicals in general use.

With regard to ligand structure, general diversity among the active steroids was limited. Conservation of the C20, C21, and C22 branch proved an absolute requirement, with compounds lacking such a motif, including the acknowledged ecdysteroid metabolites rubrosterone (Figure 6A) and poststerone (Figure 6B), notable for their inactivity (Dinan 2003). It is, therefore, a feature of this screen that vertebrate steroid families—such as sex hormones (e.g., estradiol; Figure 6C), corticosteroids (e.g., corticosterone; Figure 6D), and the

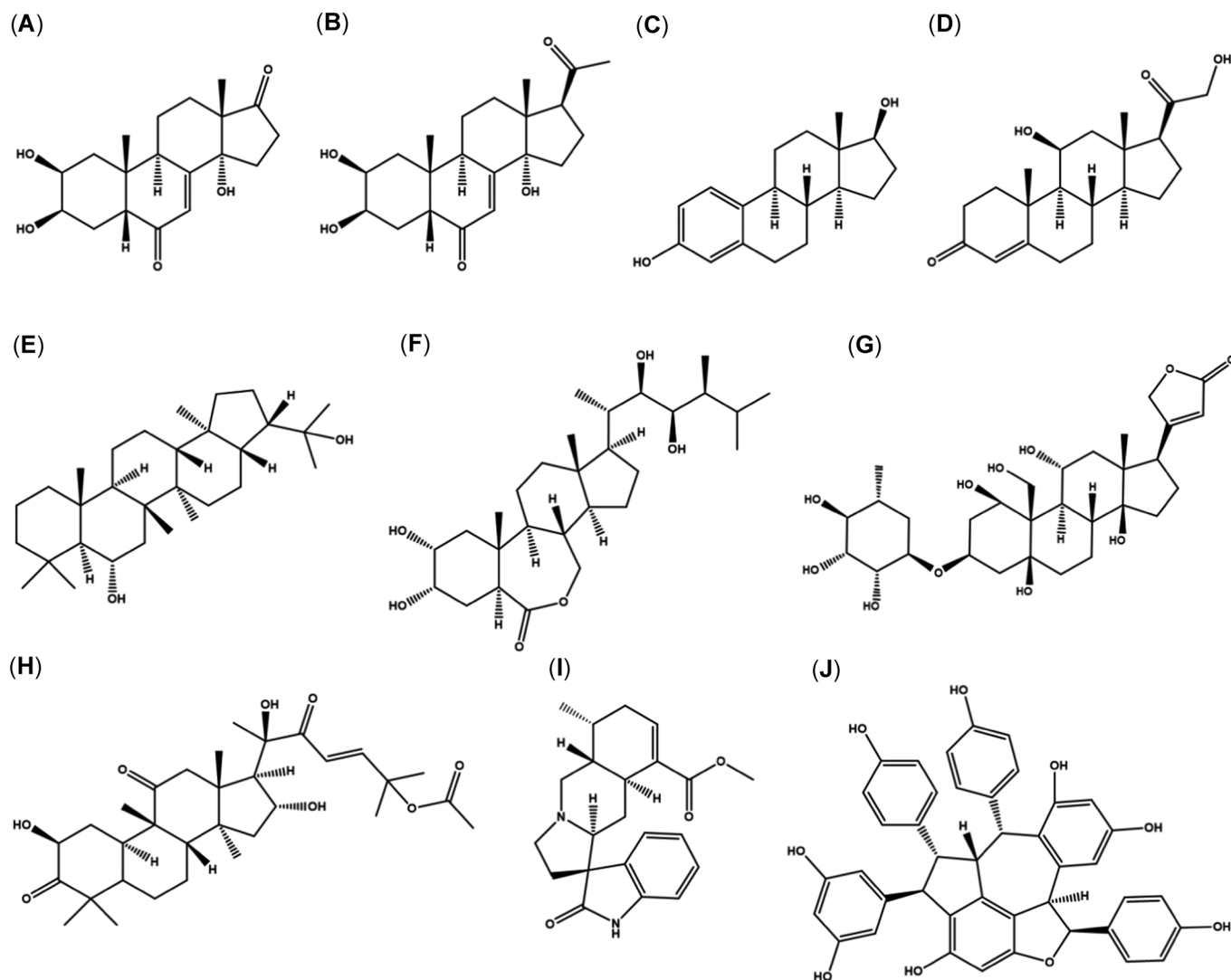


FIGURE 6: Structures of rubrosterone (A), poststerone (B), and compounds representing predicted classes of nonbinding natural products including estradiol (C), corticosterone (D), zeorin (E), brassinolide (F), and ouabain (G), alongside a selection of *in vitro* ecdysone receptor antagonists, cucurbitacin B (H), isomitraphylline (I), and suffruticosol A (J).

synthetic derivatives of each—would be predicted to be nonactive. Further excluded are the prokaryotic hopanoids (e.g., zeorin; Figure 6E), alongside the phytosteroidal brassinolide (Figure 6F) and cardiac glycoside (e.g., ouabain; Figure 6G) products. Accordingly, no data could be recovered attesting to the ability of any of these classes to bind the EcR. Many steroidal ecdysteroids described in the literature have their origins in selected plant species, in which they are speculated to play a role in defence against herbivore grazing (Dinan 2001).

Whereas the diacylhydrazine backbone is held in common, the methylene- γ -lactam and tetrahydroquinoline motifs are absent entirely from the adverse outcome–level rule-set. Such inequalities can be accounted for through consideration of the aims of the original studies, with the great majority of nonsteroidal compounds derived from research directed at the construction of novel EcR agonists through rational design. The primary interest of these studies was to assess direct binding affinity (i.e., the molecular initiating event), as opposed to assessing downstream effects that may arise within the organism (i.e., the key events and adverse outcomes). It has been established that the diacylhydrazines, including tebufenozide and methoxyfenozide, commercial insecticides, function through this route (Carlson et al. 2001). The activities of tetrahydroquinolines and methylene- γ -lactams have been demonstrated in various in vitro assays and, despite displaying promising performance at this level, further development toward their commercial adoption has yet to be pursued.

Although less intensively examined, evidence suggesting EcR antagonist capacity has emerged in a limited number of compounds. Twenty-six active chemicals were recovered from the literature—ultimately too few for derivation of reliable alerts given the structural diversity present (Supplemental Data, Table S8; Dinan et al. 1996, 1997, 2001a, 2001b). Prominent in this set were phytosteroids of the cucurbitacin and withanolide families (represented by cucurbitacin B; Figure 6H). Similarities between these compounds and the endogenous ecdysteroids are readily apparent, the primary variation being in side chain substitution. A trio of mitraphylline-like oxindole alkaloids was further present (isomitraphylline; Figure 6I), as were 2 benzodioxole derivatives and 6 stilbenoids (represented by suffruticosol A; Figure 6J).

CONCLUSIONS

Key structural and physicochemical characteristics of more than 500 experimentally determined compounds have been discerned, to form rule-sets expressed as workflows within KNIME, enabling identification of compounds likely to exert activity through the EcR within arthropods, and within insects in particular. Homology analysis by SeqAPASS revealed high levels of sequence conservation in the EcR ligand binding domains present in these species, indicating that reliable extrapolation is possible. A consideration of the assay systems employed in generation of the data allowed distinctions between molecular initiating event–level (receptor activation) and

adverse outcome–level (gross organism) effect. An ability to correctly identify binders was demonstrated at both levels, while screening of an external chemical inventory allowed further identification of compounds exhibiting a potential to influence signaling through the receptor. Combined, these data incorporate the bulk of publicly accessible knowledge concerning structure–activity relationships at the EcR. It is intended that their utility in rapidly and accurately identifying potential endocrine-disrupting substances might reduce uncertainty in environmental risk assessment.

Supplemental Data—The Supplemental Data are available on the Wiley Online Library at <https://doi.org/10.1002/etc.4733>.

Acknowledgment—The authors thank the Norwegian Institute for Water Research's Computational Toxicology Program (www.niva.no/nctp) and the Norwegian Research Council for funding (project NFR-221455 “Adverse Outcome Pathways for Endocrine Disruption in *Daphnia magna*, a conceptual approach for mechanistically-based Risk assessment (EDRISK).” Further information may be found at www.niva.no/EDRISK.

Author Contribution Statement—K.E. Tollefsen conceived and designed experiments. C.L. Mellor, J.W. Firman, and C. LaLone performed experiments. J.W. Firman constructed workflows. J.W. Firman wrote the manuscript. K.E. Tollefsen, M.T.D. Cronin, and C. LaLone provided editorial and technical assistance.

Data Availability Statement—Data pertaining to this manuscript (KNIME workflows) are deposited in figshare at https://figshare.com/articles/EDRisk_MIE_knwf/11365835.

REFERENCES

- Anderson E, Veith GD, Weininger D. 1987. SMILES: A line notation and computerized interpreter for chemical structures. EPA/600/M-87/021. Technical Report. US Environmental Protection Agency, Washington, DC.
- Asada M, Kato Y, Matsuura T, Watanabe H. 2014. Visualization of ecdysteroid activity using a reporter gene in the crustacean, *Daphnia*. *Mar Environ Res* 93:118–122.
- Bergman Å, Heindel JJ, Jobling S, Kidd KA, Zoeller RT. 2013. State of the science of endocrine disrupting chemicals—2012. Final Report. World Health Organisation, Geneva, Switzerland.
- Berthold MR, Cebon N, Dill F, Gabriel TR, Kotter T, Meinl T, Ohl P, Sieb C, Thiel K, Wiswedel B. 2008. KNIME: The Konstanz Information Miner. [cited 2019 November 11]. Available from: www.knime.com
- Billas IM, Moras D. 2005. Ligand-binding pocket of the ecdysone receptor. *Vitam Horm* 73:101–129.
- Billas IML, Iwema T, Garnier JM, Mitschler A, Rochel N, Moras D. 2003. Structural adaptability in the ligand-binding pocket of the ecdysone hormone receptor. *Nature* 426:91–96.
- Birnbaum LS. 2013. State of the science of endocrine disruptors. *Environ Health Perspect* 121:A107–A107.
- Birru W, Fernley RT, Graham LD, Grusovin J, Hill RJ, Hofmann A, Howell L, James PJ, Jarvis KE, Johnson WM, Jones DA, Leitner C, Liepa AJ, Lovrecz GO, Lu L, Nearn RH, O'Driscoll BJ, Phan T, Pollard M, Turner KA, Winkler DA. 2010. Synthesis, binding and bioactivity of γ -methylene γ -lactam ecdysone receptor ligands: Advantages of QSAR models for flexible receptors. *Bioorg Med Chem* 18:5647–5660.

- Carlson GR, Dhadialla TS, Hunter R, Jansson RK, Jany CS, Lidert Z, Slaweck RA. 2001. The chemical and biological properties of methoxyfenozide, a new insecticidal ecdysteroid agonist. *Pest Manag Sci* 57:115–119.
- Chan YH, Chu KH, Chan KM. 2019. Ecdysteroid-mimicking compounds act as both agonists and antagonists to the crustacean ecdysone receptor. *Chemosphere* 237:124551.
- Colborn T, Saal FSV, Soto AM. 1993. Developmental effects of endocrine-disrupting chemicals in wildlife and humans. *Environ Health Perspect* 101:378–384.
- Darmstra T, Barlow S, Bergman Å, Kavlock R, Van Der Kraak G. 2002. Global assessment of the state-of-the-science of endocrine disruptors. Final Report. World Health Organization, Geneva, Switzerland.
- De Wilde R, Swevers L, Soin T, Christiaens O, Rouge P, Cooreman K, Janssen CR, Smaghe G. 2013. Cloning and functional analysis of the ecdysteroid receptor complex in the opossum shrimp *Neomysis integer* (Leach, 1814). *Aquat Toxicol* 130–131:31–40.
- Diamanti-Kandarakis E, Bourguignon JP, Giudice LC, Hauser R, Prins GS, Soto AM, Zoeller RT, Gore AC. 2009. Endocrine-disrupting chemicals: An Endocrine Society scientific statement. *Endocr Rev* 30:293–342.
- Dinan L. 2001. Phytoecdysteroids: Biological aspects. *Phytochemistry* 57:325–339.
- Dinan L. 2003. Ecdysteroid structure-activity relationships. In Rahman A, ed, *Studies in Natural Products Chemistry*. Elsevier, Amsterdam, The Netherlands, pp 3–71.
- Dinan L, Bourne P, Whiting P, Dhadialla TS, Hutchinson TH. 2001a. Screening of environmental contaminants for ecdysteroid agonist and antagonist activity using the *Drosophila melanogaster* B₁₁ cell in vitro assay. *Environ Toxicol Chem* 20:2038–2046.
- Dinan L, Bourne PC, Meng Y, Sarker SD, Tolentino RB, Whiting P. 2001b. Assessment of natural products in the *Drosophila melanogaster* B₁₁ cell bioassay for ecdysteroid agonist and antagonist activities. *Cell Mol Life Sci* 58:321–342.
- Dinan L, Nakagawa Y, Hormann RE. 2012. Structure–activity relationships of ecdysteroids and non-steroidal ecdysone agonists. In Dhadialla TS, ed, *Advances in Insect Physiology*. Academic, Cambridge, MA, USA, pp 251–298.
- Dinan L, Whiting P, Alfonso D, Kapetanidis I. 1996. Certain withanolides from *Lochroma gesnerioides* antagonize ecdysteroid action in a *Drosophila melanogaster* cell line. *Entomol Exp Appl* 80:415–420.
- Dinan L, Whiting P, Girault JP, Lafont R, Dhadialla TS, Cress DE, Mugat B, Antoniewski C, Lepesant JA. 1997. Cucurbitacins are insect steroid hormone antagonists acting at the ecdysteroid receptor. *Biochem J* 327:643–650.
- European Bioinformatics Institute. 2020. ChEMBL Database. Hinxton, UK.
- Evenseth LM, Kristiansen K, Song Y, Tollefsen KE, Sylte I. 2019. In silico site-directed mutagenesis of the *Daphnia magna* ecdysone receptor identifies critical amino acids for species-specific and inter-species differences in agonist binding. *Comput Toxicol* 12:100091.
- Fay KA, Villeneuve DL, LaLone CA, Song Y, Tollefsen KE, Ankley GT. 2017. Practical approaches to adverse outcome pathway development and weight-of-evidence evaluation as illustrated by ecotoxicological case studies. *Environ Toxicol Chem* 36:1429–1449.
- figshare. 2020. KNIME workflows enabling identification of compounds with potential for binding at ecdysone receptor. [cited 2019 December 2]. Available from: https://figshare.com/articles/EDRisk_MIE_knwf/11365835
- Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP. 2011. ChEMBL: A large-scale bioactivity database for drug discovery. [cited 2019 November 11]. Available from: <http://www.rdkit.org/>
- Giacoppo JdOS, Carregal JB, Junior MC, da Cunha EFF, Ramalho TC. 2017. Towards the understanding of tetrahydroquinolines action in *Aedes aegypti*: Larvicide or adulticide? *Mol Simulat* 43:121–133.
- Gunamalai V, Kirubakaran R, Subramoniam T. 2004. Hormonal coordination of molting and female reproduction by ecdysteroids in the mole crab *Emerita asiatica* (Milne Edwards). *Gen Comp Endocrinol* 138:128–138.
- Harada T, Nakagawa Y, Akamatsu M, Miyagawa H. 2009. Evaluation of hydrogen bonds of ecdysteroids in the ligand–receptor interactions using a protein modeling system. *Bioorg Med Chem* 17:5868–5873.
- Hill RJ, Billas IML, Bonneton F, Graham LD, Lawrence MC. 2013. Ecdysone receptors: From the Ashburner model to structural biology. In Berenbaum MR, ed, *Annual Review of Entomology*, Vol 58. Annual Reviews, Palo Alto, CA, USA, pp 251–271.
- Kasuya A, Sawada Y, Tsukamoto Y, Tanaka K, Toya T, Yanagi M. 2003. Binding mode of ecdysone agonists to the receptor: Comparative modeling and docking studies. *J Mol Model* 9:58–65.
- Kortenkamp A, Martin O, Faust M, Evans R, McKinlay R, Orton F, Rosivatz E. 2011. State of the art assessment of endocrine disruptors. Final Report. European Union, Brussels, Belgium.
- LaLone CA, Villeneuve DL, Lyons D, Helgen HW, Robinson SL, Swintek JA, Saari TW, Ankley GT. 2016. Sequence alignment to predict across species susceptibility (SeqAPASS): A web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. *Toxicol Sci* 153:228–245.
- Landrum G. 2016. RDKit: Open-source cheminformatics software. [cited 2019 July 20]. Available from: <http://www.rdkit.org/>
- Mellor CL, Steinmetz FP, Cronin MTD. 2016. Using molecular initiating events to develop a structural alert based screening workflow for nuclear receptor ligands associated with hepatic steatosis. *Chem Res Toxicol* 29:203–212.
- Nakagawa Y, Henrich VC. 2009. Arthropod nuclear receptors and their role in molting. *FEBS J* 276:6128–6157.
- Organisation for Economic Co-operation and Development. 2012. Test No. 211: *Daphnia magna* reproduction test. Technical Report. Paris, France.
- Rajathi A, Pandiarajan J, Krishnan M. 2010. Effect of RH-2485 on development, metamorphosis and synthesis of major proteins in female silkworm *Bombyx mori*. *Biologia* 65:903–913.
- Richard AM, Judson RS, Houck KA, Grulke CM, Volarath P, Thillainadarajah I, Yang C, Rathman J, Martin MT, Wambaugh JF, Knudsen TB, Kancheria J, Mansouri K, Patlewicz G, Williams AJ, Little SB, Crofton KM, Thomas RS. 2016. Toxcast chemical landscape: Paving the road to 21st century toxicology. *Chem Res Toxicol* 29:1225–1251.
- Rodriguez EM, Medesani DA, Fingerman M. 2007. Endocrine disruption in crustaceans due to pollutants: A review. *Comp Biochem Physiol A* 146:661–671.
- Skakkebaek NE, Toppari J, Soder O, Gordon CM, Dival S, Draznin M. 2011. The exposure of fetuses and children to endocrine disrupting chemicals: A European Society for Paediatric Endocrinology (ESPE) and Pediatric Endocrine Society (PES) call to action statement. *J Clin Endocrinol Metab* 96:3056–3058.
- Smith HC, Cavanaugh CK, Friz JL, Thompson CS, Saggars JA, Michelotti EL, Garcia J, Tice CM. 2003. Synthesis and SAR of cis-1-benzoyl-1,2,3,4-tetrahydroquinoline ligands for control of gene expression in ecdysone responsive systems. *Bioorg Med Chem Lett* 13:1943–1946.
- Soin T, De Geyter E, Mosallanejad H, Iga M, Martin D, Ozaki S, Kitsuda S, Harada T, Miyagawa H, Stefanou D, Kotzia G, Efrose R, Labropoulou V, Geelen D, Iatrou K, Nakagawa Y, Janssen CR, Smaghe G, Swevers L. 2010a. Assessment of species specificity of molting accelerating compounds in Lepidoptera: Comparison of activity between *Bombyx mori* and *Spodoptera littoralis* by in vitro reporter and in vivo toxicity assays. *Pest Manag Sci* 66:526–535.
- Soin T, Swevers L, Kotzia G, Iatrou K, Janssen CR, Rougé P, Harada T, Nakagawa Y, Smaghe G. 2010b. Comparison of the activity of non-steroidal ecdysone agonists between dipteran and lepidopteran insects, using cell-based EcR reporter assays. *Pest Manag Sci* 66:1215–1229.
- Song Y, Evenseth LM, Iguchi T, Tollefsen KE. 2017a. Release of chitinase as an indicator of potential molting disruption in juvenile *Daphnia magna* exposed to the ecdysone receptor agonist 20-hydroxyecdysone. *J Toxicol Environ Health A* 80:954–962.
- Song Y, Tollefsen KE. 2018. Ecdysone receptor agonism leading to incomplete ecdysis associated mortality. [cited 2019 July 20]. Available from: <https://aopwiki.org/aops/4>
- Song Y, Villeneuve DL, Toyota K, Iguchi T, Tollefsen KE. 2017b. Ecdysone receptor agonism leading to lethal molting disruption in arthropods: Review and adverse outcome pathway development. *Environ Sci Technol* 51:4142–4157.
- Stanley J, Preetha G. 2016. *Pesticide Toxicity to Non-target Organisms*. Springer, Dordrecht, The Netherlands.
- Steinbeck C, Han YQ, Kuhn S, Horlacher O, Luttmann E, Willighagen E. 2003. The Chemistry Development Kit. [cited 2019 July 20]. Available from: <https://cdk.github.io/>
- Steinmetz FP, Mellor CL, Meil T, Cronin MTD. 2015. Screening chemicals for receptor-mediated toxicological and pharmacological endpoints: Using public data to build screening tools within a KNIME workflow. *Mol Inform* 34:171–178.

- Sumiya E, Ogino Y, Miyakawa H, Hiruta C, Toyota K, Miyagawa S, Iguchi T. 2014. Roles of ecdysteroids for progression of reproductive cycle in the fresh water crustacean *Daphnia magna*. *Front Zool* 11:60.
- Tice CM, Hormann RE, Thompson CS, Friz JL, Cavanaugh CK, Michelotti EL, Garcia J, Nicolas E, Albericio F. 2003a. Synthesis and SAR of α -acylaminoketone ligands for control of gene expression. *Bioorg Med Chem Lett* 13:475–478.
- Tice CM, Hormann RE, Thompson CS, Friz JL, Cavanaugh CK, Saggars JA. 2003b. Optimization of α -acylaminoketone ecdysone agonists for control of gene expression. *Bioorg Med Chem Lett* 13:1883–1886.
- Tollefsen KE, Scholz S, Cronin MT, Edwards SW, de Knecht J, Crofton K, Garcia-Reyero N, Hartung T, Worth A, Patlewicz G. 2014. Applying adverse outcome pathways (AOPs) to support integrated approaches to testing and assessment (IATA). *Regul Toxicol Pharmacol* 70: 629–640.
- Tyler CR, Jobling S, Sumpter JP. 1998. Endocrine disruption in wildlife: A critical review of the evidence. *Crit Rev Toxicol* 28:319–361.
- Yamanaka N, Rewitz KF, O'Connor MB. 2013. Ecdysone control of developmental transitions: Lessons from *Drosophila* research. In Berenbaum MR, ed, *Annual Review of Entomology*, Vol 58. Annual Reviews, Palo Alto, CA, USA, pp 497–516.
- Yang C, Tarkhov A, Maruszyk J, Bienfait B, Gasteiger J, Kleinoeder T, Magdziarz T, Sacher O, Schwab C, Schwoebel J, Terfloth L, Arvidson K, Richard A, Worth A, Rathman J. 2015. New publicly available chemical query language, CSRML, to support chemotype representations for application to data mining and modelling. [cited 2019 July 20]. Available from: <https://chemotyper.org/>
- Yokota H, Eguchi S, Nakai M. 2011. Development of an *in vitro* binding assay for ecdysone receptor of mysid shrimp (*Americamysis bahia*). *Aquat Toxicol* 105:708–716.
- Zotti MJ, Christiaens O, Rouge P, Grutzmacher AD, Zimmer PD, Smagghe G. 2012. Sequencing and structural homology modeling of the ecdysone receptor in 2 chrysopids used in biological control of pest insects. *Ecotoxicology* 21:906–918.