



An extensive ecdysteroid CoMFA

Laurence Dinan^a, Robert E. Hormann^{b,*} & Ted Fujimoto^b

^aDepartment of Biological Sciences, University of Exeter, Perry Road, Exeter EX4 4QG, U.K.; ^bRohm and Haas Co., 727 Norristown Road, Spring House, PA 19477-0904, U.S.A.

Received 11 May 1998; Accepted 17 August 1998

Key words: CoMFA, 3D QSAR, ecdysone, ecdysteroid, ligand-binding, receptor

Summary

The ecdysteroid agonist activity of 71 HPLC-purified ecdysteroids was measured in the *Drosophila melanogaster* B_{II} tumorous blood cell line assay. The resultant log(ED₅₀) values, spanning almost 6 orders of magnitude, were used to construct a comparative molecular field analysis (CoMFA) model in which conformations were selected by homology to the crystal structure of ecdysone. Model A was constructed by utilization of the region-focused electrostatic indicator field ($q^2 = 0.631$, $r^2 = 0.903$, 5 components, 4 outliers). Model B made use of region-focused electrostatic and steric indicator fields along with MlogP ($q^2 = 0.694$, $r^2 = 0.892$, 5 components, 4 outliers). The model and its underlying bioassay data support a pharmacophore hypothesis in which ecdysteroid binding is understood to be due principally to the summation of localized interactions from approximately six specific loci. This is in contrast to previous structure-activity relationship hypotheses which are formulated in terms of the presence or absence of essential functional groups, without which ecdysteroid receptor affinity would be completely absent. The present CoMFA model is utilized to predict the activities of heretofore unknown ecdysteroids.

Introduction

20-Hydroxyecdysone (20E) is the steroid hormone responsible for onset and regulation of molting in almost all arthropods [1]. Ecdysteroids, structurally diverse congeners of 20E, occur in arthropods as well as in plants (phytoecdysteroids) [2]. The molt-regulating activity of 20E and its natural analogs is expressed through association with the proteins comprising the ecdysone receptor complex (members of the steroid receptor superfamily), and subsequent promotion of gene expression [3]. Ecdysteroid quantitative structure–activity relationships (QSARs) are important for the design of new ligands, understanding the ecology of phytoecdysteroids, and proper utilization of the expanding body of structural information on ecdysteroid receptor proteins.

The ecdysteroid receptor (EcR) proteins of *Drosophila melanogaster* are the most extensively studied of those from any invertebrate. There are three isoforms (A, B1 and B2) which arise from different

promoters and alternative splicing of a single-copy gene. All three isoforms possess common DNA- and ligand-binding domains [3b]. The active form of the receptor is a heterodimer with the ultraspiracle protein (USP), yet another member of the steroid receptor superfamily, for which no ligand with ecdysteroid-like activity is known [3d, 3e]. However, recently juvenile hormones have been shown to bind to USP [4]. Binding of the steroid to the EcR protein alters the affinity of the EcR/USP complex for short stretches of DNA (ecdysteroid response elements) associated with ecdysteroid-responsive genes, thereby modifying transcription of the genes [3a].

The binding event of ecdysteroids with the EcR/USP complexes from *Drosophila melanogaster* can be indirectly, but reliably, measured by the B_{II} tumorous blood cell line assay [5], in which the extent of ecdysteroid-mediated cell aggregation can be measured turbidometrically. In this study, we report the structure, binding affinities, and comparative molecular field analysis (CoMFA) [6] of 71 ecdysteroids. The molecular features which are important for defin-

*To whom correspondence should be addressed.

ing ecdysteroid activity and, ultimately, binding to the ecdysteroid receptor complex are described. The resultant model is used to predict activities for an ecdysteroid test set as well as hypothetical ecdysteroids. We believe that this study represents the first uniform treatment of a truly extensive set of ecdysteroids.

QSAR analysis of the ecdysteroids is a matter of current interest. Ecdysone was first isolated in 1954 by Karlson [7]. Since that time, numerous ecdysteroids from both plant and animal sources have been identified and examined in various whole insect and tissue culture assays. To date, over 250 ecdysteroids are known [2]. As a result, a qualitative and general understanding of the molecular features responsible for ecdysteroidal action can be constructed. However, a consistent, unified activity scale has long eluded investigators. This has been due to sample inavailability, assay infidelity, and limitations of commonly used QSAR methods. Sample inavailability has been a serious limitation since steroid specimens are often obtained only in minute quantities after toilsome purifications. Structural assignments sometimes remain ambiguous. Assay infidelity is a concern due to the obfuscating factors of uptake and metabolism inherent in whole organism studies. Finally, historically popular QSAR methods such as Hansch analysis are not designed to specifically handle the problems associated with flexible, noncongeneric structures [8]. Consequently, structure–activity analysis of ecdysteroids has historically had a patchwork quality. The present study strives to address each of these issues in turn. First, a diverse collection of ecdysteroids of widely disparate origin, high purity, and clear structural assignment was assembled. Second, an assay known to be largely free of metabolism and transport ambiguities, the B_{II} assay, was employed. Third, CoMFA, a QSAR tool adequate for the task of noncongeneric structures, was utilized.

The purposes of this investigation are multifold: to (i) define an extensive steroid set with associated receptor data; (ii) provide a benchmark QSAR for the ecdysteroids by generating a CoMFA model based on a simple alignment procedure for the 71 ecdysteroids; and (iii) refine previous qualitative structure–activity relationship (SAR) hypotheses with a QSAR model, and provide a graphical description of the molecular features relevant for ecdysteroid binding. The model's predictive ability is tested against a series of ecdysteroids. This model provides a point of comparison for alternative QSAR methods, in particular two-dimensional (2D) descriptor QSAR and refined conformational treatments in CoMFA.

Methods

Sourcing

Ecdysteroids were isolated from plant sources, purchased from commercial sources, or synthesized. Most were generously provided by other researchers (listed in the Acknowledgements section). Ecdysteroid structures appear in Table 1.

Purification

The purity of all ecdysteroid samples was initially assessed on an analytical gradient reversed-phase HPLC system (25 cm \times 4.6 mm i.d. Spherisorb ODS-2 column, 5 μ m particle size, eluted at 1 ml/min with a linear gradient from 30% methanol/water to 100% methanol over 30 min, followed by elution with methanol for 10 min). Each of the compounds was then purified on a semi-preparative RP-HPLC system (25 cm \times 9.2 mm i.d. Spherisorb ODS-2, 5 μ m particle size) eluted isocratically with an appropriate methanol/water mixture. The purity of the RP-purified sample was determined on an NP-HPLC system (15 cm \times 4.6 mm i.d. Apex II DIOL column, 5 μ m particle size, eluted isocratically with 2–10% methanol in dichloromethane at 1 ml/min) and, if necessary, further purified in this system.

Quantification

Ecdysteroid concentrations were determined spectrophotometrically as methanolic solutions using published values for the molar extinction coefficients [2a].

Data set and assay

The agonist and antagonist activities of each ecdysteroid were determined in a bioassay based on the ecdysteroid-responsive *Drosophila melanogaster* B_{II} cell line growing in wells of microtiter plates [5]. Briefly, ecdysteroids were prepared as stock solutions (10^{-3} M to 10^{-9} M) in methanol. Aliquots (20 μ l) of each dilution were transferred to wells of a microtiter plate with (antagonist assay) or without (agonist assay) 20 μ l 5×10^{-7} M 20-hydroxyecdysone in methanol. Solvent was allowed to evaporate in a sterile air flow. Then 200 μ l of cell suspension in Schneider's medium (ca. 4×10^5 cells/ml) was added to each well and the covered plate was incubated in a humid environment at 25 °C for 7 days. Cellular response was measured turbidometrically at 405 nm with an Anthos IIa microplate reader and the specificity of the response was confirmed by examination of

Table 1. Ecdysteroid training set for CoMFA analysis

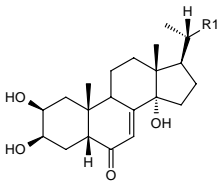
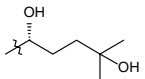
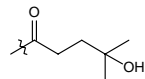
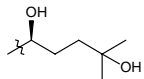
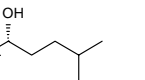
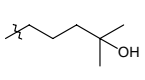
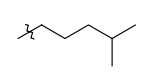
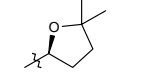
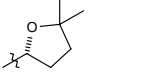
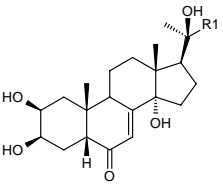
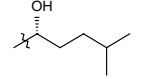
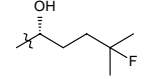
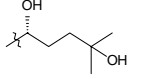
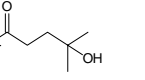
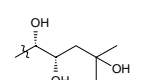
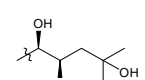
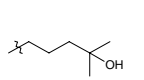
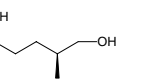
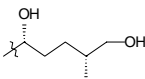
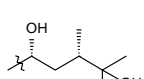
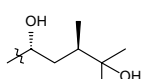
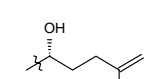
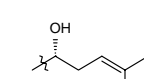
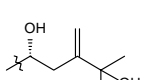
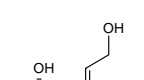
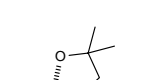
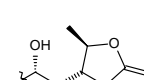
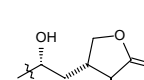
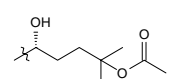
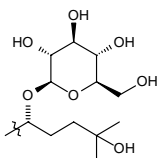
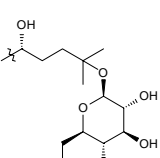
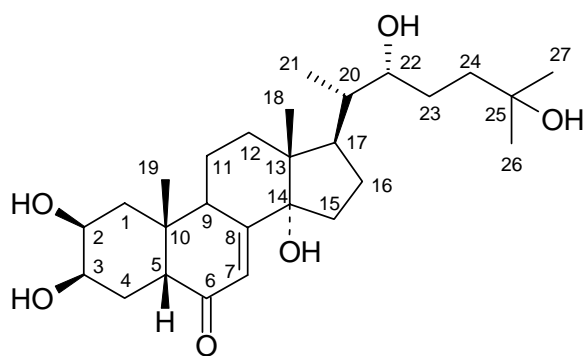
 Ecdysone Core	 1	 2	 3	 4
	 5	 6	 7	 8
 20-Hydroxyecdysone Core	 9	 10	 11	 12
	 13	 14	 15	 16
 17	 18	 19	 20	 21
 22	 23	 24	 25	 26
 27	 28	 29		

Table 1. (continued)

5, 20-Dihydroxyecdysone Core	30	31	32
2-Substituted-20-hydroxyecdysone Core	33	34	35
20-Iso-ecdysone Core			
	38	39	40

the appearance of the cells in the wells with an inverted microscope. Absorbance values at 405 nm (A_{405}) were related to those of untreated cells and 20E-treated cells (10^{-7} M final concentration; full response). None of the tested ecdysteroids possessed antagonistic activity. The agonist assay was repeated with a narrower distribution of concentrations around the active concentration apparent from the preliminary assay. The potencies of the ecdysteroids are compared by means of the ED_{50} values, i.e. the concentration of the ecdysteroid required to reduce the A_{405} value by 50% of the A_{405} (zero response) – A_{405} (full response). The ED_{50} values appear in Table 2.

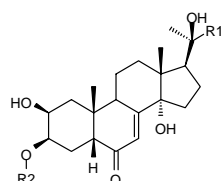
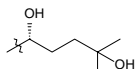
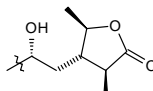
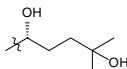


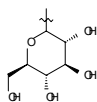
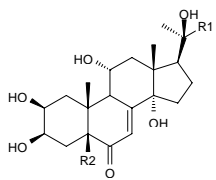
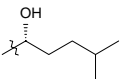
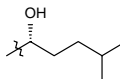
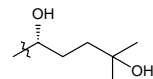
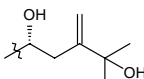
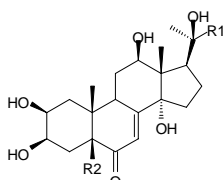
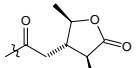
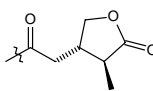
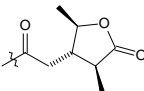
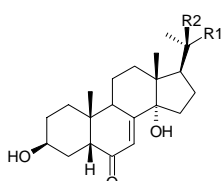
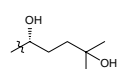
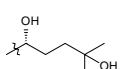
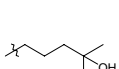
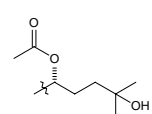
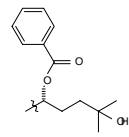


Structure 1.

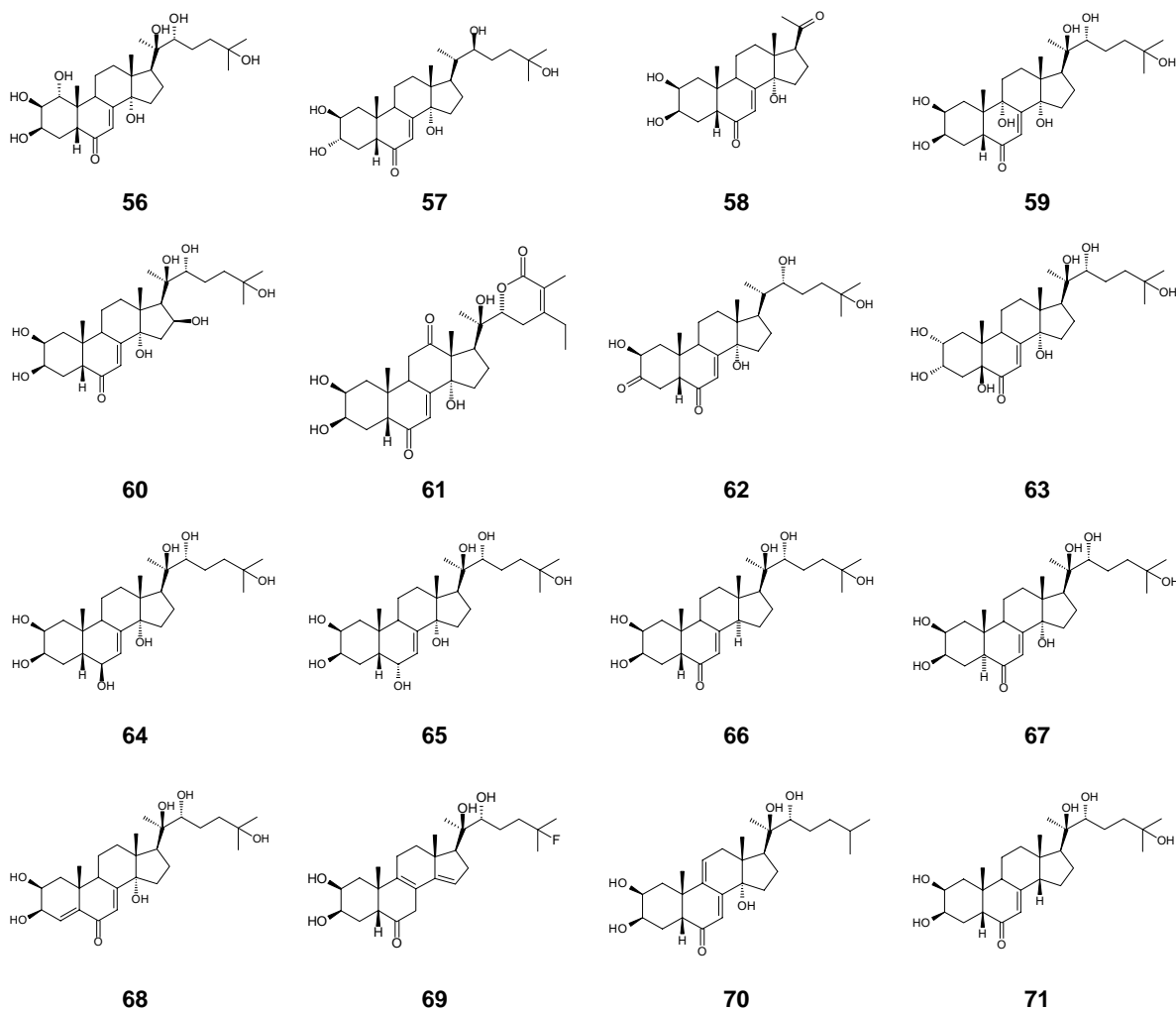
Conformation selection

Molecular modeling was performed using SYBYL Molecular Modeling Software, version 6.3 [9] running

Table 1. (continued)

	R1					
	R2					
		41	42	43		
	R1					
	R2	H	OH	H	H	
		44	45	46	47	
	R1					
	R2	OH	OH	H		
		48	49	50		
	R1					
	R2	H	OH	OH	H	OH
		51	52	53	54	55

Core Modification and Unusual Substitution



on an SGI R10000. The Advanced CoMFA module is part of the SYBYL software package. Where parameter values are not indicated, SYBYL default values were used. Ecdysone (**1**), in its crystal structure conformation [10], was selected as the reference molecule. Other structures in their modeled conformations were constructed by minimal modification of the ecdysone template, followed by energy minimization. Minimizations were performed utilizing the standard Tripos force field [11] including charges,

using conjugate gradient minimization until convergence was achieved at a gradient of 0.05 kcal/mol Å. Charges were generated by the Gasteiger–Hückel method [12]. Generally speaking, perturbation of the modified structures was small. Ecdysone itself was also minimized using similar conditions (see Structure 1).

Several groups of set members still required certain arbitrary conformational decisions: the lactone-bearing cyasterones (**25**, **26**, **42**, **50**) and sengosterones

Table 2. *Drosophila* B_{II} cell line $-\log(\text{ED}_{50})$ values for training set

No.	Steroid	ED_{50} (M)	$-\log(\text{ED}_{50})$	MlogP
1	Ecdysone	1.1×10^{-6}	5.96	2.20
2	22-Dehydroecdysone	4.5×10^{-8}	7.35	2.10
3	22- <i>epi</i> -Ecdysone	$3.5, 5.3 \times 10^{-6}$	5.36 ^a	2.20
4	25-Deoxyecdysone	1.0×10^{-8}	8.00	2.97
5	22-Deoxyecdysone	$\sim 10^{-5}$	5.30 ^b	2.97
6	22,25-Dideoxyecdysone	$\sim 10^{-6}$	6.30 ^b	3.76
7	20-(2',2'-Dimethyl-(22 <i>S</i>)-furanyl)ecdysone	7.3×10^{-6}	5.14	2.97
8	20-(2',2'-Dimethyl-(22 <i>R</i>)-furanyl)ecdysone	1.0×10^{-5}	5.00	2.97
9	Ponasterone A	3.1×10^{-10}	9.51	2.20
10	25-Fluoroponasterone	5.1×10^{-9}	8.29	2.30
11	20-Hydroxyecdysone	7.5×10^{-9}	8.12	1.44
12	22-Dehydro-20-hydroxyecdysone	1.7×10^{-7}	6.77	1.35
13	Geradiasterone	4.0×10^{-7}	6.40	0.69
14	22,23- <i>epi</i> -Geradiasterone	4.0×10^{-5}	4.40	0.69
15	22-Deoxy-20-hydroxyecdysone	1.4×10^{-8}	7.85	2.20
16	(25 <i>S</i>)-Inokosterone	2.7×10^{-7}	6.57	1.44
17	(25 <i>R</i>)-Inokosterone	1.5×10^{-7}	6.82	1.44
18	24- <i>epi</i> -Makisterone	2.2×10^{-7}	6.66	1.64
19	Makisterone A	1.3×10^{-8}	7.89	1.64
20	25,26-Didehydroponasterone A	1.5×10^{-7}	6.82	2.10
21	Stachysterone C ($\Delta 24[25]$)	1.4×10^{-8}	7.85	2.10
22	24(28)-Dehydromakisterone A	4.0×10^{-9}	8.40	1.54
23	24(28)-Dehydroamarasterone B	5.2×10^{-7}	6.28	1.74
24	Shidasterone	$1.5, 4 \times 10^{-6}$	5.56 ^a	2.20
25	Cyasterone	$1.2, 5.3 \times 10^{-8}$	7.49 ^a	1.80
26	29-Norcyasterone	1.2×10^{-8}	7.92	1.61
27	Viticosterone E	1×10^{-7}	7.00	1.80
28	20-Hydroxyecdysone-22- β -D-glucopyranoside	4.7×10^{-5}	4.33	-0.66
29	20-Hydroxyecdysone-25- β -D-glucopyranoside	8.5×10^{-6}	5.07	-0.66
30	Polypodine B	1.0×10^{-9}	9.00	0.69
31	25-Deoxypolypodine B	1.0×10^{-8}	8.00	1.44
32	5 β -Hydroxystachysterone C, 1 ^c	3.5×10^{-8}	7.46	1.35
33	5 β -Hydroxy-25,26-didehydroponasterone A ^c	6.3×10^{-8}	7.20	1.35
34	Sengosterone	9.0×10^{-8}	7.05	1.06
35	29-Norsengosterone	1.3×10^{-7}	6.89	0.87
36	20-Hydroxyecdysone-2-acetate	4×10^{-7}	6.40	1.80
37	20-Hydroxyecdysone-2- β -D-glucopyranoside	2×10^{-5}	4.70	-0.66
38	22-Dehydro-20- <i>iso</i> -ecdysone	3.0×10^{-6}	5.52	2.10
39	20- <i>iso</i> -Ecdysone	1.0×10^{-4}	4.00	2.20
40	20- <i>iso</i> -22- <i>epi</i> -Ecdysone	1.0×10^{-4}	4.00	2.20
41	20-Hydroxyecdysone-3-acetate	4.7×10^{-7}	6.33	1.80
42	Cyasterone-3-acetate	4.3×10^{-7}	6.37	2.17
43	20-Hydroxyecdysone-3- β -D-glucopyranoside	1.3×10^{-5}	4.89	-0.66
44	Ajugasterone C	$3.0, 8.0 \times 10^{-8}$	7.26 ^a	1.44
45	Muristerone A	2.2×10^{-8}	7.66	0.69
46	Turkesterone	$3.0 \times 10^{-7}, 1.3 \times 10^{-6}$	6.10 ^a	0.69
47	Atrotosterone C	3.0×10^{-6}	5.52	0.80

Table 2. (continued)

No.	Steroid	ED ₅₀ (M)	−log(ED ₅₀)	MlogP
48	22-Dehydro-12-hydroxysengosterone	2.7×10^{-6}	5.57	0.24
49	22-Dehydro-12-hydroxynorsengosterone	1.3×10^{-5}	4.89	0.05
50	22-Dehydro-12-hydroxycyasterone	1.3×10^{-6}	5.89	0.97
51	2-Deoxyecdysone	5.0×10^{-5}	4.30	2.97
52	2-Deoxy-20-hydroxyecdysone	6.6×10^{-7}	6.18	2.20
53	2,22-Dideoxy-20-hydroxyecdysone	$\sim 5 \times 10^{-5}$	4.60 ^b	2.97
54	2-Deoxyecdysone-22-acetate	1.4×10^{-6}	5.85	3.31
55	2-Deoxy-20-hydroxyecdysone-22-benzoate	6.3×10^{-6}	5.20	2.67
56	1- <i>epi</i> -Integristerone A	2.5×10^{-7}	6.60 ^a	0.69
57	3- <i>epi</i> -22- <i>iso</i> -Ecdysone	4.4×10^{-6}	5.36	2.20
58	Poststerone	$\sim 2 \times 10^{-5}$	4.98 ^b	1.62
59	9,20-Dihydroxyecdysone	1.6×10^{-5}	4.80	0.69
60	Malacosterone	9.0×10^{-6}	5.05	0.69
61	Ajugalactone	1.6×10^{-7}	6.80	1.63
62	3-Dehydroecdysone	6.0×10^{-6}	5.22	2.10
63	Rapisterone D	1.0×10^{-9}	9.00	0.69
64	6 β -Hydroxy-20-hydroxyecdysone	1.7×10^{-7}	6.77	1.55
65	6 α -Hydroxy-20-hydroxyecdysone	2.0×10^{-6}	5.70	1.55
66	14-Deoxy(14 α -H)-20-hydroxyecdysone	3.0×10^{-8}	7.52	2.20
67	(5 α -H)20-Hydroxyecdysone	3.3×10^{-6}	5.48	1.44
68	4-Dehydro-20-hydroxyecdysone	2.8×10^{-7}	6.55	1.35
69	2 β ,3 β ,20 R ,22 R -Tetrahydroxy-25-fluoro-5 β -cholest-8,14-dien-6-one	7.2×10^{-9}	8.14	2.98
70	5-Deoxykaladasterone	5.2×10^{-10}	9.28	2.10
71	14-Deoxy(14 β -H)-20-hydroxyecdysone	8.3×10^{-7}	6.08	2.20

^a Average between two concentrations of observed activity.

^b Midpoint between concentrations of observed activity and absence of activity.

^c Assignment of **32** vis-à-vis **33** based on analogy to HPLC elution order of **20** and **21**.

(**34**, **35**, **48**, **49**), structures with nonrigid functionalization other than hydroxyl at C-2, C-3 and C-22, steroid **27** (25-acetyl), lactone **61**, and steroid **78** (17-acetoxy).

For the cyasterones and sengosterones, two C-23/C-24 rotomers come into question, each of which overlaps a different lactone carbon with C-25 of ecdysone. On the principle of minimal modification, the conformation in which the carbonyl carbon of the cyasterone coincides with the *pro-S* carbon (C-27) of ecdysone was chosen. For 29-norcyasterone (**26**), the carbonyl carbon was aligned with the *pro-S* carbon of ecdysone, and the α -CH₃ was aligned with the *pro-R* carbon (Figure 1).

For structures with glucosyl groups at C-2, C-3, and C-22 (**28**, **37**, **43**), the glucosyl moieties themselves were placed in the all-equatorial conformation, and glucosyl ether linkage torsions were chosen which pointed the sugar moiety as much away from the steroid core and chain as possible, a condition that

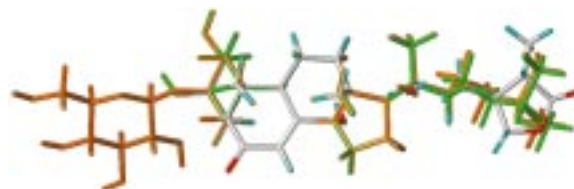


Figure 1. Alignment of representative ecdysteroids ecdysone **1** (green), 29-norcyasterone **26** (white), and 20-hydroxyecdysone-3- β -D-glucopyranoside **43** (orange).

could be satisfied by several conformations. Structures with acetyl or benzoyl groups in corresponding positions were arranged such that the carbonyl oxygen maximally coincided with the glycosidic ring oxygen. In each case, this implied setting the ester in the *s-trans* configuration.

Two structures bear non-hydroxyl functionality at C-2: **36** and **37**. The glucosyl group of **37** was oriented and minimized such that the C2-O-C α -O torsion was 179° and the C1-C2-O-C α torsion was 60°. The

acetyl group of **36** was correspondingly oriented such that the C2-O-C(O)-O torsion was -165° and the C1-C2-O-C(O) torsion was 81° . Three structures bear non-hydroxyl functionality at C-3: **41**, **42** and **43**. For these compounds, the 3-acetyl-substituted steroids **41** and **42** were oriented and minimized such that the C2-C3-O-C(O) torsions were 161° and the C3-O-C(O)-O torsions were 164° , while the glycosyl group of **43** was correspondingly oriented such that the C2-C3-O-C(α) torsion was 136° and the C3-O-C(α)-O torsion was 176° .

Three structures bear non-hydroxyl functionality at C-22: **28**, **54** and **55**. The glycosyl group of **28** was oriented and minimized such that the C22-O-C α -O torsion was -157° , and the C20-C22-O-C(α) torsion was 93° . The acetyl group of **54** was oriented in the *s-trans* conformation with a C20-C22-O-C(O) torsional angle of 80° , and a C22-O-C(O)-O torsion of -164° . Benzoyl derivative **55** was correspondingly aligned in an *s-trans* conformation such that its C20-C22-O-C(O) torsion was 78° and its C22-O-C(O)-O torsion was -178° . One structure bears non-hydroxyl functionality at C-2, C-3, and C-22: 20-hydroxyecdysone-2,3,22-triacetate **81**. Each of the acetates were aligned with the 2-, 3-, and 22-acetate groups of **36**, **41**, and **54**, respectively, and minimized to achieve the final conformation for superposition.

The C-25 acetoxy group of compound **27** was placed in an *s-cis* conformation such that the group bisected the gem-dimethyl substituents. The lactone group of **61** was oriented such that the ring oxygen aligned with O-22 of the ecdysone crystal structure. Finally, the carbon-oxygen double bond of the 17-acetoxy group of **78** was aligned as much as possible with the C-22 hydroxyl group of ecdysone.

Alignment

The template molecule, ecdysone, was oriented in the coordinate lattice such that the first and second principal axes were aligned with the x and y coordinates, respectively. Within the SYBYL environment, this is best accomplished with the ORIENT BEST VIEW routine. The remainder of the steroid training set, excepting three structures, was aligned based on the 17 carbons comprising the steroid tetracyclic core, by executing the DATABASE ALIGN ALGORITHM in Advanced CoMFA, using the ecdysone crystal structure as the reference. Three structures not aligned in this manner contained core modifications (**68**, **69**, **70**), and were superimposed on ecdysone using 25 pairs of atoms comprising the core ring system plus the

chain with equal weights. Among the set of inactive steroids, seven (**82**, **83**, **86**, **87**, **88**, **89**, **90**) were superimposed on ecdysone using equal weights for the 25 pairs of atoms comprising the core ring system plus the eight chain atoms. Homobrassinolides **91** and **92** were aligned by the FIT routine using equal weighting of the nine atoms of rings C and D.

CoMFA field generation and optimization

CoMFA regions and fields were generated using SYBYL version 6.3. Tripos standard, indicator, and hydrogen bonding fields were examined. The Tripos standard fields are default fields generated using the Tripos force field and an sp^3 carbon atom as probe. The indicator field class is similar to the standard steric (Lennard-Jones) and electrostatic (Coulombic) potentials, but differs in that all lattice energies with magnitudes below the indicated threshold are replaced with a value of zero and energies at or above this threshold are replaced with the field cut-off value. When both steric and electrostatic fields are included in a single CoMFA column, the greater of the steric and electrostatic cut-offs is used [13a,b]. Hydrogen-bond fields are indicator fields in which lattice points with steric energies below the cut-off are assigned the cut-off value if they are close to hydrogen-bond accepting or donating atoms [13a,c]. Log P calculations were performed according to the method of Moriguchi et al. (MlogP) [14a] implemented by Blake [14b].

Parameters such as energy cut-offs (steric indicator: 10–25 kcal/mol; electrostatic indicator: 1–3 kcal/mol), dielectric function (r or $1/r$), grid spacing (1, 2, or 3 Å), and grid displacements (± 0.5 Å in the x,y, and/or z directions) were systematically varied (exploration of parameters for the hydrogen bond and standard fields was performed with slightly modified versions of the reported data set). Cut-off transitions were smooth and field values were eliminated within the steric cut-off for each row.

Partial least-squares (PLS) methodology [15] was used to analyze the relationship between the CoMFA field descriptors and $-\log(ED_{50})$. Field and parameter optimization, region focusing (vide infra), and elimination of outliers were guided by full cross-validated (leave-one-out) PLS analyses, by which cross-validated r^2 (q^2) values were maximized and standard error of prediction ($SPRESS$) values were minimized [6a]. Unless otherwise indicated, for non-cross-validated analyses, the number of components chosen was based on the lowest PRESS (predictive residual sum of squares) [6b]. Once satisfactory cross-

Table 3. Summary of non-region-focused indicator fields optimized for cross-validated r^2 (q^2) values for ecysteroid affinities in the B_{II} receptor assay

Fields ^a	Energy cut-offs (kcal/mol)	Dielectric function	q^2	No. of components (std error pred)	No. of columns used
Steric	25	1/r	0.307	3 (1.14)	209
Elec ^b	2	1/r	0.359	5 (1.12)	408
Both	2.25/25	1/r	0.350	4 (1.12)	1036
Elec/ster	2.0/25	1/r	0.413	5 (1.07)	617
with MlogP					
Steric	25	1/r	0.405	4 (1.07)	210
Elec	2	1/r	0.332	5 (1.14)	409
Both	2.25/25	1/r	0.405	3 (1.06)	1037
Elec/ster ^c	2.0/25	1/r	0.418	5 (1.06)	618

^aCoMFA region = default; minimum $\sigma = 0.5$; step size = 2 Å; probe atom type = C_{sp3} (+).

^b Region-focused for model A.

^c Region-focused for model B.

Table 4. Standard error of prediction (cross-validated PLS analyses, $-\log(\text{ED}_{50})$) and % variance (non-cross-validated PLS analyses, $-\log(\text{ED}_{50})$) for models A and B

	Number of components				
	1	2	3	4	5
Model A					
Standard error of prediction	1.051	0.989	0.903	0.866	0.821
% variance	0.527	0.662	0.788	0.845	0.903
Model B					
Standard error of prediction	1.134	0.905	0.817	0.772	0.736
% variance	0.321	0.638	0.762	0.830	0.892

validated models were obtained, non-cross-validated PLS analyses were performed using the optimum number of components from the cross-validated run. These final non-cross-validated analyses (models A and B) were used for prediction purposes.

Region focusing, the relative weighting of lattice points, was performed using the full PLS analysis based on the optimum unfocused analysis and its corresponding region; the number of components used was usually the number recommended by SAMPLS (sample-distance partial least-squares) [16]. All reported grid spacing is at 2 Å; whenever attempted, a grid spacing of 1 or 3 Å resulted in inferior PLS results. Grid points were weighted by discriminant power, which is based on relative lattice point contribution to the model components. Weighting by modeling power (amount of variation accounted for by

model) and by (standard deviation) \times (coefficient) was also explored, but gave inferior results [17].

To identify outlying members of the training set, for the best two models, q^2 was optimized by systematically excluding the worst-predicted members of the training set, singly or in combination. Candidate outliers were subsequently evaluated for structural uniqueness. Once a satisfactory edited training set was obtained, the CoMFA columns were recalculated and focused anew; alternatively, the already focused region was refocused (the CoMFA default region (x: -18.082 to 15.19 Å; y: -12.87 to 11.00 Å; z: -10.03 to 9.59 Å; 4 Å extension beyond training set members) remained unaltered upon outlier elimination).

Results

CoMFA field optimization

Standard steric and electronic fields (alone and in combination), indicator steric and electronic fields (alone and in combination), and hydrogen-bonding fields were tested in the CoMFA model and optimized as a function of energy cut-off, dielectric function, CoMFA region displacement, minimum σ , step size, and probe atom type. For standard and hydrogen-bonding fields generated from preliminary versions of the data set, the optimum PLS analyses yielded q^2 values of less than 0.3; these could not be improved much by subsequent region focusing or combinations of region-focused fields, and were no longer considered. Indicator fields yielded more promising statistical results. Table 3 summarizes the optimized parameters and q^2 values for indicator fields. The best cross-validated PLS analysis ($q^2 = 0.418$, 5 components) utilized both steric and electronic fields with MlogP. We chose this analysis and the PLS analysis based on the electrostatic indicator field alone ($q^2 = 0.359$, 5 components) for further consideration.

Optimization of the indicator electrostatic analysis (model A)

The electrostatic indicator PLS analysis of Table 3 ($q^2 = 0.359$) was region-focused. The best results were obtained using discriminant power (exponential weight = 0.4, minimum $\sigma = 0.2$, 268 columns, 5 components, $q^2 = 0.488$). Inclusion of MlogP at this stage gave an inferior q^2 value of 0.321. The analysis without MlogP was optimized on q^2 by elimination of outliers. The most satisfactory PLS analysis yielded a q^2 value of 0.618 (5 components, $SPRESS = 0.836$) with elimination of steroids **15**, **22**, **30**, and **60**. Elimination of one outlier (**15**) or two outliers (**15**, **60**) gave meaningful analyses ($q^2 = 0.547$, $SPRESS = 0.938$; $q^2 = 0.585$, $SPRESS = 0.899$ respectively, 5 components).

The analysis with four outliers was chosen and further improvement was attempted by focusing a second time after elimination of the outliers ($q^2 = 0.631$, 5 components, $SPRESS = 0.821$; 260 CoMFA columns; weighting = discriminatory power^{0.3}). A non-cross-validated PLS analysis was performed; the final parameters and statistics appear in Tables 4 and 5 as model A. Despite large residuals, we chose to retain members **32**, **49**, and **51** to increase the robustness of the model.

Table 5. Statistical summary for PLS analyses of model A (electrostatic indicator field) and model B (electrostatic and steric indicator fields; MlogP)

Analysis	Model A	Model B
No. of training set members	67 ^a	67 ^b
No. of CoMFA columns	260	381
Minimum σ	0.1	0.1
Optimal no. of components	5	5
q^2	0.631	0.694
Standard error of prediction	0.821	0.736
Standard error of estimate	0.421	0.437
r^2	0.903	0.892
F value	113.8 ^c	100.9 ^c
Probability of $R^2 = 0$	0	0
Relative contributions		
Steric	0	0.370
Electrostatic	1.0	0.541
MlogP	0	0.089

^a Outliers: **15**, **22**, **30**, **60**.

^b Outliers: **30**, **53**, **60**, **63**.

^c $n_1 = 5$, $n_2 = 61$.

Optimization of the electrostatic indicator/steric indicator/MlogP analysis (model B)

The steric indicator field (Table 3), which contributes to the optimized combined steric/electrostatic/MlogP PLS analysis, was region-focused. The best results were obtained using a discriminant power exponential weighting of 0.5. This region-focused steric indicator field was then combined with the region-focused electrostatic indicator field and MlogP (minimum $\sigma = 0.3$, 289 columns, 5 components, $q^2 = 0.487$, $SPRESS = 0.991$). This analysis was then optimized on q^2 by elimination of outliers. The most satisfactory PLS analysis yielded a q^2 value of 0.653 (5 components, $SPRESS = 0.783$) with elimination of steroids **30**, **53**, **60**, and **63**. Elimination of one outlier (**63**) or two outliers (**53**, **63**) gave meaningful analyses ($q^2 = 0.524$, $SPRESS = 0.936$; $q^2 = 0.572$, $SPRESS = 0.883$, respectively, 5 components).

The analysis with four outliers was chosen and further improvement was attempted by focusing a second time after elimination of the outliers ($q^2 = 0.694$, 5 components, $SPRESS = 0.736$; 381 CoMFA columns; weighting = discriminatory power^{0.4}). A non-cross-validated PLS analysis was performed; the final parameters and statistics appear in Table 4 and 5 as model B. Despite a residual greater than 2σ , we chose to retain members **1**, **11**, and **22** to increase the robustness of the model.

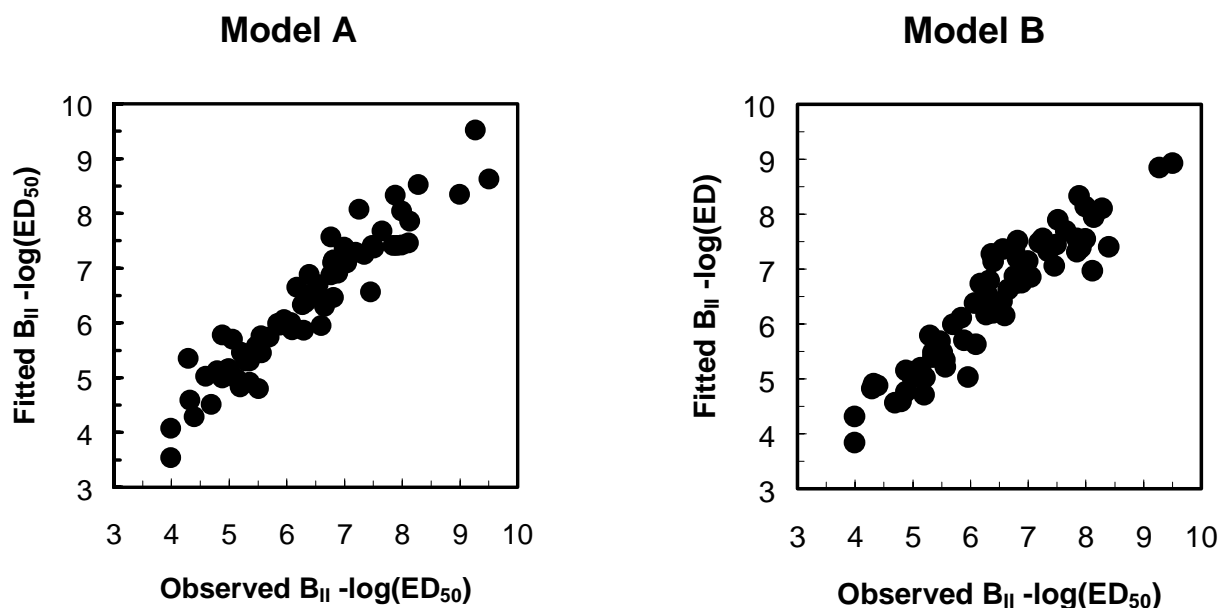


Figure 2. Comparison of actual vs. fitted $B_{II} - \log(ED_{50})$ values for models A and B.

Final models

Among a family of possible models, the PLS analyses described in Table 5 were chosen to represent this particular CoMFA approach applied to the ecdysteroid data set. PLS analysis utilizing the electrostatic indicator field alone (model A) resulted in a q^2 of 0.631 and an r^2 of 0.903 with a standard error of estimate of 0.421, when four training set outliers were removed ($n = 67$). PLS analysis utilizing the steric indicator field together with the electrostatic indicator field and MlogP (model B) resulted in a q^2 of 0.694 and an r^2 of 0.892 with a standard error of estimate of 0.437, when four training set outliers were removed.

The fitted $B_{II} - \log(ED_{50})$ values and residuals for both models are listed in Table 6. Graphs of the fitted versus actual values for both models are depicted in Figure 2. The electrostatic contour maps for model A and the electrostatic/steric contour maps for model B appear in Figures 3 and 4, respectively; ponasterone A is shown as a representative molecule. For model A, regions where positive charge is favored (blue) appear on the β -face between the C-2 and C-3 O-H dipoles, enveloping C-24, and wrapping around and extending to the end of the chain. Smaller blue regions appear at C-21 and below C-17. Important regions where positive charge is disfavored (red) appear at O-20, and between C-26 and C-27. Less significant red regions appear on the β -side of C-5, C-6 and C-8, and scattered around the terminus of the chain. For model B,

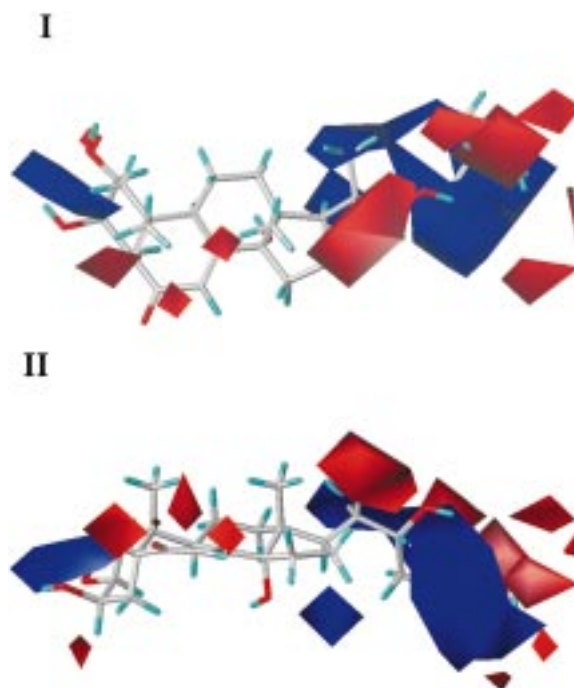


Figure 3. CoMFA electrostatic field contour plot ($\text{stdev} \times \text{coeff}$) for model A (electrostatic indicator field). Blue polyhedra represent regions where positive charge is favorable; contribution level = 75%. Red polyhedra represent regions where negative charge is favorable; contribution level = 30%. Ponasterone A **9**, one of the most active ecdysteroids, is depicted. Panel I is a view from the β -face; panel II is a view towards the C-6/C-14 edge of the steroid set.

Table 6. Measured $B_{II} - \log(ED_{50})$ values, predictions, and residuals for model A (electrostatic indicator field) and model B (electrostatic and steric indicator fields; MlogP)

Steroid	Actual	Model A		Model B	
		Calculated	Residual	Calculated	Residual
1	5.96	6.06	-0.10	5.03	0.93
2	7.35	7.25	0.11	7.32	0.03
3	5.36	4.92	0.45	5.39	-0.03
4	8.00	7.42	0.58	7.54	0.46
5	5.30	5.30	0.00	5.78	-0.48
6	6.30	5.87	0.43	6.53	-0.23
7	5.14	4.98	0.16	5.19	-0.05
8	5.00	5.16	-0.16	5.01	-0.01
9	9.51	8.63	0.88	8.92	0.59
10	8.29	8.53	-0.24	8.10	0.19
11	8.12	7.46	0.66	6.96	1.16
12	6.77	7.57	-0.80	7.37	-0.60
13	6.40	6.88	-0.48	7.13	-0.73
14	4.40	4.29	0.12	4.87	-0.47
15	7.85			7.31	0.54
16	6.57	6.45	0.12	7.36	-0.79
17	6.82	6.46	0.36	7.51	-0.69
18	6.66	6.30	0.36	6.62	0.04
19	7.89	8.33	-0.44	8.32	-0.43
20	6.82	7.14	-0.32	7.18	-0.36
21	7.85	7.41	0.44	7.55	0.31
22	8.40			7.40	1.00
23	6.28	6.33	-0.05	6.16	0.12
24	5.56	5.76	-0.20	5.34	0.23
25	7.49	7.41	0.08	7.43	0.06
26	7.92	7.41	0.51	7.38	0.54
27	7.00	7.38	-0.38	7.14	-0.14
28	4.33	4.59	-0.26	4.90	-0.57
29	5.07	5.70	-0.63	5.07	0.00
30	9.00				
31	8.00	8.04	-0.04	8.13	-0.13
32	7.46	6.56	0.90	7.05	0.41
33	7.20	7.28	-0.08	7.48	-0.28
34	7.05	7.09	-0.04	6.85	0.20
35	6.89	6.91	-0.02	6.74	0.15
36	6.40	6.52	-0.11	6.20	0.21
37	4.70	4.51	0.19	4.56	0.14
38	5.52	5.54	-0.02	5.44	0.08
39	4.00	4.07	-0.07	4.30	-0.30
40	4.00	3.53	0.47	3.83	0.17
41	6.33	6.37	-0.04	6.78	-0.45
42	6.37	6.44	-0.07	7.27	-0.90
43	4.89	4.99	-0.10	4.76	0.13
44	7.26	8.07	-0.81	7.55	-0.29
45	7.66	7.67	-0.01	7.69	-0.03
46	6.10	5.87	0.23	5.62	0.48
47	5.52	4.80	0.72	5.48	0.04
48	5.57	5.45	0.12	5.21	0.36

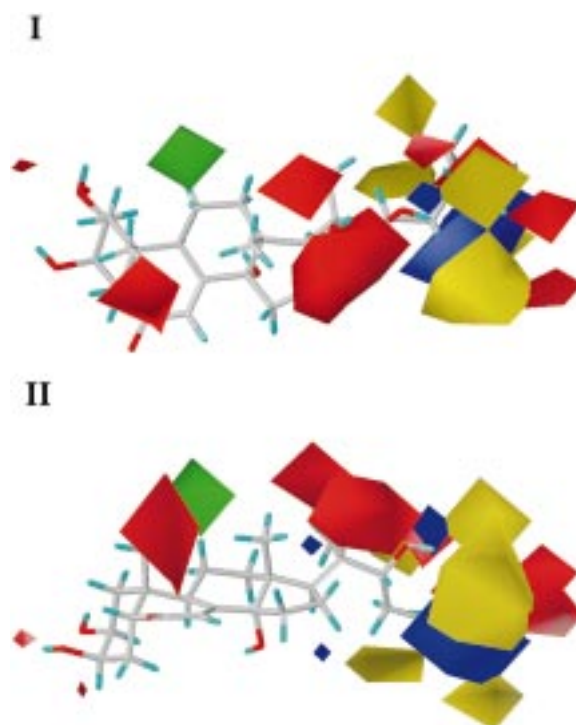


Figure 4. CoMFA electrostatic field contour plot (stdev \times coeff) for model B (electrostatic and steric indicator fields; MlogP). Blue polyhedra represent regions where positive charge is favorable; contribution level = 70%. Red polyhedra represent regions where negative charge is favorable; contribution level = 25%. Green polyhedra represent sterically favored regions; contribution level = 75%. Yellow polyhedra represent sterically disfavored regions; contribution level = 25%. Ponasterone A **9**, one of the most active ecdysteroids, is depicted. Panel I is a view from the β -face; panel II is a view towards the C-6/C-14 edge of the steroid set.

the situation is similar. Regions where positive charge is favored (blue) appear on the near side of C-24 and extending to C-26. Negative-charge-favored regions are more significant, however, occurring on the β -face of C-6, near O-20, on the β -face of C-26/C-27 and extending to H-25. In model B, regions where steric bulk is favored appear between α - and β -11-H. Regions where steric bulk is disfavored (yellow) appear around the perimeter of the chain.

Discussion

Choice and application of the CoMFA method

The goal of this investigation is to establish an initial ecdysteroid CoMFA model on an extensive basis set. In so doing, longstanding qualitative ideas about ecdysteroid structure–activity relationships (SARs)

Table 6. (continued)

Steroid	Actual	Model A		Model B	
		Calculated	Residual	Calculated	Residual
49	4.89	5.78	−0.89	5.14	−0.25
50	5.89	5.96	−0.07	5.69	0.20
51	4.30	5.35	−1.05	4.81	−0.51
52	6.18	6.65	−0.47	6.72	−0.54
53	4.60	5.02	−0.42		
54	5.85	5.99	−0.14	6.10	−0.25
55	5.20	4.83	0.37	4.71	0.50
56	6.6	5.95	0.65	6.15	0.46
57	5.36	5.30	0.06	5.47	−0.11
58	4.98	5.15	−0.17	5.12	−0.14
59	4.80	5.13	−0.33	4.58	0.22
60	5.05				
61	6.8	7.10	−0.30	6.74	0.06
62	5.22	5.46	−0.24	5.01	0.21
63	9.00	8.35	0.65		
64	6.77	6.87	−0.10	6.86	−0.09
65	5.70	5.74	−0.04	5.98	−0.28
66	7.52	7.36	0.16	7.89	−0.37
67	5.48	5.57	−0.09	5.68	−0.20
68	6.55	6.69	−0.14	6.41	0.15
69	8.14	7.85	0.29	7.94	0.20
70	9.28	9.52	−0.24	8.84	0.44
71	6.08	6.00	0.08	6.38	−0.30

are clarified, some initial predictions are made, and a point of comparison for alternative analyses is established.

CoMFA has certain distinct advantages over classical QSAR methods, namely, the ability to obtain useful information from small data sets, three-dimensional (3D) visualization, the ability to adequately handle noncongeneric structure sets, and utilization of conformation as a structure–activity factor. The last two aspects, however, frequently introduce the necessity for arbitrary decisions in the construction of a model, such as choice of conformation and molecular alignment. Additionally, CoMFA models require the adjustment of numerous parameters (grid size, field-type energy cut-offs, filters, etc.), which often may be arbitrary. This inherent difficulty is the subject of much current work in the development of adjunctive methods to conventional CoMFA [18] as well as further investigation in 2D descriptor methods applied to noncongeneric series [19]. In the final analysis, however, any QSAR model is useful and instructive insofar as it satisfactorily explains the available data, provides

physicochemical insights, and proves to be a reliable tool for prediction.

Conformation, alignment, superposition

As prescribed by the ecdysone crystal structure, the chair conformation was chosen for the A-ring. Structure alignment, based on the steroid ring system, was straightforward due to conformational rigidity. In cases where atom identity did not permit automated alignment, manual superposition using as many as 25 carbons, including side chain carbons, was performed. This approach gave quite satisfactory results.

Concerning side chain conformations resulting from energy minimizations, generally speaking, C-20, C-21, and C-22 remained quite consistently positioned across the steroid set, but larger displacements were observed towards the end of the side chain. For glycosidic and acetylated steroids, conformations were arbitrarily chosen which placed the groups in a direction as distant as possible from the nucleus. Alignment was arbitrary but self-consistent, based on correspondence of the carbonyl oxygen of the acetyl group and the ring oxygen of the glucosyl group. After conformations and alignments were chosen, *no subsequent manipulations of conformations and alignments were made in an attempt to optimize PLS analyses.*

Examining and testing the models

Contour map and functional group values. According to standard CoMFA protocol, regions in 3D space most sensitive to structural changes may be visualized by a contour map derived from the product of standard deviation and coefficient evaluated for each grid point. Such a contour map is depicted in Figure 3 (model A) and Figure 4 (model B). Examination of this CoMFA contour map is most instructive when considered together with the changes in $-\log(\text{ED}_{50})$ associated with specific changes in molecular features. The electrostatic contours for both models are quite similar and will be considered together. Two general regions are particularly worthy of note: the red region (negative-charge-favored) located at the C-20 hydroxyl and extending towards the C-22 hydroxyl, and the large blue region (positive-charge-favored) on or near the distal portion of the chain. These regions have a correspondence with trends associated with the presence of particular functionality. These trends, along with others, are tabulated in Table 7, where we have measured the change in $\log(\text{ED}_{50})$ associated with the listed structural changes for each of the indicated ecdysteroid pairs. Average values for Δ

$\log(\text{ED}_{50})$ are also listed. From Figures 3 and 4, and Table 7, one may infer that the red region near O-20 and extending towards O-22 appears as a result of the favorable effect of hydroxyl substitution at C-20 and C-22. We attribute the large blue region passing through (model A) or along (model B) the side chain to the accumulation of generally unfavorable effects resulting from oxygen substitution at C-24, C-26, and C-27, as well as the lactone-bearing cyasterones and sengosterones. The unfavorable effect of C-24/C-25 desaturation may possibly also be reflected in this blue region. Thus, both models A and B are dominated by very similar electrostatic fields.

Several other regions are undoubtedly meaningful, but evade straightforward interpretation. These include the blue region which appears on the β -face between the C-2 and C-3 hydroxyl dipoles, and red regions which appear on the β -face of C-6 and at H-25/C-26. Curiously, in model A, small red regions also appear on the β -face of C-5 and C-8. Displacements of the steroid nucleus due to core structural changes such as in **67–71** probably contribute to the presence of the contours near the core. The blue region along the C-2 and C-3 hydroxyl dipoles may be due in part to the presence of highly active rapisterone D (**63**) (inversion at C-2 and C-3) in the data set, and may also possibly be an indication of the detrimental effects of acetoxy or glucosyl substitution at C-2 and C-3. The red region near C-5 is probably not due to 5-OH substitution since both experimental data and the models interpret hydroxyl substitution at this position to generally depress activity (with some exceptions). Interpretation of the red region near C-26 is confounded by conformational variability at the terminus of the chain. It is interesting to speculate, however, on the possibility of a favorable effect of the lactone ring oxygen of the cyasterones and sengosterones, located in this region of space. The remaining electrostatic contours are difficult to interpret and may not be meaningful.

We understand the steric contour map of model B (green = sterically favored; yellow = sterically disfavored) to indicate simply that there are some general restrictions on steric bulk around the length of the side chain. There is probably some permissiveness at the terminus of the chain. Interestingly, a sterically favored region appears between α - and β -H-11. Four steroids in the data set bear α -hydroxyl substitution at the 11 position. Displacements of the steroid nucleus due to core structural changes may also contribute to this feature. In summary, model B may be understood as a refinement of model A in which steric

requirements and general hydrophobic requirements, as indicated by MlogP, are better defined.

Outliers. Structures **15**, **22**, **30**, and **60** are outliers in model A (predicted values: 5.7, 6.5, 6.9, 7.0, respectively). When entered as test structures, the predictions for the first three are underestimated; the last is overestimated. Several observations help explain these poor predictions. 22-Deoxy-20-hydroxyecdysone (**15**) lacks the 22-OH group; model A overestimates the contributing value of this group to binding by at least 0.5 $\log(\text{ED}_{50})$ units, based on simple group additivity arguments for the 22-OH group. Also, the modeled conformation of steroid **15** locates the 25-OH group in the large positive-charge-favored blue region, the distal portion of the chain being twisted slightly relative to ponasterone A, as shown in Figures 3 and 4. Although this twist is present in a number of steroids with hydroxyl substitution at C-25 (and which are undoubtedly responsible for the blue region), the effect on activity predictions may be particularly acute for **15**. Polypodine B (**30**), which bears a 5-OH group but is otherwise the same as 20E, is a surprisingly effective ligand. From experimental measurements, the 5-OH group depresses activity by about 0.25 $\log(\text{ED}_{50})$ units on average, but for **30** activity improves by 0.9 units. Arguments concerning the 25-OH group similar to those for **15** may also apply here. Steroids **22** (C-24-*exo*-methylene group) and **60** (hydroxyl substitution at C-16) are unusual structures. Also, the simple minimization procedure for conformation selection may be inappropriate for **22**; the double bond appears in the midst of the unfavorable blue region. As in model A, steroids **30** and **60** also appear as outliers in model B (predicted values: 6.3 and 6.6, respectively). The remaining two outliers in model B are steroids **53** and **63** (predicted values: 7.1 and 6.0, respectively). Steroid **53** is overpredicted by 2.5 $\log(\text{ED}_{50})$ units; **63** is underpredicted by 3 units. Both models A and B generally underestimate the depression of ligand efficacy in 2-deoxy steroids such as **53**, which also, incidentally, lacks the 22-OH group (cf. steroid **15**). Rapisterone D (**63**) is surprisingly active and has an unusual inversion of configuration at both C-2 and C-3.

Other structures which have significant underpredictions but which we chose to retain (model A: **32**; model B: **1**, **11**, **22**) may suffer from unsuitable conformations at the chain terminus, bearing oxygen functionality or unsaturation located in the large blue region along the side chain (Figures 3 and 4). In model

A, **49** and **51** are overpredicted. Steroid **51** lacks the 2-OH group; comments similar to those for **53** apply.

Test set. Predictions were made for a group of seven test steroids, which were analyzed without reference to the predicted values in the *Drosophila* B_{II} assay. The results appear in Table 8. Steroids **72** and **73** are reasonably well-predicted. Steroid **74** is underestimated partly because both models interpret the 5-hydroxyl group to be detrimental. For structures **75** and **76**, slight hydrolysis to much more active steroids may be responsible for artificially elevating the observed binding coefficients. The low activity of rubrosterone (**77**) is accurately assessed by both models. The overprediction of steroid **78** may be a result of a conformation which places the acetyl carbonyl group in alignment with the C-22-OH dipole of the training set steroids. This conformational selection may be biased.

Inactives. Steroids which have been tested in the B_{II} assay but which are inactive are listed in Table 9. In general, both models overestimate the activity of these compounds, which include homobrassinolides, steroids with inversion at C-2 and C-3, or steroids with unfunctionalized side chains. Note that the inactive steroids usually lack at least three pharmacophore elements (vide infra), or otherwise bear a significant alteration to the steroid core. Steroids **82** and **83** are notable examples.

Notable ecdysteroids from the literature. Several notable ecdysteroids which have been reported in the literature but which were not available for the training set for this model deserve comment. Table 10 depicts the structures and predicted activities. Pinnasterol (**93**) and pinnasterol-2-acetate (**94**) are significantly active in the *Sarcophaga* assay [22]. These steroids do not bear the *cis*-A/B-ring junction, have inverted stereochemistry at C-3, and lack the 14-OH group. Predictions from both models A and B affirm high activity. These structures indicate that the bridgehead double bond at C-5 provides an acceptable disposition of the 2- and 3-hydroxyl groups, and along with structures **68–71** suggest that there may be a number of alternative ring systems which display the ecdysteroidal pharmacophore in the appropriate geometry.

14-Deoxymuristerone A (**95**) was observed to be 70–80 times more active in the standard *Drosophila* Kc cell assay compared to 20E [21]. Models A and B place this structure among the most active, but do not predict the superlative activity measured by

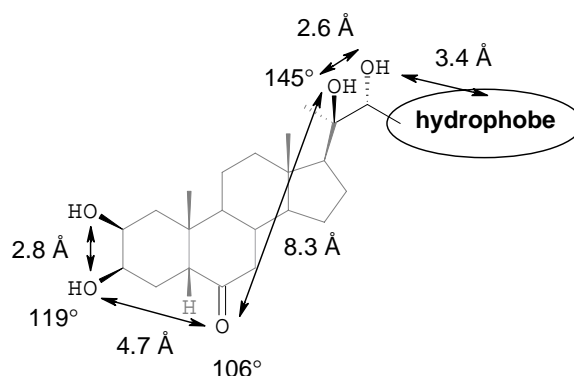


Figure 5. Pharmacophore hypothesis; dimensions are based on 20-hydroxyecdysone.

Cherbas and Nakanishi. This may be due to the fact that both models interpret the 5-hydroxyl group as being detrimental, when in fact this group is sometimes attenuating and sometimes accentuating.

S and *R*-26-Iodoponasterone A (**96** and **97**, respectively) are twice as active as ponasterone A (16 times more active than 20E) in the Kc cell assay [22]. Neither of our models accurately predicts the unusually high binding observed for 26-iodoponasterone A, although the model A prediction for **97** is quite high. This may be due to a poor assessment of the steric contribution of the large iodine, or coincidence of the iodine with the positive-charge-favoring region.

Pharmacophore hypothesis. Until now, our knowledge of ecdysteroidal SARs has been fragmentary due to a large number of assay systems that have been studied and a relatively small range of analogs. It has been possible to draw only outline conclusions from assays where the receptor is likely to be the major factor determining activity [23]. These conclusions are that: (i) the A/B ring junction is essential; (ii) the 7-en-6-one moiety is essential; (iii) a full eight-carbon acyclic steroid side chain is essential; (iv) a free 14- α -hydroxyl group is essential; (v) the 2 β -, 3 β -, and 20*R*-hydroxy groups greatly enhance activity; and (vi) C-25 hydroxyl substitution diminishes activity. Exceptions to (i) and (iv) such as pinnasterol (**93**) and 14-deoxymuristerone A (**95**) have generally been regarded as anomalies.

Ecdysteroid structure–activity data may now be couched in the language of a pharmacophore hypothesis, due to the quantity and uniformity of data presented here. We graphically summarize our current pharmacophore hypothesis in Figure 5.

Table 7. Functional group contributions to $B_{II} -\log(ED_{50})$

Functional group	$\Delta -\log(ED_{50})$ (steroid 1-steroid 2), ...				Average
+ 2 β -OH	-1.66 (51,1)	-1.94 (52,11)	-3.25 (53,15)		-2.28
3 β -OH \rightarrow 3=O	0.74 (1,62)				0.74
β -C-3 \rightarrow α -C-3	0.0(57,3)				0.0
+ 5 β -OH	-0.88 (11,30)	1.51 (9,31)	0.39 (21,32)	-0.38 (20,33)	0.25
	0.44 (25,34)	1.03 (26,35)	-0.40 (44,45)	0.32 (50,48)	
6=O \rightarrow 6-OH (α or β)	1.35 (11,64)	2.42 (11,65)			1.89
+ 11 β -OH	2.02 (11,46)	0.34 (31,45)	2.25 (9,44)	2.88 (22,47)	1.87
+ 14 α -OH	-0.6 (66,11)				-0.60
20R-C \rightarrow 20S-C	1.96 (1,39)	1.83 (2,38)	1.36 (3,40)		1.72
+ 20-OH	-2.16 (1,11)	0.58 (2,12)	-0.56 (8,24)	-1.51 (4,9)	-1.35
	-2.55 (5,15)	-1.88 (51,52)			
+ 22R-OH	-0.66 (5,1)	-0.27 (15,11)			-0.47
22R-OH \rightarrow 22S-OH	0.60 (1,3)				0.60
22-OH \rightarrow 22=O	-1.39 (1,2)	1.35 (11,12)			-0.04
+ 22-OH, + 25-OH	0.34 (6,1)				0.34
+ 24-CH ₃ (R or S)	1.46 (11,18)	0.23 (11,19)			0.85
+ 25-OH	2.04 (4,1)	1.39 (9,11)	-1.0 (31,30)	1.16 (44,46)	0.89
25-OH \rightarrow 24/25 C=C	0.27 (11,21)	1.54 (30,32)			0.91
<i>trans</i> \rightarrow <i>cis</i> C/D junction	1.44 (66,71)				1.44

Our hypothesis is described as follows. Steroid receptor binding is the result of the incremental contributions of a number of molecular features, each of which may contribute several $-\log(ED_{50})$ units to binding. These features include heteroatoms at C-2, C-3, C-20, a large dipole at C-6, and a moderately bulky hydrophobic group attached to C-22, all at approximately the distances and angles represented by 20E (Figure 5). The heteroatoms are presently identified as oxygen, but other heteroatoms may be anticipated in the future as well. As yet, there is no clear requirement that any of this functionality be capable of hydrogen-bond donation. We anticipate that ecdysone activity will be very sensitive to small changes in geometry around the C-20 and C-22 dipoles. The bond order for the heteroatoms may be changed. One or more of these six pharmacophore elements may be absent without loss, albeit possible depression, of receptor activation.

A corollary to this hypothesis is that substantial substituent and ring junction variations in the steroid scaffold are tolerated as illustrated by the training set members in the final section of Table 1, in particular steroids **67–71**. Specifically, the A/B ring junction may bear a bridgehead double bond (**68**) or be present in a *trans*-configuration (**67**). Unsaturation at the 8/9 and 14/15 positions (**69**) as well as the 9/11 position

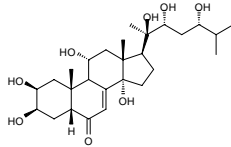
(**70**) is well tolerated. Conversion of the 6-keto group to a 6-hydroxy group (either epimer) results in depression, but not loss, of activity (**64**, **65**). Even the C/D ring junction may be present in the *cis*-configuration (**71**). There are clear limitations, however, as illustrated by the complete loss of activity when the B-ring is expanded by one atom, as exemplified by homobrassinolides **91** and **92**. It remains to be seen if a completely different scaffold bearing several or all of the six pharmacophoric elements may suitably replace the steroid nucleus.

Another aspect of this pharmacophore hypothesis is the permissibility of additional functionality, chain branching, and ring fusion. Examples of additional alkylation and hydroxylation, fluorination, chain cyclization, and acetyl and glycosidic derivatizations are all found among active molecules in the training set (Tables 1 and 2) and test set (Table 8).

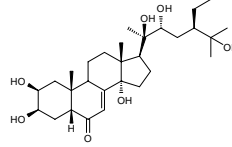
Previous ecdysteroid SAR concepts, expressed in terms of essential or nonessential functional groups, should now be modified. We base the following conclusions on both measured activities in the B_{II} assay as well as predictions for hypothetical steroids (Table 11). First, a *cis*-A/B-ring junction is not essential for activity, cf. **67** and **68**. We predict that 5 α -ponasterone A (**98**) should have a $-\log(ED_{50})$ of

Table 8. Test set: predicted and actual *Drosophila* B_{II} cell line $-\log(\text{ED}_{50})$ values

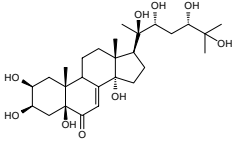
No.	Steroid	Actual	Predicted	
			Model A	Model B
72	Punisterone	6.08	6.99	6.20
73	Makisterone C	6.70	6.33	7.57
74	5 β -Hydroxyabutasterone	7.64	6.49	5.72
75	20-Hydroxyecdysone-3 β -D-xylopyranoside	5.8	24.48	5.70
76	Ponasterone A-3 β -D-xylopyranoside	5.80	4.99	5.03
77	Rubrosterone	<4	4.30	4.46
78	Dihydorubrosterone-17-acetate	<4	6.18	6.48



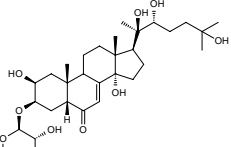
72



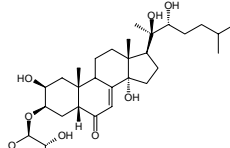
73



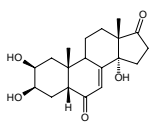
74



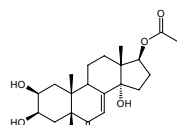
75



76



77



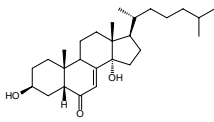
78

about 5. Second, a 7-en-6-one moiety is not essential, although absence of *both* the double bond and the keto group does seem to eliminate activity (cf. Table 9). For the hypothetical steroid 6-deoxyponasterone A (**99**), we predict a $-\log(\text{ED}_{50})$ value of 6.9 and 8.3 for models A and B, respectively (Table 11). This translates to a group value of about 2 $\log(\text{ED}_{50})$ units for the 6-keto group. For 7,8-dehydroponasterone A (**100**), we predict a $-\log(\text{ED}_{50})$ value of 8.4 and 9.2 for models A and B respectively, indicating a group value of about 0.7 $\log(\text{ED}_{50})$ units. Steroids with a hydroxyl group at the 6 position (**64** and **65**) are weakly or moderately active. Third, a full eight-carbon acyclic steroid side chain is not essential. In our assay, posterone (**58**), which has only a vestigial acetyl group at C-17, has a $-\log(\text{ED}_{50})$ of 5.3. Steroids **101** and **102** are predicted to be active. Additionally, numerous cyclic substituents at C-17 are fairly active. Although steroids with these substituents may contain a full

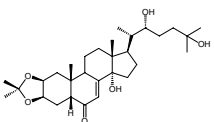
eight carbons, they should by no means be considered full chain analogs, since the conformational space available is quite limited compared to their acyclic counterparts. Fourth, a 14- α -hydroxyl group is not essential. This was perhaps anticipated with the report of 14-deoxymuristerone A. The 14-hydroxyl group may be absent with (**71**) or without inversion (**66**), or it may be replaced with unsaturation (**69**). We predict for the hypothetical steroids 14-deoxyponasterone A (**103**) and 5-hydroxy-14-deoxyponasterone A (**104**) a $-\log(\text{ED}_{50})$ value of 8.7 and 8.9, respectively (model A). A number of inactive steroids lack the 14-hydroxyl group, but these have other detrimental features as well. We estimate the contribution of the 14-hydroxyl group to be about 0.6 $\log(\text{ED}_{50})$ units. Fifth, we concur with previous analyses that the 2 β -, 3 β - and 20*R*-hydroxy groups, although not essential, greatly enhance activity. From measured activities, we evaluate the 2 β -hydroxyl group to be worth approximately 2

Table 9. Inactive ecdysteroids: predicted and actual *Drosophila* B_{II} cell line $-\log(\text{ED}_{50})$ values

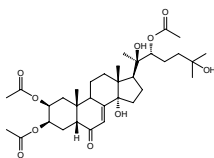
No.	Steroid	Actual	Predicted	
			Model A	Model B
79	Ketodiol	≤ 5.2	5.53	6.47
80	Ecdysone-2,3-acetonide	≤ 4.5	7.60	6.59
81	20-Hydroxyecdysone-2,3,22-triacetate	< 4.8	6.02	6.74
82	2 β ,3 β ,6 α -Trihydroxy-5 β -cholestane	≤ 4.6	4.90	7.76
83	2 β ,3 β ,6 β -Trihydroxy-5 β -cholestane	≤ 4.6	5.57	8.14
84	2,14,22,25-Tetradecoxy-5 α -cdysone	≤ 4.6	5.82	6.95
85	5 α -Ketodiol	≤ 4.6	4.22	5.66
86	2 α ,3 α ,22S,25-Tetrahydroxy-5 α -cholestan-6-one	< 4	5.62	5.76
87	Castasterone	< 3	4.67	5.39
88	24- <i>epi</i> -Castasterone	< 4	5.53	6.78
89	Bombycosterol	≤ 4.6	4.69	6.06
90	Calonysterone	neg	7.37	7.77
91	<i>iso</i> -Homobrassinolide	< 5	7.09	5.52
92	28-Homobrassinolide	< 4	5.99	6.00



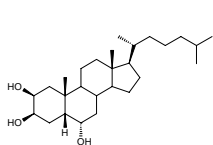
79



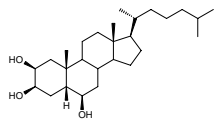
80



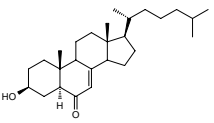
81



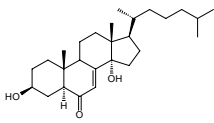
82



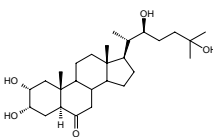
83



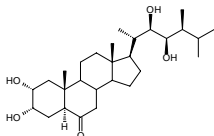
84



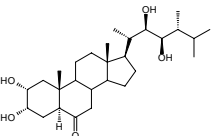
85



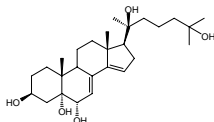
86



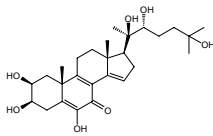
87



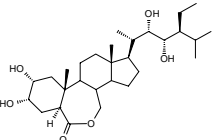
88



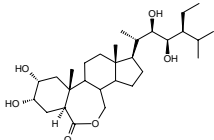
89



90



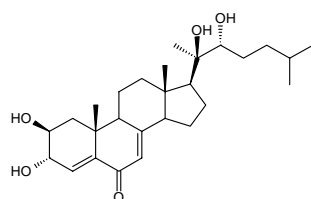
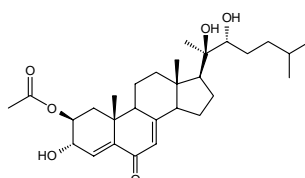
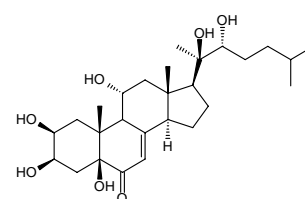
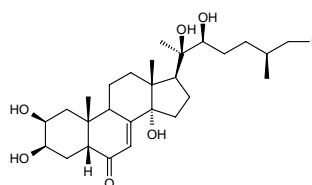
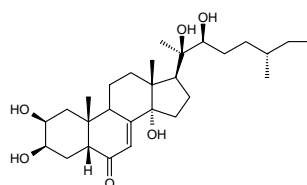
91



92

Table 10. Anomalous and notable literature steroids: predicted *Drosophila* B_{II} cell line $-\log(\text{ED}_{50})$ values

No.	Steroid	Actual	Predicted	
			Model A	Model B
93	Pinnasterol	See text	8.09	7.43
94	Pinnasterol-2-acetate	See text	8.02	7.66
95	14-Deoxymuristerone A	70–80 \times^a	8.54	8.30
96	(25 <i>S</i>)-26-Iodoponasterone A	16 \times^a	6.05	7.25
97	(25 <i>R</i>)-26-Iodoponasterone A	16 \times^a	6.54	8.23

^aKc cell assay; relative to 20E.**93****94****95****96****97**

$\log(\text{ED}_{50})$ units (Table 7). We predict the 3 β -hydroxyl group to be worth about 0.5–1 $\log(\text{ED}_{50})$ units (Table 11), and we estimate the 20*R*-hydroxyl group to be worth 1.3 $\log(\text{ED}_{50})$ units on average (Table 7). Hydroxy substitution at C-25 can be either beneficial or detrimental, but on average, diminishes activity by approximately 0.9 $\log(\text{ED}_{50})$ units. The 22*R*-hydroxy group enhances activity by ca. 0.5 $\log(\text{ED}_{50})$ units, but inversion at C-22 depresses activity. Oxidation to the 22-ketone seems to have a variable effect experimentally (Table 7); both models interpret such oxidation to be favorable when comparing training set structures, but detrimental when considering hypothetical structures **106–108**.

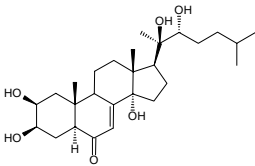
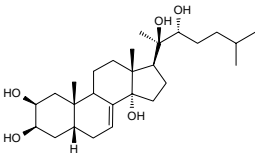
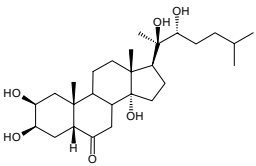
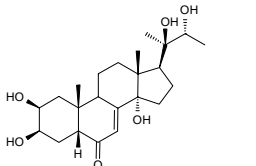
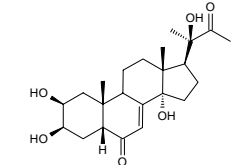
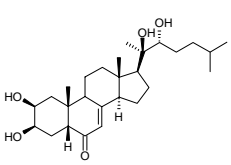
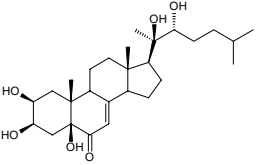
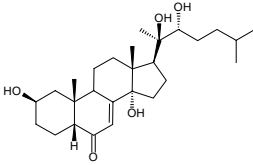
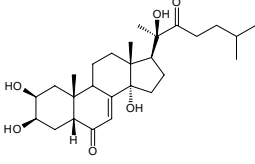
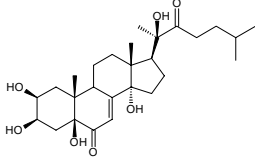
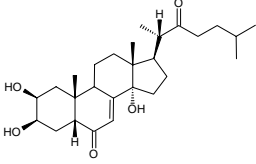
What is the bioactive conformation for ecdysone activity? Although this CoMFA analysis does not directly address this question, the steroid conformations used here resemble the Huber–Hoppe crystal structure of ecdysone. It is possible to say that this general chain

conformation together with a chair conformation for the A-ring is not unreasonable. Preliminary CoMFA studies using the active analog approach on a subset of this training set continue to suggest that the bioactive conformation resembles the ecdysone crystal structure. However, clearer, more confident hypotheses must await a comprehensive study.

Non-steroidal diacylhydrazines are agonists of ecdysone in *Drosophila* cell-based assays such as the B_{II} cell line used in this study [24]. It is conceivable that molecules of this structural class, although not congeneric to the steroids, may possibly bind at the ecdysteroid receptor complex at a common or overlapping locus with that of the steroids. If so, CoMFA analysis on a mixed training set of steroids and diacylhydrazines could reveal the nature of this overlap. Indeed, Nakagawa and co-workers [25] have conducted a CoMFA analysis of this type. We may anticipate an analogous study on this training set with use of the

Table 11. Examination of pharmacophore hypothesis: predicted *Drosophila* B_{II} cell line $-\log(\text{ED}_{50})$ values and structures of hypothetical steroids

No.	Steroid	Predicted	
		Model A	Model B
98	5 α -Ponasterone A	4.37	5.13
99	6-Deoxyponasterone A	6.87	8.33
100	7,8-Dehydroponasterone A	8.40	9.20
101	2 β ,3 β ,14 α ,20R,22R-Pentahydroxy-24-nor-5 β -cholan-7-en-6-one	7.09	6.70
102	2 β ,3 β ,14 α ,20R-Tetrahydroxy-24-nor-5 β -cholan-7-ene-6,22-dione	6.08	6.21
103	14-Deoxyponasterone A	8.74	8.58
104	5-Hydroxy-14-deoxyponasterone A	8.85	8.14
105	3-Deoxyponasterone A	8.01	8.91
106	22-Dehydroponasterone A	7.37	7.47
107	5-Hydroxy-22-dehydroponasterone A	7.16	6.97
108	20-Deoxy-22-dehydroponasterone A	7.41	7.39

			
98	99	100	101
			
102	103	104	105
			
106	107	108	

active-analog approach or some other pharmacophore-based mapping technique [26] towards the goal of clarifying the steroid bioactive conformation.

Distinct binding events associated with different affinities at multiple receptor types are not addressed by this study. Although there are at least three ecdysteroid receptor proteins in *Drosophila*, it is however doubtful that all three isoforms are present in the B_{II} cell line. Western blotting shows the presence of only

one band at 105 kDa. Correspondingly, only one band is observed for USP, the partner to EcR. Furthermore, all isoforms have the same ligand-binding domain. Additionally, our dose-response curves do not appear to detect multiple affinities for any single ecdysteroid. Despite this evidence, improvements in ecdysteroid QSAR modeling may benefit from unequivocally homogeneous receptor preparations.

Conclusions

In conclusion, an expansive collection of ecdysteroids have been tested in a uniform assay based on the *Drosophila* B_{II} cell line. A satisfactory CoMFA analysis using region focusing on indicator electrostatic and steric fields was obtained on these data even when the training set is aligned on the basis of simple analogy to a reference ecdysone crystal structure. The electrostatic field is most significant with notable regions located at O-20 (negative charge enhancing), H-25/C-26 (negative charge enhancing), between O-2 and O-3 (positive charge enhancing), and encircling the distal portion of the side chain (positive charge enhancing). Important steric regions are located at various points along the periphery of the side chain (steric bulk detracting) and near C-11 (steric bulk enhancing). Incorporation of MlogP as an independent variable improves the steric/electrostatic model.

Satisfactory results are achieved when predicting activities for an ecdysteroid test set and calculating weak binding constants for steroids known to be inactive.

Based on measured activity values and predicted values for hypothetical steroids, functional group additivity values are assigned to various important molecular features. Ecdysteroid activity can be understood in terms of the summation of the incremental contributions of at least six molecular features: heteroatom substitution at C-2, C-3, C-6, C-20, C-22, and a hydrophobic group adjoining C-22. No single feature is essential for activity, as has been previously described in earlier qualitative SAR descriptions.

Improvements in QSAR analysis of the ecdysteroids may be possible using pharmacophore-based conformational selection and alignment, or alternative QSAR methods of current interest. More fundamental improvements may be attained should biological activity data be obtained in assays which distinguish among possible distinct binding loci.

Acknowledgements

We wish to thank Umirzak Baltaev (Institute of Petrochemistry and Catalysis, Ufa, Russia), Francisco Camps, Josefina Casas and Josep Coll (CSIC, Barcelona, Spain), Juraj Harmatha (Institute of Organic Chemistry and Biochemistry, Prague, Czech Republic), Toshio Honda (Hoshi University, Japan), Dennis Horn (Melbourne, Australia), Rene Lafont (De-

partment of Biochemistry, École Normale Supérieure, Paris, France), Ted Molinski (University of California, Davis, CA, U.S.A.), Patrick Roussel (Department of Chemistry, University of Exeter, U.K.), Satyajit Sarker (Department of Biological Sciences, University of Exeter, U.K.), Karel Sláma (Institute of Entomology, České Budějovice, Czech Republic), and Apichart Suksamrarn (Department of Chemistry, Ramkhamhaeng University, Bangkok, Thailand) for their generous supply of ecdysteroid specimens. We are grateful to Pensri Whiting for excellent technical assistance.

References

- Koolman, J. (Ed.) Ecdysone: From Chemistry to Mode of Action, Thieme, Stuttgart, 1989.
- Lafont, R.D. and Wilson, I.D. (Eds.) The Ecdysone Handbook, 2nd ed., The Chromatographic Society, Nottingham, 1996.
 - Lafont, R., Arch. Insect Biochem. Physiol., 35 (1997) 3.
 - Horn, D.H.S. and Bergamasco, R., In Kerkut, G.A. and Gilbert, L.I. (Eds.) Comprehensive Insect Physiology Biochemistry and Pharmacology, Vol. 7, Pergamon, Oxford, 1985, pp. 185–248.
- Cherbas, P. and Cherbas, L., In Gilbert, L.I., Tata, J.R. and Atkinson, B.G. (Eds.) Metamorphosis, Academic Press, New York, NY, 1996, pp. 175–221.
 - Koelle, M.R., Talbot, W.S., Segraves, W.A., Bender, M.T., Cherbas, P. and Hogness, D.S., Cell, 67 (1991) 59.
 - Talbot, W.S., Swyryd, E.A. and Hogness, D.S., Cell, 73 (1993) 1323.
 - Yao, T.-P., Segraves, W.A., Oro, A.E., McKeown, M. and Evans, R.M., Cell, 71 (1992) 63.
 - Yao, T.-P., Forman, B.M., Jiang, Z., Cherbas, L., McKeown, M., Cherbas, P. and Evans, R.M., Nature, 366 (1993) 476.
- Jones, G. and Sharp, P.A., Proc. Natl. Acad. Sci. USA, 94 (1997) 13499.
- Clément, C.Y., Bradbrook, D.A., Lafont, R. and Dinan, L., Insect Biochem. Mol. Biol., 187 (1993) 23.
- Cramer III, R.D., Patterson, D.E. and Bunce, J.D., J. Am. Chem. Soc., 10 (1988) 5959.
 - Cramer III, R.D., Depriest, S., Patterson, D. and Hecht, P., In Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications, ESCOM, Leiden, 1993, pp. 443–485.
 - Green, S.M. and Marshall, G.R., Trends Pharmacol. Sci., 16 (1995) 285.
- Butenandt, A. and Karlson, P., Z. Naturforsch., 9b (1954) 389.
- Alternative methods to CoMFA which are suitable for noncongeneric structures are described in
 - Kier, L.B., Molecular Connectivity in Structure–Activity Analysis, Wiley, New York, NY, 1986.
 - Hopfinger, A., In Charifson, P.S. (Ed.) Practical Application of Computer-Aided Drug Design, Marcel Dekker, New York, NY, 1997, p. 105.
- SYBYL 6.3, Tripos Associates, St. Louis, MO, U.S.A.
- Huber, R. and Hoppe, W., Chem. Ber., 98 (1965) 2403.

11. Clark, M., Cramer III, R.D. and van Opdenbosch, N., *J. Comput. Chem.*, 10 (1989) 982.
12. a. Sigma charges are calculated by the Gasteiger method: Gasteiger, J. and Marsili, M., *Tetrahedron*, 36 (1980) 3219.
b. Pi charges are calculated by the Hückel method: Tripos Associates, St. Louis, MO, U.S.A.
13. a. SYBYL Ligand-Based Design Manual, v. 6.3, 1996, p. 64.
b. Kroemer, R.T. and Hecht, P., *J. Comput.-Aided Mol. Design*, 9 (1995) 205.
c. Bohacek, R.S. and McMartin, C., *J. Med. Chem.*, 35 (1992) 1671.
14. a. Moriguchi, I., Hirono, S., Liu, Q., Nakagome, I. and Matsushita, Y., *Chem. Pharm. Bull.*, 40 (1992) 127.
b. MlogP implemented in SYBYL programming language by Blake, J., Pfizer Inc., v. 1.2, 8 November, 1994.
15. Dunn, W.J., Wold, S., Edlund, V. and Helberg, S., *Quant. Struct.-Act. Relatsh.*, 3 (1984) 131.
16. Bush, B.L. and Nachbar, R.B., *J. Comput.-Aided Mol. Design*, 7 (1993) 587.
17. SYBYL Ligand-Based Design Manual, v. 6.3, 1996, p. 78. Weighting by (standard deviation) \times (coefficient) is tantamount to image enhancement.
18. a. Hopfinger, A.J., Wang, S., Tokarski, J.S., Jin, B., Albuquerque, M., Madhav, P.J. and Duraiswami, C., *J. Am. Chem. Soc.*, 119 (1997) 10509.
b. Cho, S.J. and Tropsha, A., *J. Med. Chem.*, 38 (1995) 1060.
c. Norinder, U., *J. Chemometr.*, 10 (1996) 95.
d. Pastor, M., Cruciani, G. and Clementi, S., *J. Med. Chem.*, 40 (1997) 1455.
19. a. Tropsha, A., Cho, S.J. and Zheng, W., In *Rational Drug Design*, ACS Symposium Series, American Chemical Society, Washington, DC, 1999, in press.
b. Zheng, W. and Tropsha, A., *J. Chem. Inf. Comput. Sci.* (1999), in press.
20. Kukuzawa, A., Kumagai, Y., Masamune, T., Furusaki, A., Katayama, C. and Matsumoto, T., *Tetrahedron Lett.*, 22 (1981) 4085.
21. Cherbas, P., Trainor, D.A., Stonard, R.J. and Nakanishi, K., *J. Chem. Soc., Chem. Commun.*, (1982) 1307.
22. a. Cherbas, P., Cherbas, L., Lee, S.-S. and Nakanishi, K., *Proc. Natl. Acad. Sci. USA*, 85 (1988) 2096.
b. Lee, S.-S., Nakanishi, K. and Cherbas, P., *J. Chem. Soc., Chem. Commun.*, (1991) 51.
23. Dinan, L., In Koolman, J. (Ed.) *Ecdysone: From Chemistry to Mode of Action*, Thieme, Stuttgart, 1989, pp. 345–354.
24. a. Dhadialla, T.S., Carlson, G.R. and Le, D.P., *Annu. Rev. Entomol.*, 43 (1998) 545.
b. Wing, K.D., *Science*, 241 (1988) 467.
25. a. Nakagawa, Y., Shimizu, B.-I., Oikawa, N., Akamatsu, M., Nishimura, K., Kurihara, N., Ueno, T. and Fujita, T., In Hansch, C. and Fujita, T. (Eds.) *Classical and Three-Dimensional QSAR in Agrochemistry*, ACS Symposium Series 606, American Chemical Society, Washington, DC, 1995, pp. 288–301.
b. Shimizu, B.-I., Nakagawa, Y., Hattori, K., Nishimura, K., Kurihara, N. and Ueno, T., *Steroids*, 62 (1997) 638.
c. Nakagawa, Y., Hattori, K., Shimizu, B., Akamatsu, M., Miyagawa, H. and Ueno, T., *Pestic. Sci.*, 53 (1998) 267.
26. a. Mayer, D., Naylor, C.B., Motoc, I. and Marshall, G.R., *J. Comput.-Aided Mol. Design*, 1 (1987) 3.
b. Martin, Y.C., In Hansch, C. and Fujita, T. (Eds.) *Classical and Three-Dimensional QSAR in Agrochemistry*, ACS Symposium Series 606, American Chemical Society, Washington, DC, 1995, pp. 318–329.