Toward a Class-Independent Quantitative Structure−Activity Relationship Model for Uncouplers of Oxidative Phosphorylation

Simon Spycher,*,† Pavel Smejtek,‡ Tatiana I. Netzeva,§ and Beate I. Escher†

Department of Environmental Toxicology, UTOX, Swiss Federal Institute of Aquatic Science and Technology, EAWAG, CH-8600 Dübendorf, Switzerland, Department of Physics, Portland State University, Portland, Oregon 97207, and European Chemicals Bureau, Institute for Health and Consumer Protection, Joint Research Centre, 21020 Ispra (VA), Italy

Received November 2, 2007

A mechanistically based quantitative structure–activity relationship (QSAR) for the uncoupling activity of weak organic acids has been derived. The analysis of earlier experimental studies suggested that the limiting step in the uncoupling process is the rate with which anions can cross the membrane and that this rate is determined by the height of the energy barrier encountered in the hydrophobic membrane core. We use this mechanistic understanding to develop a predictive model for uncoupling. The translocation rate constants of anions correlate well with the free energy difference between the energy well and the energy barrier, $\Delta G_{\text{well-barrier,A}^-}$, in the membrane calculated by a novel approach to describe internal partitioning in the membrane. An existing data set of 21 phenols measured in an in vitro test system specific for uncouplers was extended by 14 highly diverse compounds. A simple regression model based on the experimental membrane−water partition coefficient and $\Delta G_{\text{well-barrier,A}^-}$ showed good predictive power and had meaningful regression coefficients. To establish uncoupler QSARs independent of chemical class, it is necessary to calculate the descriptors for the charged species, as the analogous descriptors of the neutral species showed almost no correlation with the translocation rate constants of anions. The substitution of experimental with calculated partition coefficients resulted in a decrease of the model fit. A particular strength of the current model is the accurate calculation of excess toxicity, which makes it a suitable tool for database screening. The applicability domain, limitations of the model, and ideas for future research are critically discussed.

Introduction

The knowledge about the molecular mechanisms underlying different modes of toxic action (MOA) varies considerably. This has consequences for the establishment of quantitative structure−activity relationships (QSARs). If the mechanisms are well-understood, it is possible to search for suitable descriptors in a focused way, and the applicability domain of the model is easier to define. If the mechanisms are poorly understood, a more statistical approach is necessary, that is, larger numbers of descriptors need to be calculated and descriptor selection methods have to be applied. This makes the definition of the applicability domain more difficult, and the interpretation of the selected descriptors can become nontransparent. Of course, in many cases, a statistical approach is unavoidable and this is also reflected in the OECD report on the principles for the validation of (Q)SARs, which states that for regulatory purposes QSARs should have "a mechanistic interpretation, if possible" (*1*).

The MOA of uncoupling of oxidative and photophosphorylation is an example of an MOA with a relatively well-understood mechanism (*2*–*4*). Therefore, taking advantage of this knowledge should facilitate the descriptor search, the definition of applicability domain, and the mechanistic interpretation of the QSAR model.

The basis for the mechanistic explanation of the uncoupling effect directly stems from Mitchell's chemiosmotic theory (*5*). The chemiosmotic theory explains how ATP is produced from inorganic phosphate and ADP by ATP synthase (ATPase) in energy-transducing membranes, that is, how the phosphorylation is *coupled* to electron transport and hydrogen transfer. As a "side effect", this theory also explains the *uncoupling* effect observed for certain weak acids. Figure 1a gives a simplified schematic picture of the uncoupling effect.

A precondition for the efficient functioning of energy-transducing membranes is that they are proton-impermeable; that is, they need to maintain a pH gradient. A weak acid can act as a proton shuttle transporting protons across mitochondrial or photosynthetic membranes. Thereby they dissipate the pH gradient across the membrane by providing an alternate entry path for protons, which leads to the uncoupling of the phosphorylation from the electron transport and hydrogen transfer. The most potent uncouplers are so efficient that the ATP production is completely uncoupled at toxicant concentrations, which are up to 20 times below the number of ATPases (*3*). This observation, among others, led to the widespread agreement that these toxicants do not specifically inhibit a receptor of energy-transducing membranes but rather act as catalysts according to the protonophoric shuttle mechanism illustrated in Figure 1a. It should be mentioned that other authors have a broader definition of uncoupling and define any energy-dissipating process competing for energy with routine functions of energy-transducing membranes as uncoupling (*4*). In this study, however, the narrower definition of weak acids acting as proton shuttles is used for uncouplers, which are sometimes

* To whom correspondence should be addressed. Tel: +41-44-823-5068. Fax: +41-44-823-5311. E-mail: simon_lukas.spycher@alumni.ethz.ch.
† Swiss Federal Institute of Aquatic Science and Technology.
‡ Department of Physics, Portland State University.
§ European Chemicals Bureau.

10.1021/tx700391f CCC: $40.75
Published on Web 03/22/2008

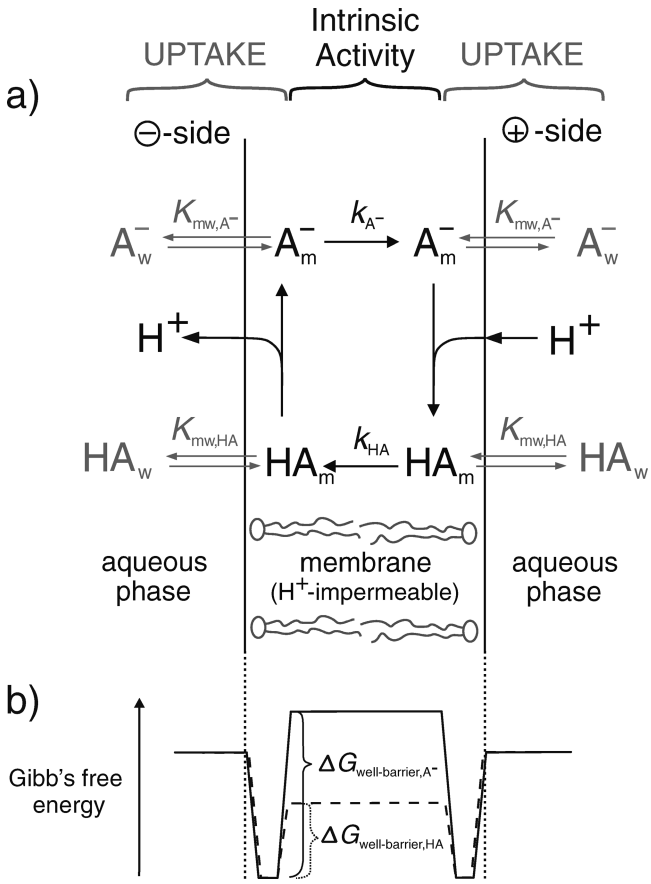

Figure 1. (a) Uncoupling process of weak acids acting as uncouplers in energy-transducing membranes. By diffusion, weak acids (HA) can transport protons from the + side to the – side, where they release a proton. Driven by the membrane potential, the anions (A^-) migrate back and can pick up another proton. The translocation rate constants, k_{HA} and k_{A^-}, determine the speed with which a compound passes through this cycle, that is, the intrinsic uncoupling activity (black letters). The gray letters indicate the partitioning constants of the uptake into the membrane and will be discussed in detail in the section on mechanisms. The gray-shaded circles indicate the polar head groups of the phosphatidylcholine molecules, while the curly lines indicate the fatty acid alkyl chains. (b) Relative change in the Gibb's free energy of the anion and the neutral species across the membrane bilayer for the uncoupler carbonyl cyanide *m*-chlorophenylhydrazone, CCCP (energy schemes adapted from refs *2* and *19*). Note that in the case of CCCP, the depth of the energy well is the same for the neutral and the charged species, respectively. This is an exception, as for most uncouplers the neutral species shows a deeper energy well than the charged species and thus a larger membrane−water partition coefficient.

also referred to as protonophoric uncouplers (*6*) or proton ionophores (*2*).

A large number of QSARs for weak acids acting as uncouplers have been published. The models based on data from specific mitochondrial tests (models based on in vivo data are omitted for brevity) cover a wide range of chemical classes ranging from phenols (*7–11*) to compounds with NH acidic groups like benzimidazoles (*12*), diarylamines (*13*), salicylanilides (*14*), and pyrroles (*15*). All of these QSARs are based on a descriptor for lipophilicity quantifying the uptake into the membrane and in most cases on an additional descriptor for the acidity of the toxicants. Data sets including compounds with bulky ortho substituents also required steric descriptors to describe the shielding of the negative charge on the oxygen or nitrogen atom, respectively, of the anion (*8*). Lipophilicity has been described by the octanol–water partitioning coefficient, log P_{ow} (or log K_{ow} in environmental sciences), or the membrane−water partitioning coefficient, log K_{mw}, measured with lipid bilayer vesicles, so-called liposomes (*9*). The regression coefficients in the above-mentioned studies varied widely. The consequence is that each QSAR can only be used within a single chemical class or even only for subsets of compounds within a chemical class.

Twenty to 30 years ago, the explanation of the mechanism for uncouplers was still controversial. To defend the chemiosmotic theory, a whole series of experimental studies with artificial bilayer membranes were published at that time (*16–20*). The development of charge-pulse (*21*) and voltage-clamp (*17*) experiments with black lipid bilayer membranes (BLM) allowed the detailed examination of the transport of weak acids. In charge-pulse studies, the membrane capacitance was charged to an initial voltage by an intense current pulse, and the subsequent decay of the potential across the membranes was measured as a function of time. In voltage-clamp studies, a constant potential was applied across the membrane and relaxations of the current were studied after the decay of the capacitance spike. While the analysis of the experimental data differs, both methods provided identical kinetic information. The uncoupling process could be described by transport models with five experimentally determined physicochemical descriptors shown in Figure 1a: acid dissociation constant, pK_a, lipid water partition coefficients of the weak acid, $K_{mw,HA}$, and its anion K_{mw,A^-}, and the translocation rate constants of the weak acid, k_{HA}, and its anion k_{A^-} (*18, 19*). With these descriptors, the transport of protons from the + side of the membrane to – side of the membrane, that is, the uncoupling process shown in Figure 1a, could be fully described and quantified. The outstanding feature of these studies was that the values of these descriptors obtained by fitting the differential equations of the transport model could be confirmed by alternative experimental methods; for example, the value of the two partition coefficients was additionally checked by equilibrium dialysis, and the values were in agreement by a factor of 2 or less.

Escher et al. studied the effect of phenolic uncouplers on chromatophores, which are isolated membrane vesicles from the photosynthetic bacteria *Rhodobacter sphaeroides* (*22*). The transport models derived from studies with artificial bilayer membranes could be extended to chromatophores and, thus, to more complex biological membrane systems. In addition, the pH dependence of the uncoupling effect could be comprehensively described by a single transport model. However, so far, neither the studies using BLMs nor the chromatophore studies has had a big impact in the field of QSAR.

The main goal of the present study was, therefore, to capitalize on these earlier experimental studies and to directly model the hitherto experimentally determined descriptors pK_a, $K_{mw,HA}$, K_{mw,A^-}, k_{HA}, and k_{A^-}. If these descriptors can be calculated with sufficient accuracy, it should be possible to model the uncoupling effect of weak acids in a single model regardless of whether the compound is a phenol, a benzimidazole, or of any other chemical class.

Materials and Methods

Chemicals. A total of 14 compounds that showed uncoupling activity (full names and abbreviations are given in Table 1) were obtained from the following sources: DTFB, DPFB, and TBrTFB were kindly provided by Prof. Zygmunt Kazimierczuk, Institute of Chemistry, Agricultural University Warsaw (Poland); TTFB was purchased from TimTec, Inc. (Newark, DE); and the other 10 compounds were purchased from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland). Chemicals were of the highest purity available, typically >98%. Unless impurities were of exceedingly higher

Table 1. Abbreviations, CAS Numbers, Full Names, and Experimental Data of 35 Compounds with the Uncoupling Activity Measured at pH 7 with Kinspec

abbr	CAS no.	compd name	pK_a	log K_{ow}	log $K_{mw,HA}$ (L/(kg_{lipid})	log K_{mw,A^-} (L/kg_{lipid})[a]	log(1/EC_{tot}) (M)	log(1/EC_w) (M)[b]	m_{lip}/V_w (g_{lipid}/L)	train/test[c]
FCCP	370-86-5	carbonyl cyanide *p*-methoxyphenylhydrazone	6.2[d]	3.68[e]	4.22[f]	4.22	7.54	8.57[g]	0.59	0
CCCP	555-60-2	carbonyl cyanide *m*-chlorophenylhydrazone	5.95[h]	3.38[e]	4.05[f]	4.05	7.21	8.09[g]	0.59	1
TTFB	2338-29-6	4,5,6,7-tetrachloro-2-(trifluoromethyl)benzimidazole	5.3[i]	4.73[j]	4.35[i]	4.35	7.16	8.28[g]	0.55	1
DTFB	2338-25-2	5,6-dichloro-2-(trifluoromethyl)benzimidazole	7.3[k]	3.49[e]	3.05[k]	3.05	5.76	5.98[g]	0.59	1
DPFB	102516-93-8	5,6-dichloro-2-(pentafluoroethyl)benzimidazole	ND	ND	ND	ND	6.11	(7.33)[g]	0.59	0
TBrTFB	2338-30-9	4,5,6,7-tetrabromo-2-(trifluoromethyl)benzimidazole	5.8[l]	4.81[e]	ND	ND	6.76	(7.56)[g]	0.59	1
fluazinam	79622-59-6	3-chloro-*N*-[3-chloro-2,6-dinitro-4-(trifluoromethyl)-phenyl]-5-(trifluoromethyl)-2-pyridinamine	7.11[m]	3.42[e,n]	ND	ND	8.75	(9.08)[g]	0.59	1
5NTFB	327-19-5	5-nitro-2-trifluoromethylbenzimidazole	6.7[l]	2.68[e]	ND	ND	5.06	(5.13)[g]	0.59	0
4NTFB	14689-51-1	4-nitro-2-trifluoromethylbenzimidazole	6.8[l]	ND	ND	ND	3.73	(3.75)[g]	0.59	1
6Cl24DNP	946-31-6	6-chloro-2,4-dinitrophenol	2.06[o]	ND	ND	ND	5.40	(5.41)[g]	0.55	0
26DBr4NP	99-28-5	2,6-dibromo-4-nitrophenol	3.39[p]	3.57[e]	ND	ND	5.25	(5.35)[g]	0.55	1
246TriNP	88-89-1	2,4,6-trinitrophenol	0.38[e]	2.03[e]	ND	ND	5.78	(5.79)[g]	0.55	1
warfarin	81-81-2	warfarin	5.00[q]	2.6[e]	3.39[q]	1.4	3.58	3.59[g]	0.55	1
246TriBP	118-79-6	2,4,6-tribromophenol	6.08[r]	4.19[e,r]	ND	ND	4.30	(4.75)[g]	0.59	1
24DCP	120-83-2	2,4-dichlorophenol	7.85[s]	3.23[t]	3.59[s]	2.69	4.04	4.62[u]	0.82	1
34DCP	95-77-2	3,4-dichlorophenol	8.59[s]	3.05[t]	3.76[s]	2.85	4.04	4.79[u]	0.82	1
35DCP	591-35-5	3,5-dichlorophenol	8.26[s]	3.62[e]	3.76[s]	2.85	4.66	5.40[u]	0.82	1
245TriCP	95-95-4	2,4,5-trichlorophenol	6.94[s]	4.19[t]	4.46[s]	2.98	4.95	6.04[u]	0.82	1
246TriCP	88-06-2	2,4,6-trichlorophenol	6.15[s]	3.72[t]	3.99[s]	2.5	3.98	4.33[u]	0.82	0
345TriCP	609-19-8	3,4,5-trichlorophenol	7.73[s]	4.41[t]	4.71[s]	3.16	5.42	6.98[u]	0.82	1
2345TeCP	4901-51-3	2,3,4,5-tetrachlorophenol	6.35[s]	4.87[t]	4.76[s]	3.9	5.98	7.15[u]	0.82	0
2346TeCP	58-90-2	2,3,4,6-tetrachlorophenol	5.4[s]	4.42[t]	4.46[s]	3.46	4.70	5.29[u]	0.82	0
PCP	87-86-5	pentachlorophenol	4.75[s]	5.24[t]	5.1[s]	4.35	5.76	7.07[v]	0.84	1
35DBC	13979-81-2	3,5-dibromo-4-methylphenol	8.28[w]	5.44[w]	4.51[w]	3.18	4.96	6.41[u]	0.87	0
bromox	1689-84-5	3,5-dibromo-4-hydroxy-benzonitrile	4.09[w]	2.97[w]	3.16[w]	2.1	4.85	4.90[u]	0.87	1
4NP	100-02-7	4-nitrophenol	7.08[t]	2.04[t]	2.72[t]	0.95	3.91	4.00[u]	0.82	1
24DNP	51-28-5	2,4-dinitrophenol	3.94[s]	1.67[t]	2.64[s]	1.9	4.12	4.15[u]	0.82	0
26DNP	573-56-8	2,6-dinitrophenol	3.7[s]	1.22[t]	2.03[s]	1.86	3.06	3.08[u]	0.82	1
34DNP	577-71-9	3,4-dinitrophenol	5.48[s]	2.23[x]	3.17[s]	1.9	4.80	4.84[u]	0.82	1
DNOC	534-52-1	2-methyl-4,6-dinitrophenol	4.31[s]	2.12[t]	2.76[s]	2.35	4.68	4.75[u]	0.82	1
DNPC	609-93-8	4-methyl-2,6-dinitrophenol	4.06[s]	ND	2.34[s]	2.26	3.36	3.42[u]	0.82	0
dinoseb	88-85-7	2-*s*-butyl-4,6-dinitrophenol	4.62[s]	3.56[t]	3.96[s]	3.35	5.94	6.40[u]	0.82	0
dino2terb	1420-07-1	2-*tert*-butyl-4,6-dinitrophenol	4.8[s]	ND	4.1[s]	3.59	6.37	7.00[u]	0.82	1
dino4terb	4097-49-8	4-*tert*-butyl-2,6-dinitrophenol	4.11[s]	ND	3.81[s]	3.23	3.90	4.28[u]	0.82	1
triclosan	3380-34-5	5-chloro-2-(2,4-dichlorophenoxy)phenol	8.05[y]	4.76[e]	ND	ND	5.07	(6.73)[y]	0.91	0

[a] Same references as for log Klipw,HA. [b] Values in brackets were determined with the uptake estimated by eqs 2, 4, and 5. [c] Training set (1); test set (0). [d] Ref *18*. [e] PhysProp-Database (http://www.syrres.com/esc/physdemo.htm) or KowWin-Demo-Database (http://www.syrres.com/esc/kowdemo.htm). [f] Ref *20* (defined for a surface of 0.7 nm^2 per lipid molecule and MW 760 g/mol lipid). [g] This study. [h] Ref *19*. [i] Ref *17* (defined for a surface of 0.7 nm^2 per lipid molecule and MW 760 g/mol lipid). [j] Ref *36* (measured D_{ow} and pK_a of 5.3 taken to obtain K_{ow}). [k] Ref *16* (defined for a surface of 0.7 nm2 per lipid molecule and MW 760 g/mol lipid). [l] Ref *37*. [m] Ref *38* [measured in 50% EtOH/50% H_2O (v/v)]. [n] Ref *39* [value for log K_{ow} taken as a mean of refs *e* (3.56) and *n* (3.27)]. [o] Ref *40*. [p] SPARC pK_a-Database (http://ibmlc2.chem.uga.edu/sparc/index.cfm). [q] Ref *41*. [r] Ref *42* [value for log K_{ow} taken as a mean of *e* (4.13) and *r* (4.24)]. [s] Ref *43*. [t] Ref *30*. [u] Ref *35*. [v] Taken as an average value of ref *35* and two older unpublished measurements. [w] Ref *44*. [x] Ref *45*. [y] Unpublished measurements.

uncoupling potency, the small fraction of impurity should not influence the measured effects. Other chemicals used for time-resolved spectroscopy are described in ref *23*.

Determination of Uncoupling Activity. The in vitro test system Kinspec, which was used in this study, works with chromatophores extracted from the photosynthetic bacterium *R. sphaeroides* (*23, 24*). In the Kinspec test, a 1 μs "single turnover" flash of light causes the build-up of a membrane potential, and the subsequent decay of the membrane potential is monitored using time-resolved spectroscopy (*23, 24*). The presence of uncouplers acting as protonophores accelerates the decay of the chromatophore membrane potential, because of their ability to transport protons across the membrane. This allows the determination of the activity of different uncouplers, which is described in detail in the accompanying text to Figure SI-1 in the Supporting Information. The endpoint for the toxic effect of uncoupling in this system (*25*) is the concentration of toxicant needed to induce an observed pseudo-first-order decay rate constant of uncoupling, $k_{uncoupling}$, of 0.5 s^{-1}. The measurements in the Kinspec system are in good agreement with other in vitro tests based on oxygen evolution of isolated mitochondria or submitochondrial particles (*9, 11, 23*) and also with in vivo effect concentrations; however, the number of compounds used for these evaluations is limited.

Toxic Mechanisms: Intrinsic Activity. The key to the understanding and the quantification of the uncoupling mechanism lies in the nature of the proton-impermeable lipid bilayer membranes. The properties of biological membranes differ from bulk solvents like octanol or hexane in several ways. First, the hydrophobicity changes over the axis perpendicular to the membrane plane, that is, over the membrane normal, which can be quantified by the change of the relative dielectric constant, ϵ_r. The relative dielectric constant, ϵ_r, drops from a value of 80 in water to values around 70 on the membrane surface (*26*) to 30 at the interface between polar head groups and hydrocarbon core (*26*) to values around 2 in the hydrophobic core (*19, 27*). This is a direct consequence of the arrangement of the phosphatidylcholine molecules in the bilayer with polar head groups forming the outer side of the membrane and the alkyl chains forming the hydrophobic core. The second difference between bulk (isotropic) solvents and membranes is also a result of the anisotropy of membranes. Membranes have highly ordered regions and regions, which are less ordered. The regions with high order lead to lateral pressures, which affect partition coefficients and permeabilities (*28, 29*). Bulky molecules partition less into these areas than could be expected from the hydrophobicity of this environment.

This arrangement leads to large differences in the Gibb's free energy for a molecule depending on the position in the membrane. In their work on the uncoupler *m*-chlorophenylhydrazone (CCCP), Kasianowicz et al. (*19*) gave a qualitative picture of the free energy of CCCP over the membrane normal, which also applies to other polar hydrophobic compounds. The free energy schemes of Figure 1b illustrate two things: First, a protonophoric uncoupler like CCCP has the lowest energy in the region of the polar head groups; that is, this is the region where such compounds are mainly accommodated. Second, for a given molecule, the energy required to cross the hydrophobic core is substantially higher for the anionic species than for the neutral species. The translocation rate constants of the weak acid, k_{HA}, and its anion k_{A^-} are strongly related to the energy differences between the region of the lowest energy (i.e., the energy well) and the region with the highest energy (i.e., the energy barrier). Efficient uncouplers like CCCP are compounds that keep this energy as low as possible. On the other hand, compounds that have to overcome a very high energy barrier should be inefficient uncouplers or even show no activity at all.

A side note: The term translocation rate constant is common in biophysics but rarely used in QSAR studies, while the related property permeability is commonly used. The difference between the two properties for the situation shown in Figure 1b is that a compound's permeability (with dimensions distance/time) from the membrane−water interface on one side to the other side is the product of the translocation rate constant (with dimension $time^{-1}$) and partition coefficients normalized to the surface (with the dimension of a distance). The permeability commonly modeled in QSAR and in drug design would additionally include the transfer through the unstirred water layer adjacent to the membrane, that is, the permeability from the outer to the inner bulk aqueous phase.

A second insight from the studies with BLMs is that the total concentration of uncoupler in the membrane remains constant during the short duration of a charge-pulse experiment (*18*). We assume that the same is valid for the chromatophores used in the present study; that is, the shuttle mechanism disturbs the equilibrium only slightly over the 100 ms of an experiment. Furthermore, we assume that prior to the experiment the equilibrium between the aqueous phase and the membrane has been established. This assumption greatly simplifies the modeling.

Toxic Mechanisms: Description of the Uptake of Weak Acids. Partition coefficients between biological membranes and water can be measured, but such measurements are expensive and the results would depend on the composition of the membrane. Good experimental approximations to real membranes are liposomes, as their partition coefficients correspond well with measured partition coefficients to the lipid fraction in biological membranes (*30*).

If a compound's liposome−water partition coefficient is known, it is therefore possible to distinguish between the aqueous effect concentration, EC_w, and the internal effect concentration, EC_m, that is, the concentration in the membrane. The logarithms of these two concentrations are related by the following expression

$$\log 1/EC_m = \log 1/EC_w - \log D_{mw} \quad (1)$$

with log D_{mw} denoting the logarithm of the pH-dependent liposome−water partition coefficient of a weak acid. Equation 1 makes clear that a compound with a high aqueous toxicity, log $1/EC_w$, but a low tendency to partition into the membrane, that is, a low log D_{mw}, has a high intrinsic uncoupling activity. On the other hand, a compound with comparable log $1/EC_w$ and large log D_{mw} values is intrinsically not very active.

There remains the question of how log D_{mw} can be calculated. If both the neutral and the charged species partition into the membrane, the following relationships applies

$$D_{mw} = f_{HA} \cdot K_{mw,HA} + (1 - f_{HA}) \cdot K_{mw,A^-} \quad (2)$$

where $K_{mw,HA}$ is the lipid water partition coefficient of the weak acid, K_{mw,A^-} is the lipid water partition coefficient of its anion, respectively, and f_{HA} is the fraction of neutral species in water, f_{HA}, which for a given pH value is defined by

$$f_{HA} = \frac{1}{1 + 10^{pH - pK_a}} \quad (3)$$

In many QSAR studies, the descriptor used for decribing the partitioning into the membrane phase is the octanol–water partitioning coefficient, K_{ow}. The fundamental differences between K_{ow} and K_{mw} have been extensively described in numerous publications and two recent reviews (*26, 31*). In summary, the uncharged species log K_{mw} and log K_{ow} show a relatively high correlation (cf. section on the calculation of descriptors for $K_{mw,HA}$), which, together with the availability of relatively accurate calculation methods, explains the enormous success of log K_{ow} in QSAR studies. However, in the case of charged species, the octanol–water partition coefficient is by orders of magnitude lower than K_{mw,A^-} and is also not linearly correlated to $K_{mw,HA}$ (*27, 30, 32, 33*). Therefore, the use of eq 2 with octanol–water partition coefficients instead of liposome−water partition coefficients would produce misleading results in QSAR studies, as has been discussed in detail elsewhere (*34*). Equations 1 and 2 allow one to cover the uptake part depicted in Figure 1a and to accurately transform the aqueous effect concentration, EC_w, to the intrinsic effect concentration, EC_m, in the membrane.

The Kinspec test system currently only allows one to work with nominal concentrations because the chromatophore vesicles are so

small (<100 nm) that they cannot be separated easily from the aqueous solution. The derivation of EC_w and EC_m from the nominal concentration, EC_{tot}, is straightforward and is given in SI-2. The assumption behind the calculation is that the fraction of compound partitioning into the protein phase is negligible as compared to the fraction partitioning into the lipid phase. This assumption has been confirmed for a series of compounds in earlier studies (*30*). However, it must be emphasized that it is highly important to have either experimental data for log D_{mw} or an accurate calculation method for log D_{mw} to obtain meaningful values for EC_w and EC_m.

Data Set. A total of 35 uncouplers measured with Kinspec was used for this study (see SI-3 for the structures). Compounds **1**–**14** have been measured in this study, while compounds **15**–**35** were measured in earlier studies (*35*). A look at the first 14 structures makes clear that we purposely chose highly diverse structures (from a chemical class perspective—not from an MOA perspective). The second criterion for choosing a structure was the availability of experimental descriptors; for example, compounds with measured liposome−water partition coefficients were preferred. The third criterion was the need to explore the limits of Kinspec and of the model regarding the pK_a. The previously measured 21 phenolic uncouplers covered a pK_a range of 3.7−8.6. While the uncoupling activity continuously becomes weaker toward higher pK_a values, no such trend was observed for low pK_a. Therefore, we chose several compounds with pK_a < 3.7 to shed more light into a possible lower pK_a limit for uncoupling.

All effect concentrations were determined at pH 7. The negative logarithm of the aqueous effect concentrations ranged over six orders of magnitude from 3.08 (26DNP) to 9.11 (Fluazinam). The log K_{ow} ranged from 1.22 (26DNP) to 5.44 (35DBC). The experimental descriptors of the 14 compounds measured in this study were collected from literature, while the experimental descriptors of the 21 phenolic uncouplers were determined in earlier studies at Eawag. References for each compound are given in Table 1.

A remarkable feature of this study is that for most compounds (cpds) experimental physicochemical descriptors were available, for example, for pK_a (34 cpds), $K_{mw,HA}$ (25 cpds), K_{mw,A^-} (25 cpds), K_{ow} (29 cpds), and for k_{HA} and k_{A^-} (7 cpds). This has several advantages: (i) Experimental descriptors can be used to establish the QSAR model, and subsequently, they can be replaced in a stepwise manner by calculated descriptors. (ii) It is possible to check directly for each descriptor whether the calculated descriptors have a sufficient quality, and thus, it is easier to track the sources of errors.

To fully exploit the second advantage, additional literature data on $K_{mw,HA}$, K_{mw,A^-}, pK_a, and K_{ow} were collected. This allowed the examination and optimization of the descriptor calculations without using the activity data. One might call these data sets descriptor calibration sets. The extensive literature search resulted in calibration sets of the following size: $K_{mw,HA}$, 136 cpds; K_{mw,A^-}, 35 cpds; pK_a, 57 cpds; and K_{ow}, 134 cpds. The CAS numbers and experimental data of these compounds can be found in Table SI-5 (showing the data for the compounds in addition to Table 1). The SMILES Code of the 35 uncouplers of Table 1 and of the totally 144 additional compounds of SI-5 are given in SI-6A and B, respectively.

For 10 compounds tested for uncoupling, no experimental data on $K_{mw,HA}$ and K_{mw,A^-} were available. In such cases, one would need calculated values to determine EC_w and EC_m from the nominal concentration EC_{tot}. However, as the calculation errors strongly affect the obtained effect concentrations, we evaluated these compounds only for their nominal concentration, EC_{tot}. The values for EC_w and EC_m of these 10 compounds are nevertheless given in Table 1, with the brackets indicating the lower confidence in the values of the aqueous effect concentration. Figure 2 shows a flowchart summarizing the approach used in this study.

Calculation of Descriptors. The calculation methods chosen should come as close as necessary to the experimental descriptors. While there are a sufficient number of calculations methods for pK_a and an abundance of methods to calculate log K_{ow}, there is no established method available to calculate experimental descriptors like the translocation rate constant of the charged species, k_{A^-}. The first choices for descriptor calculations were quantum chemical (QC) methods, however, with the goal in mind to later move to empirical calculation methods once a model with a good fit had been found. Table 2 lists the descriptors of the experimental models and the calculated descriptors that should approximate the experimental ones.

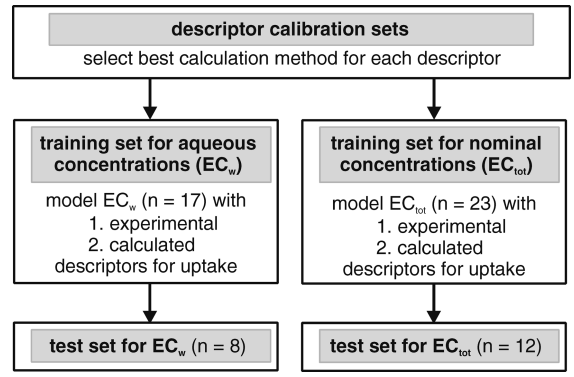

Figure 2. Workflow of the study with three main steps: (I) choice of a suitable descriptor calculation method; (II) development of a model for the training set with aqueous concentrations, EC_w, and nominal concentrations, EC_{tot}; and (III) a test set with EC_w and EC_{tot}.

Table 2. Experimental Descriptors Determining the Uncoupling Activity and Calculated Approximations of These Descriptors

descriptor	descriptors tested
pK_a	pK_a calculated with (a) QC methods and (b) empirical methods (SPARC)
log $K_{mw,HA}$	(a) direct calculation with COSMOmic (b) approximation by log K_{ow} calculated with KowWin
log K_{mw,A^-}	(a) direct calculation with COSMOmic (b) approximation by log K_{ow} calculated with KowWin and assumed constant $\Delta\log K_{mw}$ of 1.0
k_{HA}	(a) approximation by calculating $\Delta G_{well\text{-}barrier}$ (with COSMOmic) (b) approximation by calculating the partition coefficient to reference solvents representing the energy well and the hydrophobic barrier with either QC methods or empirical methods (SPARC)
k_{A^-}	(a) same methods as for k_{HA} but for charged species (b) simplification by correlating k_{A^-} with k_{HA}

The calculation method for each descriptor and the underlying assumption are described in the following section, and the calculated values for each descriptor are given in SI-7.

Acid Dissociation Constant, pK_a. An important characteristic of weak acids acting as uncouplers is that the proton is not only dissociated from OH groups, for example, phenols, but also from NH groups, for example, in the case of benzimidazoles. Two different methods covering both types of acidic compounds were used to calculate pK_a values: (a) COSMO-RS (*46*), a QC method, and (b) SPARC, an empirical method (*47*).

(a) The COSMO-RS calculations were performed as described in ref *48*. Briefly, the calculations consist of the following steps: (i) creation of start geometry using CORINA (*49*); (ii) for both the neutral species and the anion, full DFT geometry optimization with the Turbomole program package (*50, 51*) using the B-P density functional (*52, 53*) with TZVP quality basis set and applying the RI approximation (*54*) and applying the COSMO continuum solvation model in the conductor limit ($\epsilon = \infty$); (iii) analysis of tautomers and conformers if necessary; and (iv) actual COSMO-RS calculation, which takes into account the deviation of the solvent, water in the case of pK_a calculations, as compared to an ideal conductor in a model of pairwise interacting molecular surfaces. The chemical potential differences arising from these interactions are evaluated using a statistical thermodynamics algorithm for

independently pairwise interacting surfaces. More detailed descriptions of the general COSMO-RS methodology are given elsewhere (*55–58*).

(b) The empirical SPARC method is described in refs *59* and *60*. It is based on perturbation theory and accounts for resonance and electrostatic effects plus other perturbations such as H-bonding, steric contributions, and solvation.

Liposome−Water Partition Coefficient of Neutral Species, log $K_{mw,HA}$.

(a) Recently a QC method for the calculation of lipid−water partitioning (COSMOmic) has been developed (*61, 62*). COSMOmic is an extension of the COSMO-RS approach for the calculation of chemical potentials described above. In COSMOmic, the partitioning process differs from that of a bulk solvent in one fundamental way: The solvent is treated as a series of mixtures of different solvents with water, salts, and phosphatidylcholine, and this series has a spatial arrangement. The spatial arrangement of a water−phosphatidylcholine bilayer system is introduced by connecting n solvent layers representing the change of solvent composition; that is, layer 1 (the center of the membrane) consists almost completely of fatty acids alkyl chains, while toward the outer end of the system the layers contain more and more phosphatidyl head groups (but also some water) and layer n consists only of water (the bulk water phase). The percentage of each atom type was taken from a recent molecular dynamics simulation of 512 lipid molecules (dimyristoylphosphatidylcholine) in an aqueous 1 M NaCl solution with a lipid mole fraction of 0.04 (*63*).

In a COSMOmic calculation, the center of the solute is moved through the n liquid mixtures representing the membrane. In each layer, the solute is rotated to a large number of orientations. For each orientation, the chemical potential is calculated accounting also for the fact that the solute extends into the adjacent layers with different compositions. The final result of a COSMOmic run allows the calculation of a probability of occurrence of the solute for each of these n mixtures. The partition coefficient between the membrane and the water, $K_{mw,HA}$, is then obtained by summing up the probability of a solute being in mixtures 1 to $n - 1$ and dividing it by the probability of the solute being in the mixture n (the bulk water).

(b) The second option to calculate log $K_{mw,HA}$ is a linear regression between log $K_{mw,HA}$ and the logarithm of the octanol–water partition coefficient, log K_{ow}. Vaes et al. (*64*) determined such a relationship with $\log K_{mw,HA} = 0.904 \cdot \log K_{ow} + 0.515$ with a high correlation of $R^2 = 0.89$ for a set of 11 polar compounds. This analysis was repeated with the compounds of Table 1 and SI-5, which have experimental data on both $K_{mw,HA}$ and K_{ow}. Both the training set with 81 pairs of experimental $K_{mw,HA}$ and K_{ow} and the test set with 40 such pairs had almost the same regression coefficients and the same measures of goodness-of-fit and predictivity. Therefore, the training set and test set were merged to determine the final relationship used for further analysis with almost the same regression coefficients than the relationship published by Vaes et al.

$$\log K_{mw,HA} = 0.92(\pm 0.062) \cdot \log K_{ow} + 0.37(\pm 0.20) \quad (4)$$

with $n = 121$, $R^2 = 0.88$, $Q^2 = 0.87$, SD = 0.41, and $SD_{cv} = 0.42$.

Only nine compounds had larger residuals than 0.7 log units, and only two had residuals above 1.0 log units (dibutylsuccinate and bisphenol A). Thus, for the neutral species, the simplest approach to calculate log $K_{mw,HA}$ is the extrapolation from log K_{ow}. Note that in other studies lower correlations were reported, for example, ref *31* lists examples where the extrapolation from log K_{ow} leads to predicted log $K_{mw,HA}$ values, which are orders of magnitude away from the experimental value. This was mainly observed for hydrophilic compounds with log $K_{ow} < 1$, which partition into membranes by orders of magnitude more than would be expected from eq 4 while for highly lipophilic compounds with log $K_{ow} > 5.5$ the contrary seems the case (*65*). Therefore, although eq 4 is based on a large sample of 121 compounds, it can only be used safely for neutral compounds with $1 < \log K_{ow} < 5.5$. All compounds of this study fell into this log K_{ow} window; therefore, eq 4 should give reasonably good predictions (*65*). If no experimental data on log K_{ow} were available, they were calculated with the KowWin program from the EPI Suite (*66*), which is a fragment-based approach (*67*).

Liposome−Water Partition Coefficient of Anion, log K_{mw,A^-}.

(a) The log K_{mw,A^-} value can be calculated in the same fashion as described above for log $K_{mw,HA}$, as COSMOmic also allows the treatment of charged molecules.

(b) For lack of a better solution, log K_{mw,A^-} has been approximated in earlier studies from our laboratory by using the relationship log K_{mw,A^-} equals log $K_{mw,HA}$ minus the average Δlog K_{mw}, that is, the difference between log $K_{mw,HA}$ and log K_{mw,A^-}. On average, Δlog K_{mw} was observed to be close to one log unit (*26, 68*). Therefore, the approximation is

$$\log K_{mw,A^-} = \log K_{mw,HA} - \Delta\log K_{mw} \approx \log K_{mw,HA} - 1 \quad (5)$$

For the 35 compounds in Table 1 and SI-5 with experimental data on $K_{mw,HA}$ and log K_{mw,A^-}, the values for Δlog K_{mw} range from 0 to 1.98 and the average is 1.02 log units with a standard deviation of 0.64. Therefore, a considerable amount of variance is added by using eq 5, but no error larger than 1 log unit is made. Note that only four of the 35 compounds were carboxylic acids, with all of them having Δlog K_{mw} around 2. For such compounds, eq 5 is likely to cause high errors. If no experimental data for log $K_{mw,HA}$ were available, they were calculated by eq 4.

Translocation Rate Constant of Neutral Species, k_{HA}. The general description of permeability describes the passive diffusion of a compound from the bulk aqueous phase through the unstirred water layer (adjacent to membranes), through the membrane, and again through the unstirred water layer on the other side into the bulk aqueous phase. The inverse of permeability, that is, the resistance to permeation, R_{tot}, can be written as a simple sum

$$R_{tot} = 1/P_{tot} = 1/P_m + 1/P_u \quad (6)$$

where P_{tot} refers to the measured permeability, P_u refers to the unstirred water layer permeability, and P_m refers to the permeability in the membrane. In the case of uncoupling, we assume that the total concentrations in the water phase and in the membrane are not changing during our experiments; therefore, only P_m needs to be considered.

Equation 7 describes P_m as defined by Diamond and Katz (but without interfacial resistances) (*69*)

$$R_m = \frac{1}{P_m} = \int_0^d \frac{dz}{K(z)D(z)} \quad (7)$$

with d the thickness of the membrane, $K(z)$ the depth-dependent partition coefficient from water into the membrane, and $D(z)$ the diffusion coefficient in the membrane at depth z with the z-axis defined as the membrane normal. If the transport is governed, however, by a distinct barrier region within the membrane (*69–71*), which is the case for lipophilic anions (and often also for neutral species), eq 7 simplifies to

$$R = \frac{d_{barrier}}{K_{water\text{-}barrier} D_{barrier}} = \frac{1}{P_m} \quad (8)$$

with $K_{water-barrier}$ the solute partition coefficient from water to the barrier region, $D_{barrier}$ the solute diffusion coefficient through the barrier region of the membrane, and $d_{barrier}$ the thickness of the barrier region. As illustrated by Figure 1b, it is possible to distinguish two processes defining P_m: partitioning into the energy well and transfer from the energy well over the energy barrier (intramembrane permeability) or expressed in Gibb's free energy differences:

$$\Delta G_{\text{water-barrier}} = \Delta G_{\text{water-well}} + \Delta G_{\text{well-barrier}} \quad (9)$$

This distinction is also made in biophysical studies where P_m is written as the product of the linear partition coefficient, β_{HA}, with the dimension of length and the translocation rate constant, k_{HA}, with the dimension of 1/time or and β_{A^-} and k_{A^-} in the case of anions, respectively.

$$P_m = \beta_{HA} \cdot k_{HA} \quad (10)$$

β_{HA} determines the surface density $n_{HA\text{-well}}$ with units mol per surface of the HA species present in one partition well with $n_{HA\text{-well}} = \beta_{HA} \cdot C_{HA\text{-water}}$. Using eqs 8−10, one can express the translocation rate constant as

$$k_{HA} = \frac{2 \cdot D_{\text{barrier}}}{d_{\text{barrier}} \cdot d_{\text{well-well}}} \exp\left(-\frac{\Delta G_{\text{well-barrier,HA}}}{R \cdot T}\right) = \frac{2 \cdot D_{\text{barrier}}}{d_{\text{barrier}} \cdot d_{\text{well-well}}} K_{\text{well-barrier,HA}} \quad (11)$$

with $d_{\text{well-well}}$ the distance between the partition wells, R the gas constant in units of J mol^{-1} K^{-1}, and T the temperature in Kelvin. Equation 11 contains several assumptions used in the present study. First, the diffusion coefficient is roughly constant from one energy well over the barrier region to the next energy well, and second, the partitioning to the barrier region given by $K_{\text{well-barrier}}$ is the rate-limiting step, which is directly proportional to k_{HA} and k_{A^-}, respectively.

A comment related to the first assumption is necessary: The change of the diffusion coefficient, $D(z)$, over the membrane normal has been investigated recently in an MD simulation for a series of eight small organic compounds (*71*). $D(z)$ remained more or less constant from the region of the polar head groups to the barrier region, which justifies the assumption of a constant $D_{\text{barrier}}(z)$. However, for different molecules, the value of $D_{\text{barrier}}(z)$ is not the same but inversely proportional to the radius according to the Stokes–Einstein relation (for spherical molecules). For the current series of compounds, the differences in radii are smaller than a factor of 2; therefore; the effect of neglecting differences in the diffusion coefficient should be relatively small.

This said, the most important aspect of eq 11 is that it should be possible to model the differences of the translocation rate constants of different uncouplers by calculating $K_{\text{well-barrier}}$ for the neutral and the charged species while all other terms are more or less constant for all compounds.

(a) Approximation of k_{HA} by Calculating $\Delta G_{\text{well-barrier}}$. The result of a COSMOmic calculation is a probability distribution over the membrane normal, that is, over the z-axis. The free energy difference between the energy well and the energy barrier, $\Delta G_{\text{well-barrier}}$, can be calculated from the probability of a compound being located in the barrier region, p_{barrier}, and the probability of being located in the energy well, p_{well}, by taking the difference of the natural logarithms

$$\Delta G_{\text{well-barrier}} = -R \cdot T \cdot \ln\left(\frac{p_{\text{barrier}}}{p_{\text{well}}}\right) = -R \cdot T \cdot \ln(K_{\text{well-barrier}}) \quad (12)$$

with R the gas constant and T the temperature. Figure 3 shows an example of a COSMOmic probability distribution over the membrane normal. To make the picture comparable to the free energy profile of Figure 1b, the negative logarithm is plotted for both the neutral species and the anion of CCCP.

Qualitatively, the calculated energy profile corresponds well with the profiles shown in Figure 1b. The main difference is that the energy wells of the neutral species are shifted toward the center of the membrane. The descriptor used for our QSAR model is $\Delta G_{\text{well-barrier,HA}}$, which is directly proportional to the logarithm of the translocation rate constant k_{HA}. Note that CCCP is quite a special case because the partition coefficients of neutral species and anions are the same. For most compounds, the partition coefficient of the neutral species is much higher than for the charged species; thus, the depth of the energy well in the equivalent of Figures 1b and 4 would be deeper for the neutral species.

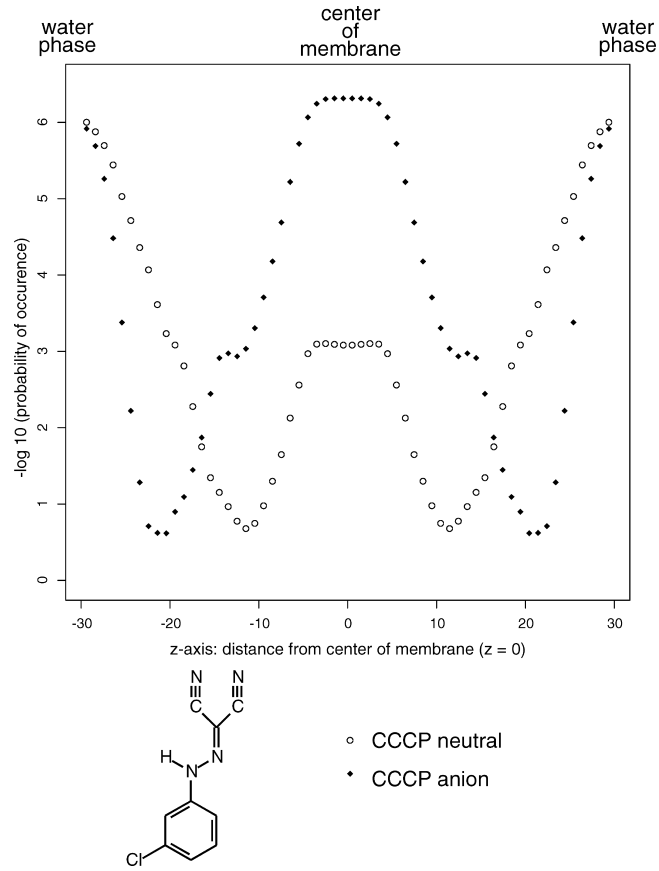

Figure 3. Calculated probability distribution for HA and A$^-$ species for the uncoupler CCCP. The z-axis represents the membrane normal. The scale is in Ångstroms ranging from 0 (the center of the membrane) to 32 (the bulk water phase). The probabilities are plotted as inverse logarithms to allow comparisons with free energy profiles.

(b) Approximation of k_{HA} by Calculating the Partition Coefficient between Reference Solvents. Mayer and Anderson proposed a model for permeability, which also uses the energy barrier as rate-limiting step (*72*). For a selection of 12 structural analogues, they showed that it might not be necessary to measure the permeability through phosphatidylcholine bilayers but took experimental partition coefficients between water and a "reference solvent" mimicking the barrier domain. Therefore, for the process shown in Figure 1, instead of calculating $K_{\text{well-barrier}}$ from the full probability profile over the membrane, one can also calculate the partition coefficient between two reference solvents, one representing the energy well and the other one representing the energy barrier. This would constitute a major simplification.

Translocation Rate Constant of Charged Species, k_{A^-}.

(a) The same approaches as described for the approximation of k_{HA} were used for charged species to calculate $\Delta G_{\text{well-barrier,A}^-}$ and partition coefficients between reference solvents.

(b) A major simplification would be the assumption that k_{HA} and k_{A^-} are correlated. Then, the measures calculated for k_{HA} could be scaled to the (lower) value corresponding to that of the charged species and possibly also be calculated with empirical methods like SPARC. Even if this assumption leads to a somewhat lower quality of the model, it could be useful for a first crude prediction and especially for database screening.

Statistical Methods. The data set was split into two-thirds training set (23 compounds) and one-third external test set of 12 compounds prior to any model building. Taylor−Butina clustering

(*73*) was used to give a more balanced split of the data. The two descriptors log D_{mw} and $\Delta G_{well\text{-}barrier}$ and the response EC_m were used as it could be expected that these three descriptors cover the variance of EC_w to a large extent (if the mechanism is unknown, splitting the data set with such an algorithm requires a larger number of more general descriptors, but this seemed unnecessary in this study). With a similarity threshold of 0.8, three large and five small clusters of compounds were obtained. Within each cluster, the compounds were ordered according to EC_w and then distributed between training and test sets.

Two criteria were used to assess the model performance: the coefficient of determination, R^2, and the residual standard deviation, SD, as defined in ref *74*. Note that for regression models usually the residual SD also accounts for the number of variables (descriptors). However, this was not done here to allow the direct comparison of regression and nonlinear models, and as no more than two descriptors were used, the difference is small.

The 95% confidence intervals of the regression coefficients were calculated as the product of the standard error and the quantile of a *t* distribution with the corresponding degrees of freedom. In many publications, the standard error is given, which is about a factor of 2 smaller.

A 5-fold cross-validation (or leave-many-out) was used for internal validation using the training set. The training set was split 19 times randomly into five subsets, that is, the cross-validation was repeated 19 times to obtain more balanced samples. The R^2 calculated in a cross-validation is denoted as Q^2 (in our case, the mean of 19 Q^2 values). For the test set, the ${Q_{ext}}^2$ was calculated as proposed in ref *74*, that is, with the mean of the experimental values $\bar{y}$ taken as the mean of the training set and not the test set. The SD calculated in the cross-validation is denoted as SD_{cv} and for the test set as SD_{ext}. In other works, SD_{cv} and SD_{ext} are also referred to as SDEP (standard deviation error of prediction) (*74*). All statistical calculations were made with R 2.1 (*75*).

Results

Uncoupling Activity. All compounds of Table 1 showed an uncoupling activity in the Kinspec system, even the three compounds with $pK_a < 3.7$ (26DBr4NP, 6Cl24DNP, and 246TriNP with pK_a values of 3.39, 2.06, and 0.38, respectively), which were measured to examine if there is a possible lower pK_a limit for uncoupling measured in Kinspec. Thus, no lower pK_a limit can be set from the Kinspec measurements. The implications will be addressed in the Discussion section.

Sensitivity Analysis of Earlier Transport Models. To determine which compound properties have the strongest influence on the activity of uncouplers, the first method applied in this study was a sensitivity analysis with existing models. The study of Escher et al. (*22*) with experimentally fitted translocation rate constants (based on fitting the transport equations under varying pH values) for seven compounds and independently determined liposome−water partition coefficients was chosen, but similar results should be obtained from studies with BLMs. The process of uncoupling is described by a transport model, which results in two linear inhomogeneous differential equations: one for the change of surface charge describing the movement through the membrane caused by potential differences (eq 43 in ref *22*) and one describing the diffusion processes caused by concentration differences (eq 41 in ref *22*). The analytical solution of these equations can be used to determine the endpoint of $k_{uncoupling} = 0.5$.

Figure 4 shows the sensitivity of the negative logarithm of the aqueous effect concentration at pH 7.0 to the five descriptors of the model, pK_a, $K_{mw,HA}$, K_{mw,A^-}, k_{HA}, and k_{A^-}. The functions of the transport equation used to generate the plots are given in SI-4.

For dino2terb (Figure 4a), it is striking that the two descriptors describing the charged species K_{mw,A^-} and k_{A^-} have the strongest effect when they differ from the experimental value (which is assumed to be the "real" value). Both descriptors have a slope coming close to one in the sensitivity analysis; that is, if the translocation rate constant is calculated 1 order of magnitude too high, the calculated effect concentration is equally by 1 order of magnitude too high. On the other hand, even an error of 1 order of magnitude in the calculation of pK_a, $K_{mw,HA}$, or k_{HA} has only a minimal effect on the effect concentration.

For CCCP (Figure 4b), which is a very strong uncoupler, the picture is more complex. For pK_a, both an overestimation and an underestimation lead to a reduced predicted uncoupling activity. This is a consequence of the pK_a value of 5.95 being close to the pH of 7 while the insensitivity of dino2terb to this descriptor is caused by the larger difference between the pH and the $pK_a = 4.8$ of dino2terb. In contrast to dino2terb, the model also becomes sensitive to the descriptors for the neutral species; however, still a little bit less than for the charged species, for example, an underestimation of k_{HA} by one log unit leads to an underestimation of EC_w of 0.5 log units while for k_{A^-} the error is still higher with about 0.8 log units. Another observation is that for CCCP an underestimation of a descriptor leads to a larger error than an overestimation. The following three conclusions can be drawn from these observations:

- To establish a mechanistic QSAR model, the search for calculation methods for the two descriptors of the charged species, K_{mw,A^-} and k_{A^-}, has the highest priority.
- Errors in the calculation of these two descriptors translate almost directly into the errors in the calculated effect concentration; that is, if the residual standard deviation of the calculation method for K_{mw,A^-} or k_{A^-} has a certain value, it is highly unlikely that the residual standard deviation of the uncoupler model is much lower.
- The two descriptors K_{mw,A^-} and k_{A^-} have such a strong influence that it might be possible to establish a linear regression model for EC_w based only on these two descriptors. The other descriptors (pK_a, $K_{mw,HA}$, and k_{HA}) might not be significant in a regression model.

Comparison of Quality of Descriptor Calculation Methods. Uptake. Several methods were used to calculate each descriptor. The decision of which method should be used for modeling the activity was made by comparing the method's ability to predict the four physicochemical descriptors pK_a, log K_{ow}, log $K_{mw,HA}$, and log K_{mw,A^-} using the compiled data sets given in Table 1 and SI-5. Table 3 summarizes the quality of the predictions of the different methods for each of these four descriptors.

The first row of Table 3 covers the QC calculation methods described in the Materials and Methods section each under subsection (a) of the corresponding descriptor, while the second row covers the empirical approaches each described under subsection (b) of the corresponding descriptor. The data sets used to test the model performances for the predictions of pK_a and log K_{ow} could easily be extended to a larger number of compounds; however, the important aspect was to test the descriptor calculation methods with a data set that is structurally similar to uncouplers and not to the biggest number of compounds possible.

The high predictive power of the QC method for the pK_a is in agreement with the original publication where almost the same values were determined (*48*). The rather high standard deviation of the empirical pK_a predictions was mainly due to the 10 NH-acidic compounds, which showed errors up to 3.5

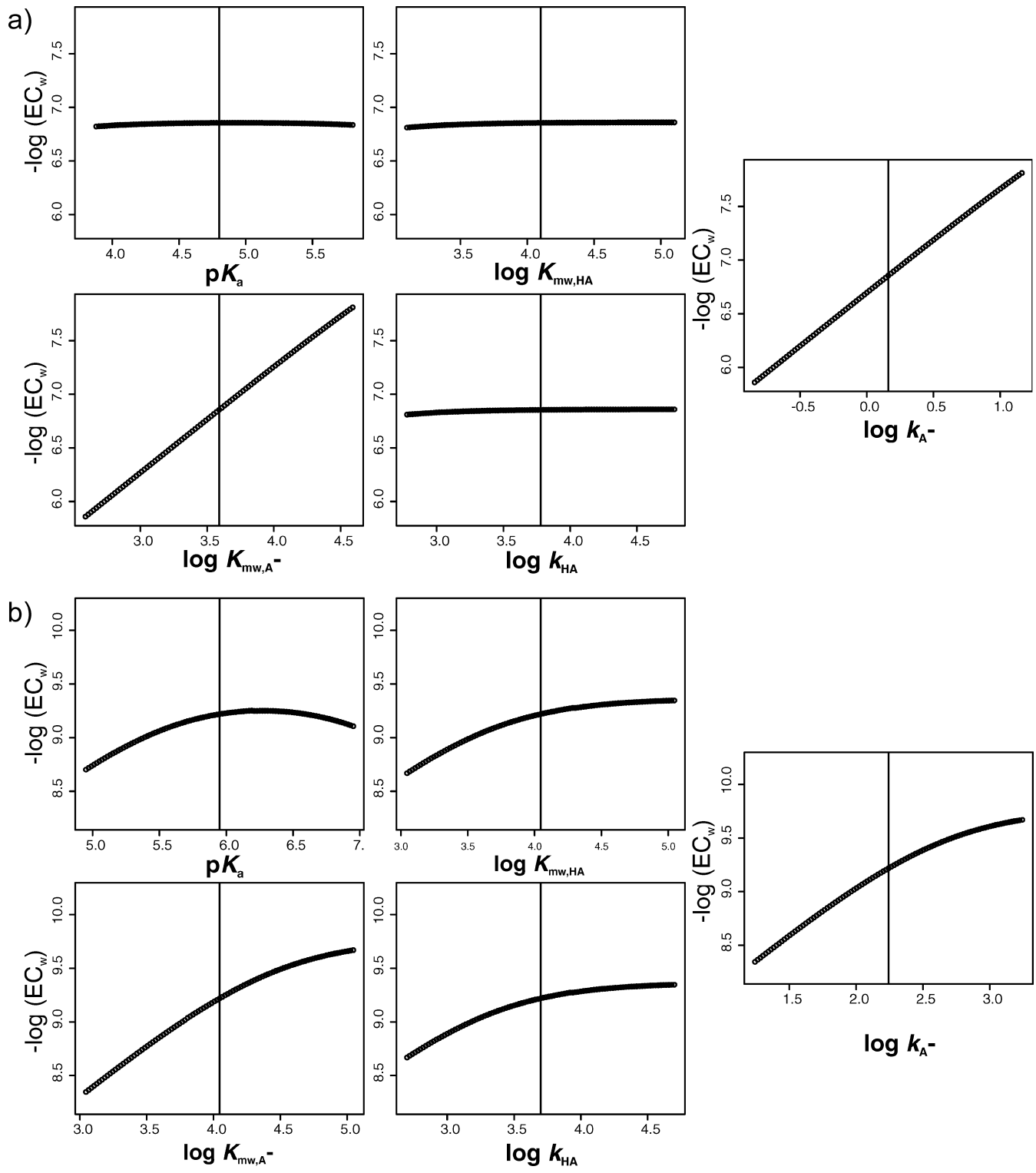

Figure 4. Sensitivity analysis for the two compounds (A) dino2terb (2-*tert*-butyl-4,6-dinitrophenol) and (B) CCCP (carbonyl cyanide *m*-chlorophenylhydrazone). Five experimentally determined model descriptors, acid dissociation constant, pK_a, lipid water partition coefficients of the weak acid, $K_{mw,HA}$, and its anion K_{mw,A^-}, and the translocation rate constants of the weak acid, k_{HA}, and its anion k_{A^-} were analyzed. Each plot shows the change in modeled EC_w upon varying a given descriptor over an order of magnitude and keeping all others constant. The vertical line in the middle of the plot shows the experimentally fitted "real" value of the descriptors.

units. The QC calculation for log K_{ow} had a relatively large SD mainly caused by three compounds with errors >2 log units (35DBC, fluazinam, and TBBPA). In the case of TBBPA, it is quite likely that the value from literature is not correct as both methods resulted in a 2 log units higher value, while in the other cases no such conclusion can be drawn. Without these three compounds, the QC method would have an $R^2 = 0.77$ and an SD = 0.61. It should be mentioned that it is difficult to draw conclusions from this method comparison in the case of pK_a and log K_{ow} because a large portion of the data were probably used to establish the empirical methods in the first place (resulting in some cases in an error of zero).

The QC predictions for log $K_{mw,HA}$ and log K_{mw,A^-} were scaled with a linear transformation using the experimental data, that is, the relationship log $K_{mw,HA}$ (experiment) = β_1·log $K_{mw,HA}$ (COSMOmic) + β_0 (which does not affect R^2 but

Table 3. Correlation and SD of Different Descriptor Calculation Methods for Descriptor Calibration Sets[a]

	pK_a (n = 57)		log K_{ow} (n = 49)		log $K_{mw,HA}$ (n = 97)		log K_{mw,A^-} (n = 35)	
	R^2	SD	R^2	SD	R^2	SD	R^2	SD
QC	0.98	0.45	0.62	0.81	0.80	0.66[b]	0.68	0.63[b]
empirical	0.87	1.26	0.80	0.59	0.86	0.56[b]	0.60	0.76[b]

[a] For the log $K_{mw,HA}$ and log K_{mw,A^-}, only a training set of 97 and 35 compounds, respectively, was used (to keep a test set for future studies). [b] Predictions fitted with the training set, that is, SD, were determined as proposed in ref *74* but y^{fit} was used instead of y^{calc}.

slightly lowers the SD). The fit equations for log $K_{mw,HA}$ and log K_{mw,A^-} are given in SI-7.

Table 3 suggests that in the absence of experimental data for the uptake, it is most sensible to calculate the pK_a with the QC method, while for the membrane−water partitioning, no clear conclusion is possible; therefore, both approaches to calculate log $K_{mw,HA}$ and log K_{mw,A^-} were used.

Intrinsic Activity. Equations 9–12 show that it is possible to calculate the translocation rate constants of molecules if the distribution over the membrane normal is known; that is, the logarithm of the translocation rate constant is approximately proportional to the free energy difference between the energy well and the energy barrier in the membrane ($\Delta G_{well\text{-}barrier}$). Only a limited number of seven experimentally determined translocation rate constants from ref *22* were available to test the quality of the calculated $\Delta G_{well\text{-}barrier}$. The fit of the logarithm of the translocation rate constant of the charged species, log(k_{A^-}) vs $\Delta G_{well\text{-}barrier,A^-}$ calculated with the QC method given in SI-7 is good with $R^2 = 0.74$ and SD = 0.38. For the neutral species, that is, for log(k_{HA}) vs $\Delta G_{well\text{-}barrier,HA}$, the correlation is lower with $R^2 = 0.43$ and SD = 0.51. The main reason for the lower R^2 is the small variability of k_{HA}, while the SD is still comparable with the one of k_{A^-}.

Quantitatively, the barrier heights of anions calculated with COSMOmic (for DMPC bilayers) are consistently higher than the barrier heights expected from experimental translocation rate constants for chromatophores (*22*) and also with lecithine-chlorodecane black lipid bilayers (*20*). An example is the anion of CCCP, which has a calculated $\Delta G_{well\text{-}barrier}$ of +8.1 kcal/mol (with COSMOmic), while the experimental value is +6.1 (*19*). This would result in a difference in k_{A^-} of a factor of 30. It is possible that this difference is caused by the different type of phospholipid used for the calculation. However, for a regression model, it does not matter if the translocation rate constants or the values of $\Delta G_{well\text{-}barrier}$ are consistently too high as long as the relative differences are correct, while for the nonlinear model, the calculated k_{HA} and k_{A^-} need to be corrected by scaling them with experimental data. The most important observation made in the section above is that fit of the (few) experimental translocation rate constants with calculated $\Delta G_{well\text{-}barrier}$ results in relatively small errors, which is an indicator that the translocation rate constants can be calculated well with COSMOmic.

Training Set. The main conclusion from the sensitivity analysis was that a regression model with only two descriptors might be sufficient to predict EC_w. The two descriptors of choice were the $\Delta G_{well\text{-}barrier}$ of the charged species calculated by eq 12 and, for the uptake, either log K_{mw,A^-} or log D_{mw} (as defined in eq 2).

Linear Regression Model for EC_w. Following the scheme shown in Figure 2 and the results shown in Table 3, different linear regression models for EC_w were tested. Table 4 summarizes the statistical quality of the regression models.

The first four models differed only by the calculation method for log D_{mw}, which was calculated according to eq 2 using either experimental pK_a values and liposome−water partitioning coefficients (R1), experimental pK_a and octanol–water partitioning coefficients (R2), and using only calculated descriptors with QC (R3) or empirical methods (R4), respectively.

Regression model R1 with log D_{mw} calculated with experimental descriptors had the best fit of all models and took the following form

$$\log(1/EC_w) = 1.09(\pm 0.29)\cdot \log D_{mw} - 0.28(\pm 0.09)\cdot \Delta G_{water\text{-}barrier,A^-} + 5.16(\pm 1.56) \quad (R1)$$

with $n = 17$, $R^2 = 0.91$, $Q^2 = 0.87$, SD = 0.47, $SD_{cv} = 0.56$, and $F = 66.95$.

The negative values of the coefficient for $\Delta G_{well\text{-}barrier,A^-}$ means that the activity decreases with increasingly high energy barrier ($\Delta G_{well\text{-}barrier,A^-}$ has positive values), while increasing values for D_{mw} increase the activity. As the regression coefficient of log D_{mw} is about one, eq R1 shows that an increase in the partition coefficient directly causes the same increase of the toxic activity. The standard deviation is comparable to the one determined for the seven compounds with experimental translocation rate constants. The largest error was caused by dino4terb (4-*tert*-butyl-2,6-dinitrophenol), which was overestimated by 1.17 log units, while all other errors were smaller than 0.7 log units. The values predicted with eq R1 vs the experimental values are shown in Figure 5.

Two additional descriptors were tested to see if they could improve eq R1, namely, pK_a and $\Delta G_{well\text{-}barrier,HA}$, that is, the translocation rate constant of the neutral species. While pK_a was not significant in the regression model and did not contribute to the fit, $\Delta G_{well\text{-}barrier,HA}$ improved the fit of the model slightly to $R^2 = 0.94$ ($Q^2 = 0.90$), SD = 0.38 ($SD_{cv} = 0.48$). However, contrary to $\Delta G_{well\text{-}barrier,A^-}$, the coefficient for $\Delta G_{well\text{-}barrier,HA}$ was positive, which physically makes no sense; therefore, only linear models with two descriptors were evaluated further.

As expected, the models with calculated log D_{mw}, that is, models R3 and R4, had a lower quality of fit and larger residual standard deviations. The decrease of fit observed for the model with QC descriptors was mainly caused by three compounds with large errors over one log unit [4NP, −1.19; CCCP, –1.24; and warfarin, +2.08 (with the + indicating overestimation and the – indicating underestimation)]. The model with empirical descriptors had less extreme errors but about the same SD.

Regression model R2 was obtained by taking experimental data on pK_a and K_{ow} and calculating log D_{mw} with eqs 4 and 5 using the data given in Table 1. As these basic experimental property data are often available in practice, this appears a useful option and the quality of the fit increased substantially. Note that two compounds did not have experimental data on K_{ow} (dino2terb and dino4terb); therefore, values calculated with the empirical method given in SI-7 (column log $K_{mw,HA}$ KowWin) had to be used. Leaving out these two compounds of model R2 would result in $R^2 = 0.82$, $Q^2 = 0.72$, SD = 0.66, and $SD_{cv} = 0.82$. The coefficients for the uptake were almost the same for models R2 and R4, which confirms that taking experimental data if available is a sensible option.

In the sensitivity analysis, the transport model showed a very strong sensitivity to log K_{mw,A^-}. Therefore, two additional descriptors for the uptake were tested, again using the experimental data: R5, the partition coefficient of the charged species, log K_{mw,A^-}, and R6, the partition coefficient proportional to the concentration of the charged species in the membrane (for a given pH). The latter can be calculated by taking only the second term in eq 2,that is,

Table 4. Quality of Fit of Linear Regression Models for EC_w at pH 7 Using the Descriptor $\Delta G_{well\text{-}barrier,A^-}$ and Different Descriptors for Uptake (Defined by Eq 2 for Models R1–R4)[a]

		EC_w train ($n = 17$)							EC_w test ($n = 8$)	
	descriptor for uptake	R^2	Q^2	SD	SD_{cv}	β_1	β_2	β_0	Q^2_{ext}	SD_{ext}
R1	log D_{mw} (exp)	0.91	0.87	0.47	0.55	1.09	−0.28	5.16	0.88	0.61
R2	log D_{mw} [pK_a (exp) and K_{ow} (exp)]	0.80	0.71	0.69	0.83	0.83	−0.30	6.70	0.74	0.88
R3	log D_{mw} (QC method)	0.73	0.58	0.80	0.99	1.51	−0.16	3.13	0.75	0.86
R4	log D_{mw} (empirical)	0.72	0.58	0.82	0.99	0.85	−0.24	6.11	0.69	0.97
R5	log K_{mw,A^-} (exp)	0.77	0.66	0.74	0.90	1.12	−0.12	3.77	0.77	0.83
R6	log $[(1 - f_{HA}) \cdot D_{mw}$ (exp)]	0.63	0.42	0.94	1.17	1.01	−0.04	3.44	0.49	1.24

[a] R2 and R4 were both calculated with eqs 4 and 5 with experimental and calculated values, respectively. The columns β_1, β_2, and β_0 denote the regression coefficients for the descriptors log D_{mw}, $\Delta G_{well\text{-}barrier,A^-}$, and the intercept, respectively.

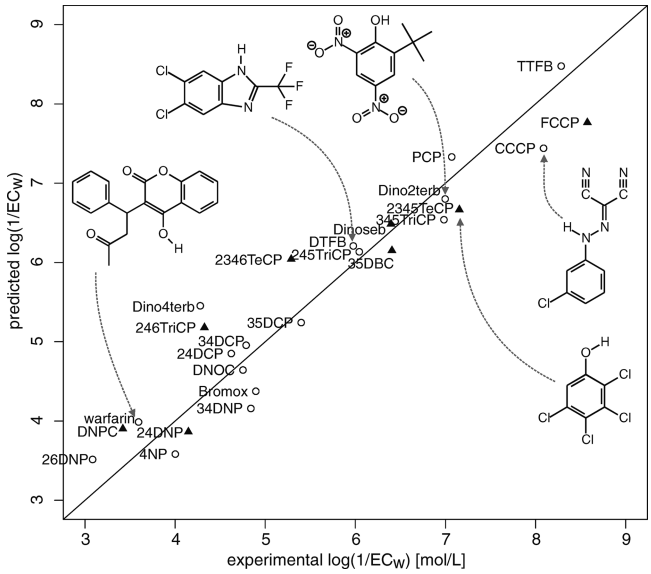

Figure 5. Experimental vs predicted aqueous effect concentrations, EC_w, at pH 7 for training and test sets with regression models based on experimental partition coefficients for the uptake (R1). Training set compounds (empty circles) and the test set compounds (black triangles). The structures are shown as examples to illustrate the diversity of the data set.

$(1–f_{HA})\cdot K_{mw,A^-}$. An example to illustrate the difference between R5 and R6 would be the two compounds 35DBP (pK_a = 8.28; log K_{mw,A^-} = 3.18) and Dino4terb (pK_a = 4.11; log K_{mw,A^-} = 3.23). While the descriptor for the uptake in R5 is nearly the same for 35DBP and Dino4terb, it is a factor of 22 higher in regression R6. Thus, regression model R6 reflects the idea that at comparable concentrations and translocation rate constants, a higher fraction of anion in the membrane should increase the activity, and a lower fraction (e.g., compounds with pK_a > 7) should decrease it.

The comparison of experimental descriptors for the uptake, that is, models R1, R5, and R6, shows that the model with log D_{mw} clearly has a better fit than the model with log K_{mw,A^-} or log $[(1 - f_{HA})\cdot K_{mw,A^-}]$, a result that will be treated in detail in the Discussion. The coefficients of equations R1−R6 are given in Table 4, and the plots of R2–R5 are given in SI-8.

Linear Regression Model for EC_{tot}. As stated in the section on the data set, 10 compounds did not have experimental data necessary to calculate log D_{mw}, and therefore for such compounds only the nominal concentration, EC_{tot} could be used to assess the quality of the model. The nominal concentration is related to the intrinsic effect concentration by eq 13 (full derivation in SI-2).

$$\log(1/EC_{tot}) = \log(1/EC_m) - \log(m_{lip}/V_w + 1/D_{mw}) \quad (13)$$

where m_{lip}/V_w is the lipid mass to water volume ratio in the in vitro assay (given in Table 1). As the lipid mass to water volume ratio can differ between different batches of chromatophores, m_{lip}/V_w should be given for each measurement. As a first test, the same 17 compounds used to model EC_w were used to fit a linear regression equivalent to eq 13,

$$\log(1/EC_{tot}) = \beta_1 \cdot \log(m_{lip}/V_w + 1/D_{mw}) + \beta_2 \cdot \Delta G_{well\text{-}barrier,A^-} + \beta_0 \quad (14)$$

where β_1 and β_2 are the regression coefficients of the two descriptors and β_0 is the intercept. The results are given in Table 5 with the regression models R7−R10. The same choice of descriptors for log D_{mw} has been used as for models R1−R4 for EC_w. The descriptors used in R5 and R6 were not tested any more because of their obvious lack of fit.

As compared to the models for EC_w, the correlations tended to decrease slightly for EC_{tot} while the SD was rather lower. This is explicable as EC_{tot} covers an about one log unit smaller range of concentrations than EC_w (from 3.1 to 7.2 M instead of from 3.1 to 8.1 M). However, as the changes are relatively small, the important conclusion from comparing model R1 with model R7 is that the concept of dividing EC_{tot} into EC_w and EC_m seems to be applicable to the test system used in this study. This applies also for the corresponding models with calculated descriptors.

When all 23 compounds of the training set were used (right half-of Table 5), the correlations are comparable for all models; however, the SD increased slightly. Three compounds had errors above one log unit (in model R8 with log D_{mw} from experimental K_{ow} and pK_a), namely, 4NTFB (+1.25), CCCP (−1.10), and warfarin (+1.12). 4NTFB was like dino2terb and dino4terb, a compound without experimental K_{ow}, and a calculated K_{ow} had to be used, which might explain the high error. Leaving out the three compounds without experimental K_{ow} of model R8 would result in R^2 = 0.83, Q^2 = 0.77, SD = 0.56, and SD_{cv} = 0.67. It is remarkable that no obvious outliers occurred and that even 246TriNP (with a pK_a of 0.38 could be fitted with these descriptors). The coefficients of eqs R7−R10 are given in Table 5, and the plots of predicted vs experimental values are given in SI-9.

Nonlinear Models for EC_w. The cyclic process shown in Figure 1 can be fully described by a transport model with the already introduced descriptors pK_a, $K_{mw,HA}$, K_{mw,A^-}, k_{HA}, and k_{A^-} for compounds with pK_a < pH (*22*). However, for uncouplers with pK_a > pH, an additional effect needs to be considered, namely, the formation of dimers of the neutral and the charged species, AHA^- dimers also called heterodimers (*16, 17, 22, 76, 77*). In a first step, a model neglecting AHA^- dimers was tested using the functions given in SI-4. The values of k_{HA} and k_{A^-} were calculated by fitting the seven experimentally derived values from ref (*22*) with $\Delta G_{well\text{-}barrier,HA}$ and $\Delta G_{well\text{-}barrier,A^-}$, respectively. The plot of predicted vs experimental values (SI-10) showed a large number of compounds

Table 5. Quality of Fit of Linear Regression Models for EC_{tot} at pH 7 Using the Descriptor $\Delta G_{well\text{-}barrier,A^-}$ and Different Descriptors for the Uptake as Defined in Eqs 2 and 13, Respectively[a]

		EC_{tot} ($n = 17$)							EC_{tot} ($n = 23$)							EC_{tot} test ($n = 12$)	
	descriptor for uptake	R^2	Q^2	SD	SD_{cv}	β_1	β_2	β_0	R^2	Q^2	SD	SD_{cv}	β_1	β_2	β_0	Q_{ext}^2	SD_{ext}
R7	log D_{mw} (exp)	0.84	0.78	0.48	0.57	−1.12	−0.28	5.15								0.79[b]	0.63[b]
R8	log D_{mw} [pK_a (exp) and K_{ow} (exp)]	0.78	0.70	0.56	0.66	−0.63	−0.28	6.73	0.78	0.72	0.65	0.74	−0.52	−0.31	7.15	0.64	0.67
R9	log D_{mw} (QC method)	0.71	0.53	0.64	0.82	−0.92	−0.25	5.41	0.75	0.64	0.69	0.84	−0.84	−0.27	5.80	0.70	0.62
R10	log D_{mw} (empirical)	0.74	0.61	0.61	0.75	−0.56	−0.27	6.81	0.75	0.66	0.69	0.81	−0.38	−0.30	7.47	0.65	0.66

[a] The columns β_1, β_2, and β_0 denote the regression coefficients for the descriptors $\log(m_{lip}/V_w + 1/D_{mw})$, $\Delta G_{well\text{-}barrier,A^-}$, and the intercept, respectively. [b] The external test set is only possible with compounds having experimental log D_{mw} (exp), that is, $n = 8$.

very close to the unit slope, and no compound predicted more than half a log unit too high, while six compounds with pK_a close or above the experimental pH had predicted effect concentrations being several orders of magnitude too low (24DCP, 35DCP, 245TriCP, 345TriCP, and 4NP), which drastically reduced the correlation.

Therefore, an approximation for the translocation rate constant of AHA^- dimers, k_{AHA^-} was needed. A very simple approach was chosen based on the observation that the simple regression model (R1) showed a very good performance without considering AHA^- dimers. Therefore, the information in the descriptors of the linear model seems also to contain k_{AHA^-}, and a likely hypothesis is that k_{AHA^-} must also be related to $\Delta G_{well\text{-}barrier,A^-}$. The four experimentally derived values for k_{AHA^-} from ref (*22*) were fitted with $\Delta G_{well\text{-}barrier,A^-}$ and used as an additional parameter in the transport model. This measure improved the correlation between the predicted and the experimental log(1/EC_w) to an R^2 of 0.85 and an SD of 0.66. Replacing the experimental values of pK_a, $K_{mw,HA}$, and K_{mw,A^-} with calculated values resulted in correlations comparable to the linear models R2−R4.

As at present, the nonlinear models did not show a better performance than the linear models, no additional tests with the test set and the EC_{tot} were made. Further improvements are possible by more closely examining the nature of the dimers either by gaining additional experimental data or by more extensive computations; that is, there might be compounds that have a low $\Delta G_{well\text{-}barrier,A^-}$ but nevertheless have a very low or even zero k_{AHA^-} because they do not form dimers for steric reasons.

Alternative Approaches to Calculate $\Delta G_{well\text{-}barrier,A^-}$. In the Materials and Methods section, possible simplifications to approximate the translocation rate constant of the charged species were described. They were based on the idea to use reference solvents representing the environment of the energy barrier and the energy well. Two partition coefficients were calculated with COSMOtherm: $K_{1,9decadiene\text{-}water}$ (*72*) and $K_{1,9decadiene\text{-}octanol}$, as a descriptor that should reflect the environment of the energy well more closely (although being of a theoretical nature as the two phases are probably miscible). The partition coefficients of both the neutral and the charged species were calculated. The following correlations of the logarithm of the partition coefficients with $\Delta G_{well\text{-}barrier,A^-}$ were observed as follows: 0.86 (log $K_{1,9decadiene\text{-}octanol,A^-}$), 0.97 (log $K_{1,9decadiene\text{-}water,A^-}$), 0.16 (log $K_{1,9decadiene\text{-}octanol,HA}$), and 0.29 (log $K_{1,9decadiene\text{-}water,HA}$).

When the two partition coefficients log $K_{1,9decadiene\text{-}octanol,A^-}$ and log $K_{1,9decadiene\text{-}water,A^-}$ were tested as descriptors in the regression model with experimental log D_{mw}, that is, the equivalent of R1, the correlation decreased slightly in the case of $K_{1,9decadiene\text{-}octanol,A^-}$ ($R^2 = 0.89$, SD = 0.52) and increased slightly in the case of log $K_{1,9decadiene\text{-}water,A^-}$ ($R^2 = 0.92$, SD = 0.44) with the regression coefficients very similar to R1 taking the values of 0.99 for log D_{mw}, 0.36 for log $K_{1,9decadiene\text{-}water,A^-}$, and an interecept of 4.89. However, given the size of the data set, these differences are too small to be considered as improvements, but these results are certainly worth further investigations also for experimental studies because $K_{1,9\text{-}decadiene\text{-}water}$ would be easier to measure than membrane permeability.

A last but important observation is that there was no correlation between $\Delta G_{well\text{-}barrier,A^-}$ and $\Delta G_{well\text{-}barrier,HA}$ (both calculated with COSMOmic) with an r^2 of 0.03.

Test Set and Final Model with Applicability Domain. Once the final choice of descriptors was made, the models were applied to predict the test sets. The test set for EC_w is rather small ($n = 8$); therefore, the Q_{ext}^2 values are only crude indicators. The values are given in Table 4, with no Q_{ext}^2 being substantially lower than the Q^2, and the SD_{ext} is only slightly larger than the SD_{cv}, which is reassuring. In the case of the QC method (R3 of Table 4), the Q_{ext}^2 was actually higher than the R^2 because one compound in the training set had an extremely high residual (warfarin, +2.08) due to the serious overestimation of log D_{mw}.

In the case of R1, the predicted values of the test set compounds are also shown in Figure 5, while the corresponding plots for the linear regression models R2−R6 are shown in SI-8, and all statistical quality criteria are given in Table 4. The same observations could be made for EC_{tot}, with the values of the regression coefficients and the statistical quality criteria given in Table 5 and the plots of R7−R10 given in SI-9.

For future predictions for EC_w, three final models are recommended. In case experimental data for D_{mw} are available, regression model R1 with all available compounds takes the values given by eq 15

$$\log(1/EC_w) = 1.15(\pm 0.24)\cdot \log D_{mw} - 0.29(\pm 0.08)\cdot \Delta G_{well\text{-}barrier,A^-} + 5.13(\pm 1.37) \quad (15)$$

with $n = 25$, $R^2 = 0.90$, $Q^2 = 0.87$, SD = 0.50, $SD_{cv} = 0.56$, and $F = 97.32$, and in case, the D_{mw} needs to be estimated by octanol–water partition coefficients (calculated or experimental). Equation 16 applies

$$\log(1/EC_w) = 0.78(\pm 0.27)\cdot \log D_{mw} - 0.35(\pm 0.12)\cdot \Delta G_{well\text{-}barrier,A^-} + 7.39(\pm 1.61) \quad (16)$$

with $n = 25$, $R^2 = 0.79$, $Q^2 = 0.73$, SD = 0.72, $SD_{cv} = 0.82$, and $F = 40.72$ or eq 17 for descriptors calculated entirely with the QC method

$$\log(1/EC_w) = 1.56(\pm 0.64)\cdot \log D_{mw} - 0.20(\pm 0.15)\cdot \Delta G_{well\text{-}barrier,A^-} + 3.42(\pm 3.04) \quad (17)$$

with $n = 25$, $R^2 = 0.74$, $Q^2 = 0.67$, SD = 0.80, $SD_{cv} = 0.92$, and $F = 31.58$.

For the QC method, again, a single compound (warfarin) reduced the R^2 and increased the SD. This was caused by the

severe overestimation of log D_{mw} of 1.5 log units. Without this compound, the measures of fit and predictivity would change to $R^2 = 0.81$, $Q^2 = 0.73$, SD = 0.68, and $SD_{cv} = 0.80$. The regression coefficients of eq 17 changed only little after removing warfarin, but their confidence intervals were strongly reduced.

Concerning the applicability domain, we suggest that the models presented in this study are not limited to the chemical classes shown in SI-3 but apply to any weak acid with the limitations described in the following section. The models have a very low number of descriptors; therefore, no statistical approach to define the applicability domain (*78*) was chosen, but some empirical guidelines to the limitations are given. The limiting factors of the model are pK_a, solubility, and size. The pK_a is limiting because if either the HA or the A^- concentration in the membrane is extremely low, the cyclic process shown in Figure 1 is not efficient any more. There is no generally agreed rule at which upper and lower pK_a uncoupling becomes irrelevant, but the compounds described in the literature (*9, 14, 15, 37*) almost all range between 3 and 9. As an example, we extended the sensivity analysis of CCCP shown in Figure 4b) for the pK_a up to 3 units above the pK_a of CCCP (from 5.95 to 9). The transport model predicts a reduction of the 1.8 log units. Therefore, a compound with a pK_a of 10 or higher is not likely to be an uncoupler even if the $\Delta G_{well\text{-}barrier,A^-}$ is quite low, but nevertheless, the regression models should not be applied outside of the pK_a window of $3 < pK_a < 9$, because they do not contain a descriptor for speciation (while the nonlinear model would account for such effects).

The second limiting factor is solubility. It is clear that highly insoluble compounds will be difficult to measure in the short equilibration time of a typical Kinspec experiment, but they nevertheless might act as uncouplers in the environment especially in mixtures. No cutoff value can be given for solubility, but it is recommended to compare the predicted EC_w for uncoupling with the aqueous solubility. If the solubility is lower, the same argumentation than the one that has been made for baseline toxicants applies (*79*).

The last factor that could cause misleading predictions is a molecule's size. This factor is limiting because $\Delta G_{well\text{-}barrier,A^-}$ is calculated with a membrane model (COSMOmic) that must have a size limit (e.g., it would not make sense to make calculations for a compound that is larger than half of the membrane). Currently, COSMOmic does not account for lateral pressures (*28, 29*), which become stronger for large compounds. This might lead to overestimation of both the partition coefficients and the $\Delta G_{well\text{-}barrier,A^-}$, even for intermediately large molecules. The largest molecule studied in this study was warfarin (length $\approx$ 12 Å). Therefore, larger molecules than warfarin should not be predicted with the present model.

The value of $\Delta G_{well\text{-}barrier,A^-}$ itself is not limiting, that is, as the coefficient is negative, the predicted effect concentration becomes very large (i.e., low toxicity) with very high values of $\Delta G_{well\text{-}barrier,A^-}$, and eventually, the uncoupling effect is less relevant than baseline toxicity, which reflects the situation in real organisms where baseline toxicity always competes with more specific MOAs. As baseline QSARs are available for almost any endpoint and also for Kinspec (*25*), high values $\Delta G_{well\text{-}barrier,A^-}$ leading to very low uncoupling activity only require comparison with the predicted baseline toxicity and do not fall out of the applicability domain of the models proposed in this study. Neither can low values as with fluazinam ($\Delta G_{well\text{-}barrier,A^-} = 1.9$), which is almost the most potent uncoupler known, be modeled sufficiently well with the regression models.

Discussion

The main goal of this study was to examine if it is possible to model the activity of uncouplers of different chemical classes in a single model. Although this is a fundamental premise of MOA-based toxicity QSARs, such models are—apart from baseline toxicity (narcosis)—still rare. Although the present data set is small, it is extremely diverse, containing compounds from five different chemical classes. The good fits of the models with experimental data for the uptake are strong indicators that this goal can be reached for uncouplers.

The most important descriptor for the intrinsic activity is the translocation rate constant (or permeability) of the anion. In this perspective, uncoupling is a singular MOA and easily misunderstood; that is, if one wants to model the permeability of a weak acid through the membrane, it is perfectly legitimate to neglect the permeability of the charged species. This applies also to uncouplers as even for the most potent uncouplers the translocation rate constant of the charged species is at least an order of magnitude smaller than for the neutral species (*20*) and for less potent uncouplers the difference is 3 orders of magnitude or more (*22*). However, as the activity of uncouplers is determined by the limiting step in the cyclic process shown in Figure 1a, the translocation rate constant of the charged species is exactly the bottleneck one needs to describe.

In this study, the descriptor calculation methods were chosen and optimized with additional data (the "descriptor calibration sets" given in SI-5 plus Table 1 with the results shown in Table 3). As toxicity data sets are often very small, this approach is very useful. If the descriptor calculation methods had been chosen and optimized with the small training set of this study, an overfitted model would have been very likely. Of course, this approach is feasible only for MOA where the mechanism and the required descriptors are known. An additional advantage of the present data set is that it allowed the use of experimental data for the uptake. This gives the modeler clear indications where the main sources of error are. As an example, in the usual scenario of making a toxicity QSAR, one would continue to search for descriptors for regression model R3 or R4 (the models with only calculated data) as the fit is not yet very "impressive", while in this study, model R1 with experimental data for the uptake showed that two descriptors are sufficient and the most important point is to improve the quality of the descriptor calculation methods.

However, the quality of the descriptor calculations is also the weakness of the presented models. The quality of the models based completely on calculated descriptors is not very high. As mentioned in the Introduction, class-based QSARs exist for almost all of the compound classes tested in this study. It is likely that these class-specific QSARs have smaller errors. For example, for the 21 phenols measured in earlier Kinspec studies, we proposed a model with an R^2 of 0.94 (*34*). Therefore, the question is justified if there is an advantage of working with such a general model. We see three main advantages: First, the general model presented here should be applicable to all weak acids, and there is a considerable number of weak acids from different chemical classes for which no class-specific QSARs exist at all. As a matter of fact, our earlier model for phenols (*34*) showed large errors when it was used for other chemical classes, especially NH-acidic compounds, a result that we attribute to large errors in the empirical method used to calculate the descriptor for the charged species. The ability to generalize is an important advantage; for example, if a database is screened for compounds that should be tested experimentally, a generally applicable model is more important than having models with

high fit but limited applicability. Second, general models have a more "collaborative spirit"; for example, if someone else needs to model permeability of charged compounds, the present work can be helpful while the models presented here can immediately benefit from other groups improving membrane−water partition models. Third, the lower fit is caused by the limits to predict log D_{mw}, while the intrinsic activity determined by $\Delta G_{well\text{-}barrier,A^-}$ can be predicted very well. The intrinsic activity directly relates to the toxic ratio, TR (or excess toxicity), given by 18. The TR is the ratio of the EC_w for baseline toxicity to the experimental EC_w, which can be orders of magnitude lower for specific MOA like uncoupling (ref *25* describes the determination of baseline toxicity with Kinspec):

$$TR = EC_w(\text{baseline})/EC_w(\text{specific}) \quad (18)$$

Indeed, for the 25 uncouplers with data on EC_w, the logarithm of the TR correlates well with $\Delta G_{well\text{-}barrier,A^-}$

$$\log(TR) = -0.31(\pm 0.08)\cdot\Delta G_{well\text{-}barrier,A^-} + 5.38(\pm 0.99) \quad (19)$$

where $n = 25$, $R^2 = 0.71$, $Q^2 = 0.66$, SD = 0.54, $SD_{cv} = 0.59$, and $F = 57.33$ and with the SD being less than a factor of 4, and only a single compound with a residual above 1 log unit (dino4terb). Therefore, the descriptor $\Delta G_{well\text{-}barrier,A^-}$ is an ideal criterion for screening of databases as there is a high interest to detect compounds with excess toxicity. We developed some simple criteria to detect potential uncouplers and applied these critieria for a database of industrial compounds, namely, the EU High and Low Production Volume Chemicals, HPVC and LPVC, respectively (Spycher, S., Netzeva, T. I., and Escher, B. I. Unpublished results). In the next sections, the results of the present study will be discussed in more detail.

One of the aims of this study was to examine the lower pK_a limit for uncouplers as it is known that compounds like 2,4,6-trinitrophenol (246TriNP) with a pK_a of 0.38 do not cause uncoupling in mitochondria (*80*), while the lowest pK_a of known phenolic uncouplers is about 4 (e.g., dino2terb, dinoseb, and 24DNP). Three phenols with $pK_a < 4$ were tested in this study, and all of them showed high activity in the Kinspec system, including 2,4,6-trinitrophenol. This discrepancy has been observed earlier with submitochondrial particles (inverted mitochondrial membranes) where an effect can be measured with 2,4,6-trinitrophenol as well. In the case of submitochondrial particles, this discrepancy has been elegantly explained by McLaughlin et al. who showed that the inversion of the potential from −175 mV for mitochondria to +175 for submitochondrial particles (plus the pH change from 8.3 to 6.3) causes this difference (*81*). As the chromatophores in Kinspec have no potential before the experiment starts, they lie somewhere in between mitochondria and submitochondrial particles. Thus, it is not possible to infer the activity in mitochondria directly from chromatophores for compounds with $pK_a < 4$. Therefore, a limitation of the linear models of this study needs to be defined, and we suggest that they should not be used for $pK_a < 3$. This limitation could probably be overcome with transport models, which account for potential and pH differences of the membranes as has been shown by Benz and MacLaughlin for the case of FCCP (*18*). Such models allow the direct calculation of the proton flux caused by a weak acid, and it should be possible to extend results determined with Kinspec to such models. However, as compared to the easily interpretable linear regression models developed in this study, they are rather complex, but introducing such models into QSAR modeling is an important area of future research. The conclusion from measuring the three compounds with $pK_a < 4$ is that both the linear model and the Kinspec describe the indispensable precondition for uncoupling (i.e., high enough translocation rate constant of the anion), but toward lower pK_a, they do not consider other limiting effects and might overestimate the real uncoupling activity. They are therefore too conservative, which is less serious than the underestimation of the activity.

When it comes to the assessment of the quality of current descriptor calculation methods, it is clear that more research is needed to improve the calculation of membrane−water partition coefficients, be it by empirical or QC methods [it should be mentioned that the QC method tested (COSMOmic) is still at an early stage of development]. The need to improve predictions for charged species is especially urgent as charged species are not only of interest for the particular MOA of uncouplers but also of general interest in many areas including toxico- and pharmacokinetics. The carefully reviewed compilation of liposome−water partition coefficients given in Table 1 and SI-5 can contribute to this task (another thoroughly evaluated compilation of partition coefficients of neutral and charged compounds mainly determined with pH metric titration can be found in ref *82*).

The observation that the fit of the seven experimentally determined translocation rate constants of phenolates with $\Delta G_{well\text{-}barrier,A^-}$ has about the same SD as the regression models with experimental data for the uptake (R1, R7, and eq 15) is reassuring. If the regression models had a lower SD, it would have been an indicator of overfitting, while a higher SD would have indicated that either the good fit of the experimental translocation rate constants applies only to the phenolates tested, that other relevant descriptors are missing, or that the functional form is not linear.

As a matter of fact, the fit of linear models is surprisingly good given that they do not account for the speciation (i.e., the fractions of HA and A^-, respectively) nor for AHA^- dimer formation, which cannot be neglected in transport models (at least for compounds with $pK_a >$ pH). It appears that the two descriptors used for the linear models contain the information of the dimer formation. If uncoupling would be related to uptake and translocation rate constant of the charged species alone, then regression models R5 and R6, which look only at the partition coefficient (R5) or the pH-dependent uptake of the charged species (R6), should work better. Therefore, it appears that the descriptor log D_{mw} also contains information that correlates with the formation and transfer of AHA^- dimers. This hypothesis is corroborated by the observation that the compounds with large errors in R6 are the same compounds as the ones with large errors in the nonlinear model with a translocation rate constant of AHA^- dimer set to zero. It is possible that the linear regression models proposed in this study generalize less well than real mechanistic models like the transport models tested, but for the 35 compounds examined in this study, the linear models did not have extreme outliers. Therefore, the discussed inconsistency of using no descriptor for the formation and transfer of AHA^- dimers does not seem a serious weakness, although it is known to be an important phenomenon. The lower fit of the models based only on calculated descriptors is consistent with the errors of calculated partition coefficients for neutral and charged species.

A large section of this study deals with the issue of comparing the results of aqueous concentrations, EC_w, and nominal concentrations, EC_{tot}. While at a first look the two effect concentrations appear very similar ($r^2 = 0.93$), it is important

to note that for compounds with high values for log D_{mw}, the solution in the test system tends to be depleted and EC_w is consistently lower than EC_{tot}. As a consequence, EC_w can be up to 40 times lower than EC_{tot} and simply taking EC_{tot} would lead to reduced value of the toxic ratios. The distinction is especially important if in vitro results are related to in vivo data with measured aqueous concentrations. A nice result is that the fit of the regression models for EC_w and EC_{tot} is comparable. The fact that the SDs of R1 and R7 are almost the same can be considered as an indirect proof of the assumption that, first, the liposome–water partition coefficients approximate the chromatophore–water partition coefficients well and, second, that there is no substantial partitioning into other compartments of the chromatophores (e.g., proteins) or even if then the relative differences between the compounds are the same.

A very low correlation was observed between the free energy, $\Delta G_{well\text{-}barrier,A^-}$ and the equivalent for the neutral species $\Delta G_{well\text{-}barrier,HA}$. This is a serious limitation if simpler empirical methods usually developed for neutral species should be used. While it is possible to develop class-specific uncoupler QSARs with the use of descriptors of the neutral species, the goal of developing class-independent or "global" QSARs apparently can only be reached with descriptors for the charged species.

Conclusions

In this study, a series of highly diverse uncouplers could be modeled with only two descriptors; thus, it appears possible to reach the goal of class-independent QSARs for the uncoupling activity of weak acids. The preconditions for such global QSARs for uncouplers are, first, the successful calculation of descriptors for the translocation rate constant of the charged species and, second, the calculation of membrane−water partition coefficients of the neutral and the charged species.

The first precondition could be met by using a novel membrane model that allows the calculation of the distribution of compounds within the membrane and, thus, to calculate the heights of energy barriers encountered by molecules crossing the interior of the membrane. The energy barrier of the anion, $\Delta G_{well\text{-}barrier,A^-}$ is the rate-limiting step of the uncoupling effect, and it could be shown that this descriptor alone already covers the largest portion of the variability of the intrinsic activity. The descriptor $\Delta G_{well\text{-}barrier,A^-}$ can also be related to the excess toxicity (toxic ratio) caused by the uncoupling activity of weak acids; therefore, $\Delta G_{well\text{-}barrier,A^-}$ is also the most important descriptor to distinguish weak acids with the potential to uncouple from weak acids without this potential.

With the data set used in this study, it was possible to take advantage of experimentally determined descriptors. For the case of uncouplers, it was possible to show that to increase the quality of QSAR models based on calculated descriptors, it is not necessary to include additional descriptors but to improve the predictions for the uptake into membranes especially for charged compounds. Such a conclusion is very hard to make if no experimental descriptors are available; therefore, such close collaborations between modeling and experimental studies are highly desirable.

Some additional work is necessary to improve the understanding of the effects of speciation, that is, to define the lower and upper limits that the pK_a uncouplers can have or even better to find models that account for speciation effects. If these challenges can be overcome, then the door would be opened to pH-dependent QSAR models for uncouplers and in our opinion also to in vitro–in vivo correlations and interspecies interpolations.

Acknowledgment. We thank Nadine Bramaz for experimental assistance, Philippe Perisset for programming the Kinspec system, Johann Gasteiger and Dimitar Hristozov of the Computer-Chemie-Centrum of the University of Erlangen for support with software and in programming questions, Björn Reineking of ETH Zürich for help in implementing the transport model in R, and Andreas Klamt and Uwe Huniar from COSMOlogic GmbH&CoKG for help with COSMOmic and COSMOtherm. We thank the 3R-Research Foundation for funding (3R-Project 95-05). Disclaimer: This paper reflects the views of the authors as individual scientists and does not necessarily represent a position of the European Commission.

Supporting Information Available: Function for evaluation of membrane potential decay observed in the Kinspec system (SI-1), derivation of expressions for transforming nominal effect concentrations into aqueous effect concentrations and intrinsic concentrations (SI-2), 35 structures of the data set studied (SI-3), equations of the nonlinear model implemented in the R environment and language (SI-4), additional experimental data used to select descriptor calculation methods (SI-5), SMILES codes of structures of (A) Table 1 and (B) SI-5 (SI-6), calculated descriptors of 35 uncouplers (SI-7), plots of regression models R2−R6 for EC_w (SI-8), plots of regression models R7−R10 for EC_{tot} (SI-9), and plot of nonlinear model with different assumptions for AHA^- dimers (SI-10). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) OECD (2004) The report from the expert group on (quantitative) structure−activity relationships [(Q)SARS] on the principles for the validation of (Q)SARs, OECD Environment Health and Safety Publications Series on Testing and Assessment, No. 49.
(2) McLaughlin, S. G. A., and Dilger, J. P. (1980) Transport of protons across membranes by weak acids. *Physiol. Rev. 60*, 825–863.
(3) Terada, H. (1990) Uncouplers of oxidative phosphorylation. *Environ. Health Perspect. 87*, 213–218.
(4) Wallace, K. B., and Starkov, A. A. (2000) Mitochondrial targets of drug toxicity. *Annu. Rev. Pharmacol. Toxicol. 40*, 353–388.
(5) Mitchell, P. (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemiosmotic type of mechanism. *Nature 191*, 144–148.
(6) Rosier, R. N. (1979) Calcium transport in vesicles energized by cytochrome oxidase. Report UR-3490-1549, Vols. *1–2*.
(7) Hansch, C., Kiehs, K., and Lawrence, G. L. (1965) The role of substituents in the hydrophobic bonding of phenols by serum and mitochondrial proteins. *J. Am. Chem. Soc. 87*, 5770–5773.
(8) Miyoshi, H., and Fujita, T. (1988) Quantitative analyses of the uncoupling activity of substituted phenols with mitochondria from flight muscles of house flies. *Biochim. Biophys. Acta 935*, 312–321.
(9) Miyoshi, H., Tsujishita, H., Tokutake, N., and Fujita, T. (1990) Quantitative analysis of uncoupling activity of substituted phenols with a physicochemical substituent and molecular parameters. *Biochim. Biophys. Acta 1016*, 99–106.
(10) Schüürmann, G., Somashekar, R. K., and Kristen, U. (1996) Structure-activity relationships for chloro- and nitrophenol toxicity in the pollen tube growth test. *Environ. Toxicol. Chem. 15*, 1702–1708.
(11) Argese, E., Bettiol, C., Giurin, G., and Miana, P. (1999) Quantitative structure-activity relationships for the toxicity of chlorophenols to mammalian submitochondrial particles. *Chemosphere 38*, 2281–2292.
(12) Tollenaere, J. P. (1973) Structure-activity relations of three groups of uncouplers of oxidative phosphorylation. Salicylanilides, 2-trifluoromethylbenzimidazoles, and phenols. *J. Med. Chem. 16*, 791–796.
(13) Guo, Z. J., Miyoshi, H., Komyoji, T., Haga, T., and Fujita, T. (1991) Quantitative analysis with physicochemical substituent and molecular parameters of uncoupling activity of substituted diarylamines. *Biochim. Biophys. Acta 1059*, 91–98.
(14) Terada, H., Goto, S., Yamamoto, K., Takeuchi, I., Hamada, Y., and Miyake, K. (1988) Structural requirements of salicylanilides for uncoupling activity in mitochondria: Quantitative analysis of structure-uncoupling relationships. *Biochim. Biophys. Acta 936*, 504–512.
(15) Gange, D. M., Donovan, S., Lopata, R. J., and Henegar, K. (1995) The QSAR of insecticidal uncouplers. *ACS Symp. Ser. 606*, 199–212.

(16) Cohen, F. S., Eisenberg, M., and McLaughlin, S. (1977) The kinetic mechanism of action of an uncoupler of oxidative phosphorylation. *J. Membr. Biol. 37*, 361–396.
(17) Dilger, J., and McLaughlin, S. (1979) Proton transport through membranes induced by weak acids: A study of two substituted benzimidazoles. *J. Membr. Biol. 46*, 359–384.
(18) Benz, R., and McLaughlin, S. (1983) The molecular mechanism of action of the proton ionophore FCCP (carbonylcyanide *p*-trifluoromethoxyphenylhydrazone). *Biophys. J. 41*, 381–398.
(19) Kasianowicz, J., Benz, R., and McLaughlin, S. (1984) The kinetic mechanism by which CCCP (carbonyl cyanide *m*-chlorophenylhydrazone) transports protons across membranes. *J. Membr. Biol. 82*, 179–190.
(20) Kasianowicz, J., Benz, R., and McLaughlin, S. (1987) How do protons cross the membrane-solution interface? Kinetic studies on bilayer membranes exposed to the protonophore S-13 (5-chloro-3-tert-butyl-2′-chloro-4′nitrosalicylanilide). *J. Membr. Biol. 95*, 73–89.
(21) Benz, R., Laeuger, P., and Janko, K. (1976) Transport kinetics of hydrophobic ions in lipid bilayer membranes. Charge-pulse relaxation studies. *Biochim. Biophys. Acta 455*, 701–720.
(22) Escher, B. I., Hunziker, R., Schwarzenbach, R. P., and Westall, J. C. (1999) Kinetic model to describe the intrinsic uncoupling activity of substituted phenols in energy transducing membranes. *Environ. Sci. Technol. 33*, 560–570.
(23) Escher, B. I., Snozzi, M., Häberli, K., and Schwarzenbach, R. P. (1997) A new method for simultaneous quantification of uncoupling and inhibitory activity of organic pollutants in energy-transducing membranes. *Environ. Toxicol. Chem. 16*, 405–414.
(24) Escher, B. I., Snozzi, M., and Schwarzenbach, R. P. (1996) Uptake, speciation, and uncoupling activity of substituted phenols in energy transducing membranes. *Environ. Sci. Technol. 30*, 3071–3079.
(25) Escher, B. I., Eggen, R. I. L., Schreiber, U., Schreiber, Z., Vye, E., Wisner, B., and Schwarzenbach, R. P. (2002) Baseline toxicity (narcosis) of organic chemicals determined by in vitro membrane potential measurements in energy-transducing membranes. *Environ. Sci. Technol. 36*, 1971–1979.
(26) Escher, B. I., and Sigg, L. (2004) Chemical speciation of organics and of metals at biological interphases. In *IUPAC Series on Analytical and Physical Chemistry of Environmental Systems* (Buffle, J., and van Leuwen, H. P., Eds.) pp 205–269, John Wiley, Chichester.
(27) Lessard, J. G., and Fragata, M. (1986) Micropolarities of lipid bilayers and micelles. 3. Effect of monovalent ions on the dielectric constant of the water-membrane interface of unilamellar phosphatidylcholine vesicles. *J. Phys. Chem. 90*, 811–817.
(28) Xiang, T. X., and Anderson, B. D. (1994) Molecular distributions in interphases: Statistical mechanical theory combined with molecular dynamics simulation of a model lipid bilayer. *Biophys. J. 66*, 561–572.
(29) Gullingsrud, J., and Schulten, K. (2004) Lipid bilayer pressure profiles and mechanosensitive channel gating. *Biophys. J. 86*, 3496–3509.
(30) Escher, B. I., and Schwarzenbach, R. P. (1996) Partitioning of substituted phenols in liposome-water biomembrane-water, and octanol-water systems. *Environ. Sci. Technol. 30*, 260–270.
(31) Avdeef, A. (2001) Physicochemical profiling (solubility, permeability and charge state). *Curr. Top. Med. Chem. 1*, 277–351.
(32) Smejtek, P., and Wang, S. (1993) Distribution of hydrophobic ionizable xenobiotics between water and lipid membranes: Pentachlorophenol and pentachlorophenate. A comparison with octanol-water partition. *Arch. Environ. Contam. Toxicol. 25*, 394–404.
(33) Avdeef, A., Box, K. J., Comer, J. E. A., Hibbert, C., and Tam, K. Y. (1998) pH-metric logP 10. Determination of liposomal membrane-water partition coefficients of ionizable drugs. *Pharm. Res. 15*, 209–215.
(34) Spycher, S., Escher, B. I., and Gasteiger, J. (2005) A quantitative structure−activity relationship model for the intrinsic activity of uncouplers of oxidative phosphorylation. *Chem. Res. Toxicol. 18*, 1858–1867.
(35) Escher, B. I., and Schwarzenbach, R. P. (2002) Mechanistic studies on baseline toxicity and uncoupling of organic compounds as a basis for modeling effective membrane concentrations in aquatic organisms. *Aquat. Sci. 64*, 20–35.
(36) Buechel, K. H., Draber, W., Trebst, A., and Pistorius, E. (1966) Inhibition of photosynthetic reactions in isolated chloroplasts by benzimidazole-type herbicides and their structure-activity relation, especially regarding the distribution coefficient and the pKa value. *Z. Naturforsch., B: Anorg. Chem., Org. Chem., Biochem., Biophys., Biol. 21*, 243–254.
(37) Jones, O. T., and Watson, W. A. (1967) Properties of substituted 2-trifluoromethylbenzimidazoles as uncouplers of oxidative phosphorylation. *Biochem. J. 102*, 564–573.
(38) Guo, Z. J., Miyoshi, H., Komyoji, T., Haga, T., and Fujita, T. (1991) Uncoupling activity of a newly developed fungicide, fluazinam [3-chloro-N-(3-chloro-2,6-dinitro-4-trifluoromethylphenyl)-5-trifluoromethyl-2-pyridinamine]. *Biochim. Biophys. Acta, Bioenerg. 1056*, 89–92.
(39) Brandt, U., Schubert, J., Geck, P., and Von Jagow, G. (1992) Uncoupling activity and physicochemical properties of derivatives of fluazinam. *Biochim. Biophys. Acta 1101*, 41–47.
(40) Robinson, R. A. (1967) Dissociation constants of some substituted nitrophenols in aqueous solution at 25 Deg. *J. Res. Natl. Bureau Stand, Sect. A: Phys. Chem. 71*, 385–389.
(41) Ottiger, C., and Wunderli-Allenspach, H. (1997) Partition behavior of acids and bases in a phosphatidylcholine liposome-buffer equilibrium dialysis system. *Eur. J. Pharm. Sci. 5*, 223–231.
(42) Kuramochi, H., Maeda, K., and Kawamoto, K. (2004) Water solubility and partitioning behavior of brominated phenols. *Environ. Toxicol. Chem. 23*, 1386–1393.
(43) Escher, B. I., Schwarzenbach, R. P., and Westall, J. C. (2000) Evaluation of liposome-water partitioning of organic acids and bases. 1. Development of a sorption model. *Environ. Sci. Technol. 34*, 3954–3961.
(44) Escher, B. I., Hunziker, R. W., and Schwarzenbach, R. P. (2001) Interaction of phenolic uncouplers in binary mixtures: Concentration-additive and synergistic effects. *Environ. Sci. Technol. 35*, 3905–3914.
(45) Abraham, M. H., Du, C. M., and Platts, J. A. (2000) Lipophilicity of the nitrophenols. *J. Org. Chem. 65*, 7114–7118.
(46) COSMOlogic GmbH&CoKG (2006) COSMOtherm, C2.1 Release 01.06, http://www.cosmologic.de.
(47) University of Georgia (2006) SPARC—Performs automated reasoning in chemistry, 3.1, http://ibmlc2.chem.uga.edu/sparc, accessed on 02-08-2007.
(48) Klamt, A., Eckert, F., Diedenhofen, M., and Beck, M. E. (2003) First principles calculations of aqueous pK_a values for organic and inorganic acids using COSMO-RS reveal an inconsistency in the slope of the pK_a scale. *J. Phys. Chem., A 107*, 9380–9386.
(49) CORINA, 3.21 MolNet GmbH, http://www.mol-net.de.
(50) Ahlrichs, R., Baer, M., Haeser, M., Horn, H., and Koelmel, C. (1989) Electronic structure calculations on workstation computers: The program system TURBOMOLE. *Chem. Phys. Lett. 162*, 165–169.
(51) COSMOlogic (2005) Turbomole 5.8, http://www.cosmologic.de/QuantumChemistry/main_turbomole.html.
(52) Becke, A. D. (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys. 38*, 3098–3100.
(53) Perdew, J. P. (1986) Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B 33*, 8822–8824.
(54) Eichkorn, K., Treutler, O., Oehm, H., Haeser, M., and Ahlrichs, R. (1995) Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett. 242*, 652–660, Erratum to document cited in CA123-93649.
(55) Klamt, A. (1995) Conductor-like screening model for real solvents: A new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem. 99*, 2224–2235.
(56) Klamt, A., Jonas, V., Buerger, T., and Lohrenz, J. C. W. (1998) Refinement and parametrization of COSMO-RS. *J. Phys. Chem. A 102*, 5074–5085.
(57) Klamt, A. (1998) COSMO and COSMO-RS. In *Encyclopedia of Computational Chemistry* (Schleyer, P. v. R. E-i.-C., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, and Schreiner, P. R. E., Eds.) pp 604–615, Wiley, Chichester, United Kingdom.
(58) Klamt, A., and Eckert, F. (2000) COSMO-RS: A novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilib. 172*, 43–72.
(59) Hilal, S. H., Karickhoff, S. W., and Carreira, L. A. (1995) A rigorous test for SPARC's chemical reactivity models: Estimation of more than 4300 ionization pKas. *Quant. Struct.-Act. Relat. 14*, 348–355.
(60) Hilal, S. H., Karickhoff, S. W., and Carreira, L. A. (2003) Prediction of chemical reactivity parameters and physical properties of organic compounds from molecular structure using SPARC, EPA/600/R-03/030.
(61) Busalla, T. (1996) Berechnung von Membranverteilungskoeffizienten. Diploma Thesis, University of Cologne, Cologne, Germany.
(62) Klamt, A. (2005) *COSMO-RS from Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam.
(63) Sachs, J. N., Crozier, P. S., and Woolf, T. B. (2004) Atomistic simulations of biologically realistic transmembrane potential gradients. *J. Chem. Phys. 121*, 10847–10851.
(64) Vaes, W. H. J., Ramos, E. U., Hamwijk, C., van Holsteijn, I., Blaauboer, B. J., Seinen, W., Verhaar, H. J. M., and Hermens, J. L. M. (1997) Solid phase microextraction as a tool to determine membrane/water partition coefficients and bioavailable concentrations in vitro systems. *Chem. Res. Toxicol. 10*, 1067–1072.
(65) Gobas, F. A. P. C., Lahittete, J. M., Garofalo, G., Shiu, W. Y., and Mackay, D. (1988) A novel method for measuring membrane-water

partition coefficients of hydrophobic organic chemicals: Comparison with 1-octanol-water partitioning. *J. Pharm. Sci. 77*, 265–272.
(66) U.S. Environmental Protection Agency (2000) EPI Suite, 1.67, http://www.epa.gov/oppt/exposure/pubs/episuitedl.htm.
(67) Meylan, W. M., and Howard, P. H. (1995) Atom/fragment contribution method for estimating octanol-water partition coefficients. *J. Pharm. Sci. 84*, 83–92.
(68) Escher, B. I., Bramaz, N., Richter, M., and Lienert, J. (2006) Comparative ecotoxicological hazard assessment of beta-blockers and their human metabolites using a mode-of-action-based test battery and a QSAR approach. *Environ. Sci. Technol. 40*, 7402–7408.
(69) Diamond, J. M., and Katz, Y. (1974) Interpretation of nonelectrolyte partition coefficients between dimyristoyl lecithin and water. *J. Membr. Biol. 17*, 121–154.
(70) Xiang, T. X., and Anderson, B. D. (1994) The relationship between permeant size and permeability in lipid bilayer membranes. *J. Membr. Biol. 140*, 111–122.
(71) Bemporad, D., Essex, J. W., and Luttmann, C. (2004) Permeation of small molecules through a lipid bilayer: A computer simulation study. *J. Phys. Chem. B 108*, 4875–4884.
(72) Mayer, P. T., and Anderson, B. D. (2002) Transport across 1,9-decadiene precisely mimics the chemical selectivity of the barrier domain in egg lecithin bilayers. *J. Pharm. Sci. 91*, 640–646.
(73) Butina, D. (1999) Unsupervised data base clustering based on Daylight's fingerprint and Tanimoto similarity: A fast and automated way to cluster small and large data sets. *J. Chem. Inf. Comput. Sci. 39*, 747–750.
(74) Worth, A. P., Bassan, A., Gallegos, A., Netzeva, T. I., Patlewicz, G., Pavan, M., Tsakovska, I., and Vracko, M. (2005) The Characterisation of (Quantitative) Structure−Activity Relationships: Preliminary Guidance, EUR 21866 EN.
(75) R Foundation for Statistical Computing (2005) R: A language and environment for statistical computing, 2.2.1, http://www.R-project.org.
(76) Finkelstein, A. (1970) Weak-acid uncouplers of oxidative phosphorylation. Mechanism of action on thin lipid membranes. *Biochim. Biophys. Acta 205*, 1–6.
(77) Barstad, A. W., Peyton, D. H., and Smejtek, P. (1993) AHA-heterodimer of a class-2 uncoupler: Pentachlorophenol. *Biochim. Biophys. Acta 1140*, 262–270.
(78) Netzeva, T. I., Worth, A., Aldenberg, T., Benigni, R., Cronin Mark, T. D., Gramatica, P., Jaworska, J. S., Kahn, S., Klopman, G., Marchant, C. A., Myatt, G., Nikolova-Jeliazkova, N., Patlewicz Grace, Y., Perkins, R., Roberts, D., Schultz, T., Stanton, D. W., van de Sandt, J. J. M., Tong, W., Veith, G., and Yang, C. (2005) Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. The report and recommendations of ECVAM Workshop 52. *ATLA 33*, 155–173.
(79) Mayer, P., and Reichenberg, F. (2006) Can highly hydrophobic organic substances cause aquatic baseline toxicity and can they contribute to mixture toxicity? *Environ. Toxicol. Chem. 25*, 2639–2644.
(80) Hanstein, W. G., and Hatefi, Y. (1974) Trinitrophenol. Membrane-impermeable uncoupler of oxidative phosphorylation. *Proc. Natl. Acad. Sci. U.S.A. 71*, 288–292.
(81) McLaughlin, S., Eisenberg, M., Cohen, F., and Dilger, J. (1978) The unique ability of picrate to uncouple submitochondrial particles but not mitochondria is consistent with the chemiosmotic hypothesis. *Front. Biol. Energ. [Pap. Int. Symp.] 2*, 1205–1213.
(82) Avdeef, A. (2003) *Absorption and Drug Development: Solubility, Permeability and Charge State*, John Wiley & Sons, Inc., Hoboken, NJ.

TX700391F