Chapter 28

Environmental Toxicity (Q)SARs for Polymers as an Emerging Class of Materials in Regulatory Frameworks, with a Focus on Challenges and Possibilities Regarding Cationic Polymers

Hans Sanderson, Kabiruddin Khan, Anna M. Brun Hansen, Kristin Connors, Monica W. Lam, Kunal Roy, and Scott Belanger

Abstract

Polymers are highly diverse and understudied materials from an environmental toxicity point of view. For the past decades, polymers have largely been out of scope regarding detailed safety assessment in most regulatory programs as they are assumed to not possess relevant toxicological properties due to their size. This regulatory exclusion is currently being reconsidered. This chapter discusses the available information about selected cationic polymers and outlines (Q)SAR ((Quantitative) Structure-Activity Relationship) approaches that could be used to develop new models to demonstrate potential aquatic toxicity of polymers. The amount of publicly available, high-quality environmental toxicity data on industrial polymers such as cationic polyquaterniums is extremely limited. Given the large size (dimension and molecular weight) of the materials, typical hydrophobicity-driven toxicity is not expected. Relevant descriptors for cationic polymers need to be identified. Molecular weight and charge density are well-known physical-chemical attributes that are suspected to be correlated with aquatic toxicity, but there might be other relevant descriptors as well.

We suggest models that predict polymer properties may be useful for estimating relevant properties regarding toxicity. Moreover, novel fragment-based 2D and 3D hologram (Q)SAR (H(Q)SAR) may prove relevant in determining these properties that can be used to derive hypotheses about toxic mechanisms and guide experimental test designs. In a regulatory context, (Q)SARs have to be transparent and scientifically robust which extends to fragment-based models that may be useful in categorizing polymers. The toxicity of category members can then be experimentally explored, and read-across strategies developed within the category.

The authors of this chapter are pursuing polymer (Q)SAR strategies in the coming years via generation of novel experimental and computational data on polyquaterniums. We will also evaluate the potential for fragment-based (Q)SARs for polymers in REACH.

Key words Polymers, Chemometric tools, Descriptors, Environmental toxicity, Cationic, Polyquaterniums

Kunal Roy (ed.), *Ecotoxicological QSARs*, Methods in Pharmacology and Toxicology,
https://doi.org/10.1007/978-1-0716-0150-1_28, © Springer Science+Business Media, LLC, part of Springer Nature 2020

1 Introduction

Polymers are large macromolecules consisting of repeating monomer units. Polymers are an exceptionally diverse group of compounds and are used in a large range of applications. Polymers may be described as linear, branched, or cross-linked. They may exist as homopolymers and have one repeated monomer, or they may be copolymers and contain two or more monomers combined in random or ordered approaches. While many polymers in commerce are synthetic, there are also natural polymers or biopolymers that are important building blocks of life, such as amino acids, proteins, and cellulose. Many polymers are soluble and dispersible in water. These are often used in consumer and personal care products, pharmaceuticals, water treatment, and wood preservation. Novel uses and applications in biomedical and nano-industries are expected to grow significantly in the coming years.

1.1 Current Regulatory View of Polymers

Historically, polymers have been subject to exemptions or reduced regulatory requirements in countries practicing chemical legislation. The assumption was that the high molecular weight and reduced reactivity of polymers in environmental compartments were viewed as lower risk to human health and the environment when compared to lower molecular weight substances. Most chemical legislations have adopted 500 Da as the highest molecular weight in scope, which is based on one component of Lipinski rules of bioavailability whereby substances that are >500 Da are considered less bioavailable. Since polymers are predominantly represented by higher molecular weight componentry, it has long been assumed that much of the polymer is not bioavailable and inert in the environment and the focus of chemical registrations and data needs has been more on lower molecular weight impurities and unreacted monomers as well as additives (non-intentionally added substances, NIAS, and intentionally added substances, IAS). The focus of most regulatory programs is generally on new polymers, not existing polymers already in commerce. However, through K-REACH, Korea is the first country to require registration of current polymers, with all existing chemistries greater than 1 metric ton requiring registration by 2030. In addition, Environment Canada has polymers included in their Chemical Management Program, and the agency recently published draft safety assessments for the polyamines (December 2016).

For new polymers requiring registration, the criteria used by many global regulatory agencies to identify "polymers of low concern" include molecular weight and levels of monomers, in addition to the presence of specific structural features or functional groups. The "polymers of low concern" concept is intended to guide prioritization of polymer review by regulators. While this

approach has been the practice of many global regulatory agencies, polymers have been exempted from chemical registrations by the European Chemicals Agency (ECHA).

Previous guidance from ECHA on polymer registration was done with the view that polymers would eventually require registration. Below is an excerpt from Article 138 Section 2 [1] (ECHA):

> "The (European) Commission may present legislative proposals as soon as a practicable and cost-efficient way of selecting polymers for registration on the basis of sound technical and valid scientific criteria can be established, and after publishing a report on the following:
>
> (a) The risks posed by polymers in comparison with other substances;
>
> (b) The need, if any, to register certain types of polymer, taking account of competitiveness and innovation on the one hand and the protection of human health and the environment on the other."

In recent years, the simplified view of the potential risks associated with polymers has received increased scrutiny, and with the polymer exemption under REACH being revisited, the expected outcome is starting approximately in the year 2023; polymers identified as "polymers requiring registration" (PRR) will come within the scope of REACH. This activity suggests the potential reapplication of current REACH methods, such as categorization and use of (Q)SAR, to characterize and estimate safety data and even to support grouping and read-across approaches for these materials. The use of (Q)SAR may be of special interest to regulators due to the limited publicly available data for polymers. It is likely many suppliers and downstream users of polymers have privately held data, but this information is often protected as confidential business information (CBI) due to the competitive environment of polymer innovation. Without unrestricted access to environmental safety data on polymers, the need for (Q)SAR development becomes more acute and relevant for polymer registration with the potential to use (Q)SAR to predict toxicity of polymers in lieu of testing.

As mentioned above, the "polymers of low concern" approach is based on the assumption that there is little toxicological concern expected for polymers due to their decreased bioavailability as a result of their significant molecular weight and their inability to cross biological membranes. Many (Q)SAR models used to estimate environmental fate and effects have not included large molecular weight chemistries in their training set; therefore, (Q)SARs are not intended to be useful predictors for large molecular weight and complex polymers. Most (Q)SARs used to estimate toxicity are based on log Kow as a surrogate for hydrophobicity/hydrophilicity or contributions of certain functional groups or structural features. Log Kow are generally estimated using fragment-based approaches, leading to a gross overestimation for high molecular weight

polymers. Furthermore, polymers are classic forms of UVCBs (Chemical Substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials) which are difficult to devise (Q)SARs for as they are not discrete chemical entities. There is, however, a somewhat historical (Q)SAR for polymers in ECOSAR in the EPI Suite [2], which is also built into the OECD (Q)SAR toolbox. These (Q)SARs were developed using a set of data from polymers exclusively.

In the mid- to late 1990s, the USEPA conducted a review of more than 10,000 polymeric substances notified to the USEPA [3] for market access, and based on this exercise, a guidance document was developed along with (Q)SAR domains for these materials. Although this guidance received a modest update in the mid-1990s [3], it remains the main reference point for the document (USEPA) [4].

The ECOSAR models are, in essence, based on chemicals with specific mechanisms of action causing excess toxicity (greater toxicity than that predicted by baseline toxicity) and compounds without a specific mechanism of action. There are a couple dozen known mechanisms of toxicity (e.g., organophosphates and others). However, the majority of compounds are nonspecific and have what is known as a narcotic mechanism of action (typically more than 75% of all chemicals) [5].

The narcotic mechanism of action, also known as baseline toxicity, is based on the assumption of disruption of the cell membrane integrity. This means that the critical cell membrane function

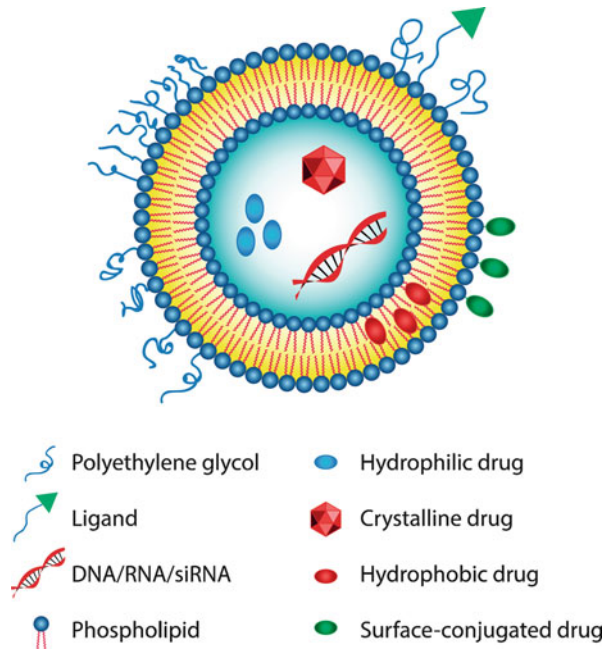

Fig. 1 Cell membrane (Colorbox)

necessary to sustain homeostasis in the chemical environment of the cell is impaired and the cell will die. The cell membrane (Fig. 1) is shared among all forms of life on Earth and is therefore a very good proxy for toxicity as the disruption of the cell membrane integrity will affect all life.

The mechanism of cell membrane disruption has not been entirely clarified, but it is usually characterized as a puncture or shift in fluidity of the lipid bilayer protecting the cell so that the functions needed to maintain homeostasis are impaired (e.g., efflux pumps or ligands are closed or lost). Some compounds can penetrate the membrane via pumps and receptor ligands in the membrane; these are typically the compounds with excess toxicity. Hence, an experimental proxy for the cell membrane was needed, and n-octanol/water partitioning coefficient was identified as a good model for the partitioning of chemicals between water and lipids such as the lipid bilayer. The log Kow expresses the ability of the compound to disrupt the cell membrane, and, hence, the most significant acute environmental toxicity descriptor was defined [6]. Toxicity is derived from Greek with the original meaning "poisons arrow" and is defined as a compound's ability to penetrate the cell membrane. The Paracelsus toxicity theorem that *dose makes the poison* we have used in toxicology for the past centuries is therefore enabled.

The aim of this chapter is to discuss the potential possibilities and challenges with the development and use of environmental toxicity (Q)SARs for polymers in a regulatory context for REACH. With current knowledge and available computational tools, we will explore how to build on the work by USEPA a quarter of a century ago and develop novel (Q)SARs for relevant polymers and used in anticipated REACH registrations. It is already clear that chemical assessments in REACH will not utilize (Q)SARs directly to satisfy registration requirements for specific endpoints (e.g., acute fish toxicity) but they may be exceptionally important in the establishment of chemical categories or groupings thereby lessening testing needs. The chapter will only take into consideration (Q)SAR and read-across as tools for risk assessment of polymers in a broad sense. Some details of the methodology involved in (Q)SAR and read-across will be discussed in Subheading 2 below after we have briefly reviewed the available data and provided a couple of examples.

2 Materials and Methods

Polymers as a group contain a wide variety of materials with differing structural attributes, functionalization, and physical/chemical properties. Polymers are composed of repeating monomer units, and copolymers are made up of more than one species of monomer.

2.1 Compounds: Polymers—A Brief Overview of Chemical Diversity and Available (Q)SARs

The copolymers are classified by how the units are arranged in the chain. The major groups include alternating, random, and block copolymers. Branched copolymers have a single main chain with one or more polymer side chains that are grafted or have branching that form other architectures. This complexity and diversity in polymeric species and structure present a significant challenge for their assessment and modeling. Polymers may contain structural alerts and/or specific functionalized properties (e.g., in pharmaceuticals and biocides) and may require specific toxicity analysis. Others may be completely toxicologically inert or have specific features and uses that warrant further assessment. It therefore makes sense to further define these materials.

According to USEPA definitions, compounds with a molecular weight greater than 1000 Daltons are too large to pass through the cellular membrane and are therefore unable to exert toxicity in a traditional manner. However, these compounds could cause mechanical effects (e.g., gill clogging) at high concentrations (e.g., typically observed at >1000 mg/L). Mechanical effects including binding to external and internal (e.g., gut) biological surfaces which are not "toxicity" in the traditional sense but for biological organisms may still be ecologically relevant. Polymer safety assessments may include separate considerations for the polymer, oligomers, and monomers depending on the polymer composition. USEPA [4] divides polymers into three categories based on the average molecular weight (MWn) and the amount of low molecular weight components (LMW):

Category 1: Polymers with low average molecular weight (MWn <1000 Daltons). These can potentially be assessed as discrete structures in EPI Suite, within the normal limitation of the software, as long as the composition and structure of the polymer is known [4, 6].

Category 2: Polymers with high average molecular weight (MWn >1000 Daltons) and large LMW material composition (≥25% with MW <1000 Daltons; ≥10% with MW <500 Daltons). The environmental toxicity of these polymers can be assessed; however, oligomers may need separate assessment to account for any increased toxicity due to their lower molecular weight [4].

Category 3: Polymers with high average molecular weight (MWn >1000 Daltons) and minimal LMW material (<25% with MW <1000 Daltons; <10% with MW <500 Daltons). These are generally assessed solely as the polymer (USEPA) [4].

The aquatic toxicity of polymers is influenced by solubility. Insoluble polymers are not expected to be toxic due to lack of bioavailability. Typical acute aquatic toxicity values for these polymers are >100 mg/L or > 10 mg/L for acute and chronic tests, respectively. However, physical or mechanical effects may occur if the insoluble polymer exists as a fine particle. Indeed, this is the case for microplastic particles [7].

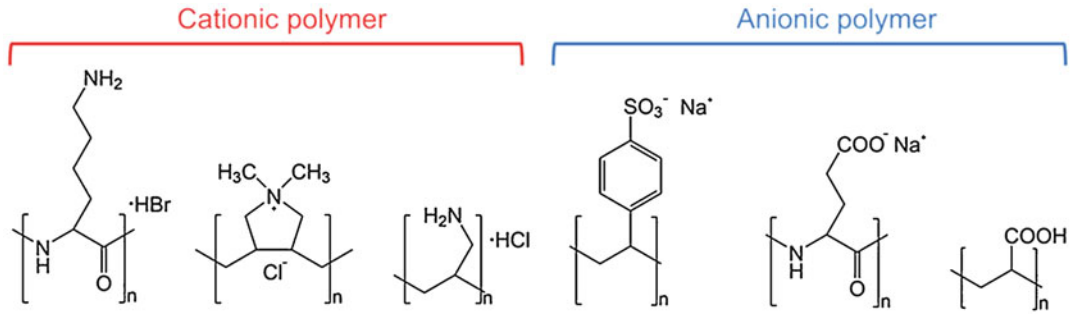

Fig. 2 Examples of cationic and anionic polymers

Polymer charge (neutral, anionic, cationic, amphoteric) can also modulate toxicity (Fig. 2). Nonionic polymers have very low water solubility and are generally believed to be of low hazard concern, unless they contain a significant amount of oligomer or if the polymer is used as a surfactant or dispersant. Anionic polymers are classified as poly(aromatic acids) or poly(aliphatic acids). Poly(aromatic sulfonate and carboxylate) polymers have moderate acute aquatic toxicities with fish, daphnids, and algae (LC50 1–100 mg/L). Poly(aliphatic acids) polymers have low toxicity to fish and daphnids (LC50 >100 mg/L), whereas algae seem to be more sensitive presumably due to chelating effects of nutrients. Due to chelation potential of many of these polymers, the mitigation potential of hard water further complicates study interpretation and design. The toxicity of both cationic and amphoteric polymers has been shown to increase with increasing cationic charge density [4]. As cationic polymers are believed to pose the greatest environmental hazard, the need for accurate aquatic acute toxicity (Q)SAR predictions is the greatest for these compounds. Other properties that may impact the toxicity of the pure polymer include physical form, particle size distribution, swellability, dispersibility, and of course in addition to these the presence and potentially weight fraction of reactive functional groups.

2.2 Cationic Polymers

Although cationic polymers are not limited to quaternary ammonium, phosphonium, and sulfonium functional groups, cationic polymers with quaternary ammonium groups are used in personal care and household cleaning products as conditioners or softeners and as flocculants in drinking water treatment plants. Therefore, there is a potential for the release into the aquatic environment. Sound environmental risk assessment, with a focus on the aquatic compartment, is of particular interest for cationic polymers, especially those with the quaternary ammonium functionality. At the time of the Boethling and Nabholz publication on polymer risk assessment, almost all of the cationic polymers reviewed contained a N-functionality [3].

Cationic polymers have a net positive charge at environmental pH and therefore have the potential to be highly sorptive and surface active. It has been suggested that cationic polymers will sorb to biological surfaces which are net negative, and previous studies have shown that cationic polymers have an impact on respiratory processes and may disrupt oxygen transfer (e.g., gill membranes of fish; Biesinger and Stokes) [8]. Total organic content (TOC) and dissolved organic carbon (DOC) have been shown to have a mitigating impact on the aquatic toxicity of cationic polymers, presumably due to sorption, but most of these investigations have been with highly charged polymers, and less mitigation may be presumed for lower cationic charged compounds. Traditional toxicity studies are conducted in clean media (e.g., standard OECD test media). The TOC and DOC levels in this media are not representative of environmental concentrations, and, therefore, the hazard values derived from these standardized studies may overestimate the environmental hazard of cationic polymers. Mitigation factors, specific to cationic polymers charge density and LMW composition, have been developed to adjust aquatic toxicity values to reflect environmental TOC levels. Based on confidential data from 53 cationic polymers, the USEPA described mitigation factors ranging from 7 to 290 [2–4]. However, these studies were conducted in the absence of analytical verification. Experimental TOC and DOC values may be important parameters to include in (Q)SAR modeling building exercises. These observations have resulted in some test conduct considerations as reflected in the OECD difficult test substance monograph (OECD) [9] and USEPA [10]. Because cationic polymers interact with anionically charged substances in general, we have observed toxicity mitigation as a function of water hardness in our laboratories (P&G, unpublished data). These are important, since toxicity often is linked to the positively charged polymers [11, 12].

In addition to DOC/TOC levels, other parameters to be considered in developing (Q)SARs for cationic polymers are physical-chemical properties that often serve as identity descriptors for the polymer. The cationic charge is typically found on a nitrogen group. For this reason, the % amine-nitrogen has been previously used as a descriptor in aquatic toxicity (Q)SARs. To further elaborate, from the historical work by Boethling and Nabholz [3], the cationic charge density based on %amine-nitrogen is because almost all the polymers submitted to the US TSCA office had their cationic charge based on nitrogen. The polymer backbone may also influence the toxicity. Cationic polymer backbone types can be carbon-based, silicon-based (e.g., Si-O), or natural (e.g., starch). The importance of backbone type and environmental hazard is not entirely clear. For fish, the toxicity silicon and carbon-based backbones are described using unique (Q)SAR equations. Natural polymer backbones are assumed to have equal or slightly less acute

toxicity than carbon backbone cationic polymers. However, daphnids have unique (Q)SAR equations for natural and carbon-based backbones, with silicon backbones having equal or slightly less acute toxicity than carbon-based backbones.

Fish and daphnid acute/chronic ratios range from 14 to 18, which suggest a narcotic mode of action [3]. The toxicity ranged from 0.006 mg/L towards algae for a carbon-based backbone polymer with a 7.8% amine-nitrogen charge density quaternary amine and 38% MW <500 to more than 1000 mg/L for a natural-based backbone with 0.07% amine-nitrogen charge density quaternary amine and 0% MW <500) [3].

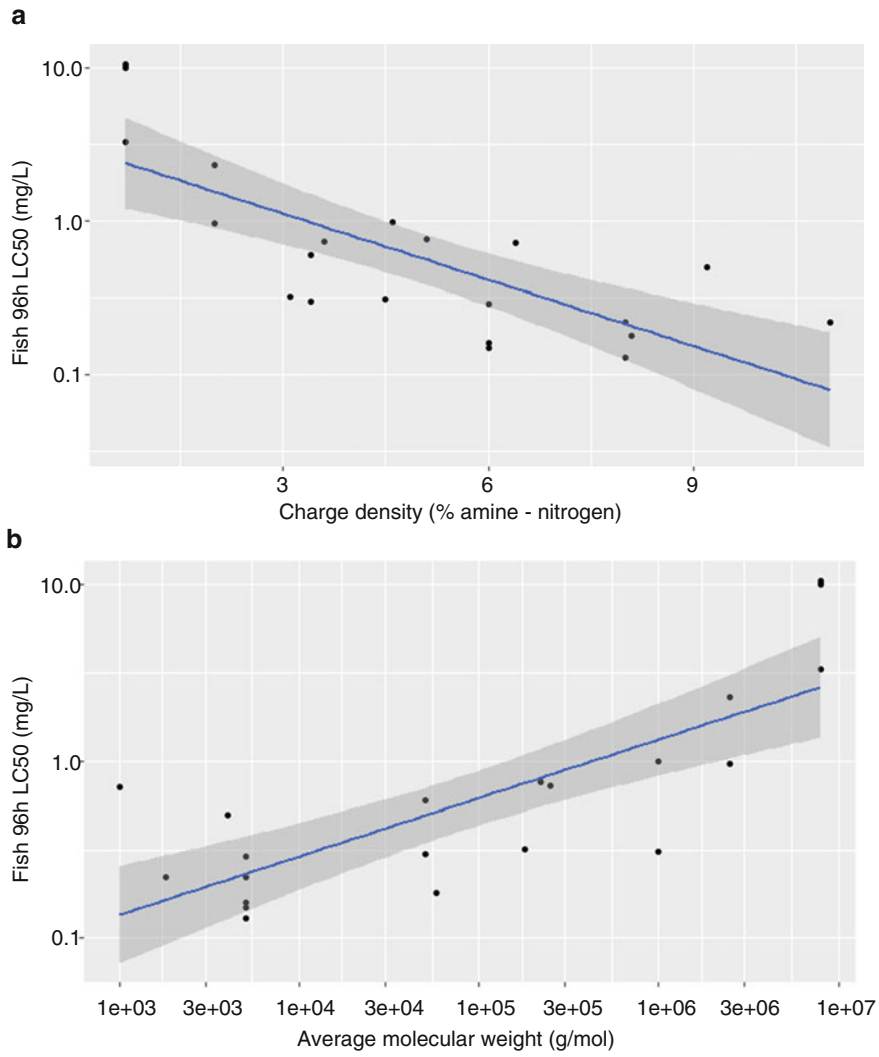

Fig. 3 Acute fish toxicity of quaternary amine cationic polymers (carbon-based backbone) as a function of (**a**) charge density (% amine-nitrogen) and (**b**) average molecular weight. Data obtained from Boethling and Nabholz (1996) [3]

It has been suggested by Boethling and Nabholz [3] that the aquatic toxicity of cationic polymers is influenced by charge density, molecular weight, and position of cation relative to the backbone. Figure 3 depicts the relationship between acute fish toxicity and (a) charge density (percent amine-nitrogen) or (b) average molecular weight in carbon-based backbone quaternary cationic polymers [3]. These plots hint towards increasing charge density leading to an increase in toxicity, whereas an increasing molecular weight corresponds with a decrease in toxicity—however much work is needed to develop reliable (Q)SARs.

In recognition of the impact of charge density of the environmental toxicity of cationic polymers, several regulatory agencies (e.g., Canada) have established a functional group equivalent weight (FGEW) cutoff of 5000 for the criteria of polymer of low concern (PLC). This concept has also been supported by the OECD review in 2009 of PLC criteria around the world. The FGEW cutoff concept can be a valuable tool in the prioritization of polymers to be selected for detailed regulatory reviews (e.g., REACH registration in EU).

2.3 Polyquaternium Cationic Polymers: A Complex Cationic Polymer Category

There is very limited data available on a specific class of polyquaternium that supports the observation that measured aquatic toxicity is influenced by charge density.

Polyquaternium cationic polymers represent a class of particular interest of cationic polymers due to their widespread use and releases to the aquatic environment. Polyquaterniums represent a very wide diversity of structures, and as of early 2019, there were approximately 40 registered active varieties with the Chemical Abstracts Service. Polyquaterniums are available as homopolymers or copolymers, and most are water soluble. Homopolymers vary in MW typically from <100,000 to 500,000 Daltons. All polyquaternium polymers contain a monomer with a quaternary ammonium functional group, such as diallyldimethylammonium chloride or trimethylammonium chloride. There is a diversity in monomer chemistries used as the copolymer for the quaternary ammonium monomer. A few examples of nonionic or anionic copolymers include vinylpyrrolidone, acrylic acid, polyvinyl alcohol, and acrylamide. Within each class of polyquaternium, the molecular weight will vary depending on the number of repeat units. While charge density remains constant for homopolymer polyquaterniums, the range in charge density or degree of substitution is dependent on the ratio of the monomers for copolymer polyquaterniums. The selection of monomers and fine-tuning of monomer ratios are necessary to give a range of physical-chemical properties and product benefits for diverse applications.

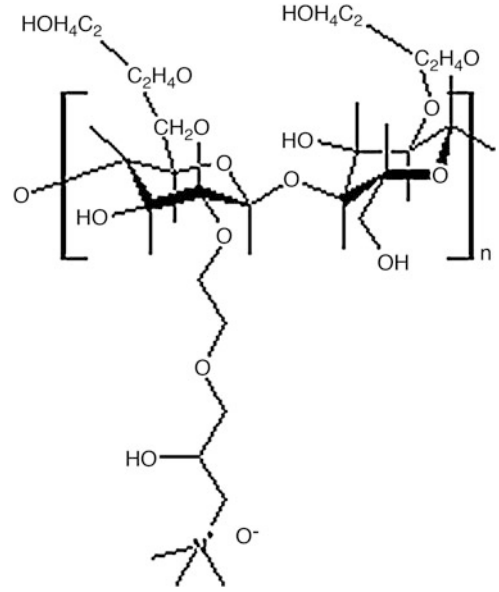

Fig. 4 Representative structure of polyquaternium-10

Table 1
Polyquaternium-10 aquatic effects data

Variant	Charge density [13] (meq/g)	%N[2]	Avg MW[a] [14] (kDa)	Viscosity (as 2% aq sol'n) (mPa/s)[2]	96 h acute fish EC_{50} (*Gambusia holbrooki*) [15] (mg/L)	72 h EC_x algae (*Chlorella sp12*) [17] (mg/L)
UCARE JR125	High (0.9)	1.5–2.2	Low (250)	75–125	1.2	$EC_{50} = 0.04$
UCARE JR30M	High (1.0)	1.5–2.2	High (600)	30,000	1.5	$EC_{10} = 0.002$ $EC_{50} = 0.05$
UCARE JR400	High (1.2)	1.5–2.2	Low (400)	300–500	2.1	$EC_{10} = 0.013$ $EC_{50} = 0.05$
UCARE LK	Low (0.3)	0.4–0.6	Low (~400)[b]	300–500	100	Not available
UCARE LR30M	Low (0.4)	0.8–1.1	High (600)	30,000	66	Not available
UCARE LR400	Low (0.6)	0.8–1.1	Low (~400)[b]	300–500	64	Not available

[a]Supplier information
[b]Estimated based on viscosity information [13–17]

Polyquaternium-10 (PQ10) is a cationic cellulose polymer with quaternary ammonium functionality, varying in charge density and MW (Fig. 4). The diversity in charge density is driven by the ratio of the monomer groups. A representative structure is illustrated in Fig. 4.

Table 1 below provides the measured and published aquatic toxicity of PQ10; the newest data is from 1991.

Based on the limited data available on aquatic effects, it could be proposed that charge density within a polymer class influences aquatic effects on fish and algae, while MW does not appear to have an impact. More information, from a well-structured toxicity investigation program, would be useful to determine the viability of the hypothesis. This observed trend supports the rationale to develop (Q)SAR to estimate aquatic effects when applicable. Since there is very limited publicly available data, it is not well understood whether a (Q)SAR developed for one polymer subclass could be leveraged by another subclass with some common structural features.

In a more recent example of research to understand whether physical properties of polymers can be used to estimate aquatic toxicity, Pereira et al. evaluated molecular weight, charge density, and integrative intrinsic viscosity of several cationic polyacrylamides to determine whether these structural features and variables could be used to predict the environmental effects [18]. The studied polyacrylamides were copolymers of acrylamide and acryloyloxyethyltrimethyl ammonium chloride with a cationic monomer content between 40 and 50% (w/w). The test species included in this study were bacteria, microalgae, macrophytes, and daphnids. While correlations were found between physical properties of the cationic polyacrylamides, the authors concluded that no clear ecotoxicity patterns correlating to physical properties were observed. While the observations may be valid for this particular group of polymers, the historical data from Boethling and Nabholz and Cumming et al. suggest there is a general relationship between certain structural features, such as charge density, and observed aquatic toxicity for cationic polymers, and in fact, (Q)SARs have been used for decades to estimate toxicity of cationic polymers by the USEPA [3, 15].

It is clear from the above that there is a strong need to explore (Q)SAR methodologies to describe the toxicity of polymers and cationic polymers in particular. The regulatory development of (Q)SARs for polymers has been advanced very little for the past decades, and publication of environmental toxicity data has also been sparse in that period. We will therefore in Subheading 3 section briefly describe possible options that may be applied in future elucidation of environmental toxicity (Q)SAR methods for polymers.

3 (Q)SAR Methods

Developing (Q)SARs based on curated PQ data is challenging as the data availability, transparency, and quality for the training set are limited and insufficient polymer descriptor information is available. The same is the case in a greater degree for cationic polymers in general [19]. And (Q)SARs are of course even more challenging for polymers in general based as they are much more diverse and data poor. Below is an outline of methods and approaches to consider.

3.1 Chemometric Tools in Ecotoxicological Evaluation of Polymers

In the recent decade, we have seen a notable rise in the use of alternative strategies in testing methods, including computational tools, for safety assessment of various organic/inorganic chemicals [20–22]. The in silico tools have demonstrated their successful application in detecting hazard potential of various chemicals belonging to several subclasses such as pharmaceuticals [23], agrochemicals, nanoparticles, and personal care products [23–27]. For emerging pollutants such as micro- and nano-sized particles [28] and polymers, such models are available to much a lower extent. There is a clear-cut deficit in the number of reports on application of in silico tools in toxicity (especially ecotoxicity) assessment of polymeric materials. The data scarcity on polymer ecotoxicity whether in silico or in vitro is evident mainly from availability of very few published studies in the literature. One possible reason is the high degree of proprietary nature for polymers and concerns with protecting confidential business information by disclosing identity descriptors for polymers in the public domain. While there are methods available that can estimate the effects of individual parent monomers [29, 30], the polymeric versions of the compounds are often left unevaluated (due to highly extensive computational requirement). Quantitative structure-activity/property relationship ((Q)SAR/QSPR) and quantitative read-across analysis (QRA) are widely accepted computational techniques, which are believed to be the most successful [2] two approaches that can be successfully implemented in identification of potent environmental pollutants among polymeric compounds (specifically, cationic polymers in view of their insufficient experimental data) using a very small amount of experimental results. It is also worth mentioning here that regardless of how statistically robust or significant a (Q)SAR/read-across model may be, it would be unavoidably associated with certain limitations [6]. These limitations are model specific, such as that a single (Q)SAR model may have its limited applicability owing to its restricted chemical domain which can be tackled by using intelligent consensus (Q)SAR approaches as proposed by Roy et al. [31]. Another major challenge in predictive toxicology is to effectively evaluate the reliability of obtained predictions of unknown/untested or not even synthesized chemicals. This limitation was also addressed recently with the introduction of prediction reliability indicator tool as proposed by Roy et al. [32]. Several commercially available tools for prediction of different endpoints for chemicals in general include TOPKAT software [33], CAESAR [34], ECOSAR [2], Toxicity Estimation Software Tool, etc.; however these tools generally do not include polymers in their training set which could be considered an important limitation [35]; hence we explore in the following sections alternative approaches.

It is well known that the most robust environmental toxicity tests are accompanied by confirmatory analytical verification of exposure. Analytical exposure determinations in aquatic toxicity tests are formally required, whenever feasible, under all typical OECD test guidelines for acute and chronic aquatic toxicity. However, limitations are also known for confirmation of exposures when polymers are tested. Indirect determinations can be useful in limited circumstances. These may include total organic carbon (sensitive down to perhaps 2 mg/L) or other alternatives such as measurement of an inorganic component such as silicon as was done in the 1990s during the programs addressing environmental safety of polydimethylsiloxane (PDMS) polymers [36]. (Q)SAR developments may be somewhat hampered by the lack of specific analytical verification of exposures until "high-end" analytical methods can be made routine and broadly available.

3.2 (Q)SAR Methodologies: Broad Classifications

The toxicity of whole polymeric structures or the structures in a monomeric form can be analyzed using (Q)SAR/QSPR methods, which can be classified as follows:

Regression-Based (Q)SAR This technique can be implemented to explore the quantitative correlation between toxicity of polymeric materials with the corresponding structural features. Multiple linear regression, partial least squares, and artificial neural networks are some of the examples of regression-based approaches. The use of regression approach for polymers is demonstrated in [37].

Classification-Based (Q)SAR For graded responses or where there is a lack of absolute quantitative toxicity data of polymers, classification-based techniques can be used to group the data into Boolean classes such as toxic, nontoxic, or moderately toxic classes. A classification-based technique like linear discriminant analysis (LDA) is also helpful in big data analysis [38].

3.3 Protocols for (Q)SAR Analysis in Polymers

(Q)SAR follows well-established protocols for developing statistically acceptable models for prediction of activity/property/toxicity chemical compounds [38, 39]. The Organization for Economic Cooperation and Development (OECD) has recommended five basic principles for (Q)SAR model development: (1) a defined endpoint, (2) an unambiguous algorithm, (3) a defined domain of applicability, (4) strict validation protocols, and (5) mechanistic interpretation, if possible [26]. The details of any (Q)SAR workflow are discussed below.

Collection of reported/generated biological data: For a (Q)SAR study involving polymeric compounds, the data collection should follow the prescribed guidelines of OECD [38, 39] which include uniform experimental conditions, uniform time of exposure for the desired effect, experiment with a standard species, analytical verification of the exposure concentrations, etc. The

data curation should be done effectively to check for duplicates/salts/ions, etc. Another important point in collecting homogenous ecotoxicity data for polymers includes ideality in experimental water conditions such as hardness, alkalinity organic carbon content (TOC and DOC), etc. that may affect the observed toxicity. In the case of algal testing and (Q)SARs, definition of specific anionic and cation components of media may also be important.

Descriptor calculation: For the descriptor calculation, in most of the cases, initially the monomeric or repeating unit is identified. The flanks of monomers are capped with hydrogen atom in order to satisfy the valence electron. Then the structure are subjected to descriptor calculating software such as Dragon [40], SiRMS [41], alvadesc [42], PaDEL-Descriptor [43], etc. to calculate molecular descriptors.

Division of the dataset: In order to obtain useful models, the collected data should be partitioned into training and test sets following unbiased methods. Some of the widely followed dataset division techniques include Kennard-Stone [44], Euclidean distance approach [45], k-medoids approach [46], and random sampling. These tools for dataset division are available, for example, at http://teqip.jdvu.ac.in/(Q)SAR_Tools/.

Feature selection: In feature selection, molecular descriptors important for the response values are identified. Some of the feature selection techniques include stepwise selection, genetic algorithm, double cross validation (DCV), and factor analysis [47]. The problems of small datasets (as in the case of polymer toxicity data, which is scarce) can be addressed to some extent using DCV. In DCV, the training set is split into calibration and validation sets, and these are utilized for model building and model selection, while the test set is exclusively used for model assessment. This process obviates the possibility of bias in descriptor selection. For ideal (Q)SAR models, the intercorrelation among the descriptors should be very less.

Modeling algorithms and chemometric tools used in (Q)SAR: The most commonly employed linear modeling algorithms include multiple linear regression (MLR) [48], univariate linear regression (ULR), ordinary least squares (OLS), partial least squares (PLS), principal component analysis (PCA) [27], principal component regression (PCR), etc.

Model validation metrics and mechanistic interpretation: Finally the developed model should be validated following internationally recognized guidelines. Some widely used validation metrics for regression models include leave-one-out (LOO) cross-validation R^2 (Q^2) and for training set evaluation and ${Q_{F1}}^2$, ${Q_{F2}}^2$, ${Q_{F3}}^2$, and concordance correlation coefficient (CCC) for test set evaluation. Some other stringent criteria for model validation include (1) mean absolute error (MAE) criteria proposed by Roy et al. [49] and (2) Golbraikh and Tropsha criteria for model

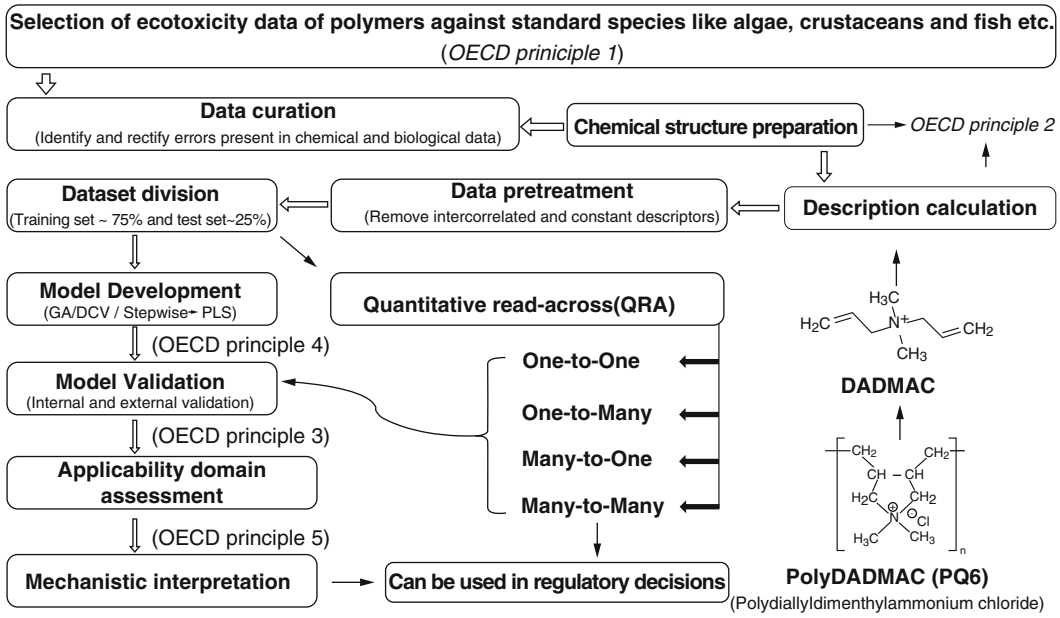

Fig. 5 Process involved in ecotoxicity study of polymers following in silico (Q)SAR and QRA approach

validation [46]. A mechanistic interpretation of a developed model is desired wherever possible. Figure 5 depicts the general outline of polymer toxicity modeling.

4 Applications of (Q)SAR to Polymers: A Literature Review—Applications of (Q)SAR in Ecotoxicity of Polymers

With the proprietary nature for many polymers, manufacturers and downstream formulators have generated aquatic effects data for stewardship reasons, but much of this data is privately held to protect confidential business information (CBI). However, there are some classes of polymers that have been studied with publicly available publications demonstrating potential toxicity of polymers with some aquatic species [19]. Several examples are presented below as case studies of (Q)SAR development for diverse polymer classes.

Acute algal toxicity: The very first and comprehensive (Q)SAR study on toxicity of polymers was conducted by Nolte et al. [19]. The data ($N = 43$) for growth rate inhibition (EC50) of algae were collected from the literature using Google Scholar and Web of Science. However, since the data was limited, the authors combined the data for two different times of exposure, i.e., 96-h and 72-h reflective of primarily USEPA and OECD algal test procedures, respectively. Three different models based on their charge separation (cationic ($N = 9$), anionic ($N = 16$), and

nonionic ($N = 17$) compounds) were developed using one theoretical descriptor following regression-based decision tree technique. More complex branched polymers, polymeric surfactants, and non-nitrogen cationic polymers were omitted from the study. The models predict that cellular adsorption, disruption of the cell wall, and photosynthesis could be the possible mechanisms of action for algal toxicity of cationic and nonionic polymers. The findings of the (Q)SAR results combined with molecular dynamics simulations proposed that nutrient depletion is likely the dominant mode of toxicity. (Q)SAR relationships for green algae growth inhibition, however with the low number of data for the generated (Q)SAR, were not statistically robust and do not comply with the quality criteria cited by Cherkasov et al. [6, 19].

4.1 Application of (Q)SAR in Toxicity Prediction of Polymers (Peptides)

Antimicrobial peptide toxicity: Langham and colleagues [51] developed (Q)SAR models to quantify and predict antimicrobial peptide toxicity against human host cells (epithelial and red blood cells) based on physicochemical properties like interaction energies and radius of gyration which were in turn calculated from molecular dynamics simulations of the peptides in aqueous solvent. For model the development 60 peptides with experimentally determined toxicities were used. Langham and colleagues [51] proposed based on the findings of molecular modeling study that physicochemical properties of peptides and interactions in a solvent are responsible for their toxicity against human cells in their native state. The developed models were then employed in predicting several other protegrin-like peptides. The (Q)SAR model could correctly rank four out of five protegrin analogues newly synthesized and tested for toxicity in laboratory.

Although quantitative structure-toxicity relationship modeling reports involving polymers are scarce, there are several reports on (Q)SAR/QSPR modeling of their biological activity and property endpoints. We report here some of them to demonstrate that similar tools may be applied to develop models to predict toxicity of polymers.

4.2 Application of (Q)SAR to Biomedical Applications of Polymers

Cellular response and protein absorption: Khan and Roy [52] developed predictive (Q)SAR models for a cellular response (fetal rate lung fibroblast proliferation) and protein adsorption (fibrinogen adsorption (FA)) on the surface of tyrosine-derived polymers designed for the purpose of tissue engineering. These polymers were synthesized using a combinatorial approach which in turn is a decade long process used in tissue engineering applications; the process is briefed in the source paper [52]. The model consists of 66 data for cellular response and 40 data for protein adsorption on polymers. The models were developed using only selected 2D descriptors having definite physicochemical meaning. To enhance the biological domain of the model, multiple (Q)SAR models were

developed and then subjected to consensus modeling as proposed by Roy et al. [31]. The final consensus models were validated using strict OECD guidelines and accepted internal and external metrics.

Cellular response: Semiempirical QSPR models were developed to predict the cellular response to the surfaces of polymers designed for tissue engineering applications by Kholodovych and colleagues [53]. The findings of the models were then compared with experimental results which showed a high degree of accuracy proving its significance for biomedical applications. Partial least squares (PLS) regression technique was used for model development using 62 polyarylates and structure-based molecular descriptors.

Bioresponse modeling: Artificial neural networks (ANN) were applied to model bioresponse to the surfaces of polymers collected from combinatorial library [54]. For analysis, 22 structurally distinct polymers were modeled against human fibrinogen adsorption. Additionally, the developed models were used to model rat lung fibroblast and normal human fetal foreskin fibroblast proliferation in the presence of 24 and 44 different polymers. The root mean square was used for the error comparison with experimental finding, and it was lower than experimental results thus proving applicability of the developed models.

Protein adsorption: Smith et al. [55] proposed a surrogate model for the prediction of protein adsorption onto the surfaces of polymers designed for tissue engineering applications. The proposed surrogate model combines machine learning, molecular modeling, and an artificial neural network. The experimental errors were estimated using Monte Carlo technique. The dataset consists of 45 polymers with measurements of human fibrinogen adsorption. A total of 106 molecular descriptors were computed using the Molecular Operating Environment (MOE) software. The surrogate model was developed in two stages: firstly the three descriptors with highest correlation to the adsorption were identified, and then these three descriptors were used as input for the second stage, i.e., for artificial neural network (ANN) to predict fibrinogen adsorption. Here, a Monte Carlo approach enabled a direct assessment of the effect of the experimental uncertainty on the results. Only the training set (nearly 50%) was employed for ANN using random sampling followed by checking of experimental error using Monte Carlo analysis. The accuracy of ANN was then compared with experimental data for the remaining polymers (the validation set). The Pearson correlation coefficient was used as validation metric. In conclusion, the surrogate model was proposed to get accurate and unambiguous predictions of polymers to check for their range of fibrinogen absorption, an essential requirement for assessing polymers for regenerative tissue applications.

4.3 Applications of (Q)SAR in Property Estimation of Polymers

In other areas of (Q)SAR development, there are a number of publications that demonstrate that quantitative structure-property relationship (QSPR) models can be developed to predict certain physical properties of polymers. Though a number of studies in the available literature exist on modeling of various properties of polymers, we have reported here a few of the recent reports.

Refractive Index Khan et al. [56] proposed robust QSPR models to predict refractive indices (RIs) of a set of 221 diverse organic polymers employing simple 2D descriptors generated by using monomeric unit. The final model consists of six theoretical descriptors developed using partial least squares (PLS) regression technique. For feature selection, double cross-validation tool was used. Use of consensus modeling for predictions from multiple modeling was also demonstrated. Finally, four small virtual libraries were selected to predict their RIs values using obtained consensus model.

Glass Transition Temperature The glass transition property of 206 diverse polymers was studied by Khan and Roy [37] using the QSPR approach since it has a direct impact on polymer stability. Five individual QSPR models were obtained using six 2D molecular descriptors following partial least squares regression and DCV as the feature selection tool. The models were extensively validated, and Y-randomization (Y-scrambling) test was performed in order to prove nonrandom and robust nature of the developed models. At last, comparison with existing QSPR models was made to demonstrate the effectiveness of the novel models.

5 Discussion of Future Avenues: Application of Fragment-Based (Q)SAR and Read-Across in Ecotoxicity Predictions of Polymers

The area of (Q)SAR modeling for the evaluation of toxicity of polymers has remained largely unexplored, which could be used to motivate and inspire (Q)SAR modelers to contribute to this dynamic and vastly underdeveloped field. A major notable point here is many of the previous modeling studies [57, 58] on polymers involve computation of quantum-chemical descriptors which can be a time-consuming process. This problem can be solved effectively by using only 2D descriptors having simple more definite physicochemical meaning in order to avoid conformational analysis, computational complexity of energy minimization, and alignment problems.

Apart from the classical methods of (Q)SAR model development, one can also apply more novel and more appropriate methods as discussed below.

Fragment-based (Q)SAR: These use molecular substructures expressed in fingerprints as descriptors in the developed models.

Fragment (Q)SARs can be implemented in the ecotoxicological modeling of polymers when studying a part of a molecule or specific group in relation with the toxicity. A widely used group-based (Q)SAR is H(Q)SAR (Hologram-(Q)SAR) [38, 39].

H(Q)SAR: This is a modern 2D FB-(Q)SAR (fragment-based) technique which utilizes molecular substructures expressed in binary pattern also termed as fingerprints in model development as variables. The method does not involve calculation of any physicochemical chemical descriptor or 3D structure generation. The process follows three steps:

1. Fragment generation for each of the training set molecules
2. Representation of the fragments in the holograms
3. Finding correlation of the molecular holograms with the corresponding activity data using training set compounds employing the PLS technique

A number of parameters affect H(Q)SAR model generation such as hologram length, fragment size, and distinction [38]. H(Q)SAR encodes all possible fragments within the molecules along with sub-fragments; thus it is helpful in understanding the fragments responsible for the toxicity of polymers in reference species. The other possible applications of H(Q)SAR in ecotoxicity of polymers include exploring individual atomic contributions to the toxicity with a visual display of active centers in the compounds.

Read-across: The read-across approach is a practice based on the assumption that structurally similar compounds exhibit similar physicochemical, environmental fate, toxicological, and ecotoxicological properties. The process starts with the grouping of similar objects (here, structures), and then the response value of one or more chemicals can be used to predict the behavior of target chemicals. Four different strategies for read-across have been proposed so far, i.e., one-to-one, one-to-many, many-to-one, and many-to-many. As per the OECD guidelines [58], the QRA prediction can be performed in following one of the four ways:

1. Using similar chemicals for the endpoint to perform read-across
2. Using a mathematical scale to check the trend in experimental results using two or more similar chemicals (e.g., trend analysis)
3. Taking an average of endpoint values of two or more source chemicals
4. If sufficient data is available, using the most conservative value from the source chemicals in the whole category

A read-across strategy can be used to estimate the toxicity for a series of cationic or anionic polymers with acceptable levels of

uncertainty. Considering that the toxicity data are available for a limited number of polymers, read-across will be very helpful for bridging data gaps. However, efforts are needed to define how similar polymers should be grouped and what key physical-chemical properties should be used in the grouping scheme. Data anchors at the extremes of the biological attribute being used to develop the read-across are important to define. Previous groupings by ECHA or EPA may be too broad, and further work is needed to refine based on the diversity of polymers within classes or subclasses. In addition, it is possible that the grouping and read-across approach may need to be customized depending on polymer class or even route of exposure. The potential impact of polymers to human health and the environment may be estimated through developing (Q)SAR models and by enabling read-across to structural analogues and avoiding or minimizing the need to conduct safety studies. This would bring benefits to time, resources, and avoiding animal testing. (Q)SARs could also be leveraged in polymer innovation and providing guidance on the design space.

6 Conclusions

It is clear from the above that regulatory programs are increasingly starting to include polymers for environmental risk assessment, chiefly in REACH, and that there has been a paucity for a couple of decades in the development of aquatic toxicity (Q)SARs by the USEPA [4] for polymers. There is hence a need to develop models for this purpose. It is also clear that polymers are very diverse and this diversity needs to be reflected in the model development and domains [6]. It is also clear that key and necessary data that are needed to do assessments or generate regression-based (Q)SARs are currently largely missing [19] and the sparse available experimental data lacks insight on experimental exposure. Moreover, regression-based (Q)SARs still require identification of the most determinant toxicity descriptors of the polymer. It is highly questionable if this is hydrophobicity since the mechanism of action is either unknown or not narcotic since the molecules are too large to exert the narcotic mechanism we normally associate with narcosis. Cationic polymers are highlighted as an example in this chapter of a class of polymers of high and down-the-drain use, more specifically polyquaterniums. The toxicity of these materials is dependent upon charge density, molecular weight, %amine-nitrogen, solubility, and type of backbone. There may be other additional and currently uninvestigated descriptors that govern the toxicity of these and other cationic polymers. We have suggested a series of non-regression-based (Q)SAR approaches that may be applied to elucidate the potential descriptors. Figure 5 outlines a process for developing (Q)SARs which when combined with the learnings from

Cherkasov et al. [6] are important methods moving forward. Using polymer properties may be useful for estimating fate, effects, and even form in the environment. For example, the glass transition temperature (Tg) [37] may be used to estimate form. If a polymer is below Tg, then it has to be a solid. If it is above Tg, then it could be a solid or liquid depending on the melting temperature of the polymer, which would determine the bioavailability and toxicological availability of the material. 3D comparative molecular field analysis and other ANN or 2D H(Q)SAR may prove highly relevant—but in a regulatory setting, the models have to be transparent in which case the fragment-based models may initially be used to identify critical toxicity and availability descriptors which can then be used to cluster the polymers. The toxicity of these clusters can then be experimentally explored and recorded and subsequently develop read-across within these. The authors of this chapter are pursuing this in the coming years via generation of novel experimental and computational data on polyquaterniums, and we will also evaluate the potential for fragment-based (Q)SARs for polymers in REACH.

Acknowledgments

This work was supported by Cefic in the LRI project ECO46: iTAP (http://cefic-lri.org/projects/eco-46-improved-aquatic-testing-and-assessment-of-cationic-polymers-itap/). KK thanks Indian Council of Medical Research, New Delhi, for financial support in the form of a senior research fellowship.

References

1. ECHA, Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Article 138 (2)
2. Mayo-Bean K, Moran K, Meylan B, Ranslow P (2012) Methodology document for the ECOlogical Structure-Activity Relationship Model (ECOSAR) class program. US-EPA, Washington D.C.
3. Boethling RS, Nabholz JV (1996) Environmental assessment of polymers under the US Toxic Substances Control Act. United States Environmental Protection Agency
4. USEPA (2013) Polymer guidance: https://www.epa.gov/sites/production/files/2015-03/documents/polyguid.pdf
5. Sanderson H, Thomsen M (2009) Comparative analysis of pharmaceuticals versus industrial chemicals acute aquatic toxicity classification according to the United Nations classification system for chemicals. Assessment of the (Q) SAR predictability of pharmaceuticals acute aquatic toxicity and their predominant acute toxic mode-of-action. Toxicol Lett 187:84–93
6. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, Dearden J, Gramatica P, Martin YC, Todeschini R (2014) (Q)SAR modeling: where have you been? Where are you going to? J Med Chem 57:4977–5010

7. Connors KA, Dyer SD, Belanger SE (2017) Advancing the quality of environmental microplastic research. Environ Toxicol Chem 36 (7):1697–1703
8. Biesinger KE, Stokes GN (1986) Effects of synthetic polyelectrolytes on selected aquatic organisms. J Water Pollut Control Fed 58:207–213
9. OECD (2019) Guidance document on aqueous-phase aquatic toxicity testing of difficult test chemicals. Series on testing and assessment. No. 23 (second edition). Paris, 81p
10. USEPA (1996) Ecological effects test guidelines OPPTS 850.1085 fish acute toxicity mitigated by humic acid. EPA712–C–96–117. Washington D.C., p 10
11. de Rosemond SJ, Liber K (2004) Wastewater treatment polymers identified as the toxic component of a diamond mine effluent. Environ Toxicol Chem 23:2234–2242
12. Liber K, Weber L, Levesque C (2005) Sublethal toxicity of two wastewater treatment polymers to lake trout fry (Salvelinus namaycush). Chemosphere 61:1123–1133
13. Cumming J, Hawker D, Matthews C, Chapman H, Nugent K (2010) Analysis of polymeric quaternary ammonium salts as found in cosmetics by metachromatic polyelectrolyte titration. Toxicol Environ Chem 92:1595–1608
14. Siebert J, Luyt A, Ackermann C (1990) A new transmission electron microscopic (TEM) method to determine differences between cationic polymers in solution. Int J Pharmaceut 61:157–160
15. Cumming JL, Hawker DW, Nugent KW, Chapman HF (2008) Ecotoxicities of polyquaterniums and their associated polyelectrolyte-surfactant aggregates (PSA) to Gambusia holbrooki. J Environ Sci Heal A 43:113–117
16. Cumming J, Hawker D, Chapman H, Nugent K (2011) The fate of polymeric quaternary ammonium salts from cosmetics in wastewater treatment plants. Water Air Soil Pollut 216:441–450
17. Cumming JL (2008) Environmental fate, aquatic toxicology and risk assessment of polymeric quaternary ammonium salts from cosmetic uses. Griffith University, Mount Gravatt
18. Pereira JL, Vidal R, Goncalves FJM, Gabriel RG, Costa R, Rasteiro MG (2018) Is the aquatic toxicity of cationic polyelectrolytes predictable from selected physical properties? Chemosphere 202:145–153
19. Nolte TM, Peijnenburg WJ, Hendriks AJ, van de Meent D (2017) Quantitative structure-activity relationships for green algae growth inhibition by polymer particles. Chemosphere 179:49–56
20. Khan K, Baderna D, Cappelli C, Toma C, Lombardo A, Roy K, Benfenati E (2019) Ecotoxicological (Q)SAR modeling of organic compounds against fish: application of fragment based descriptors in feature analysis. Aquat Toxicol 212:162–174
21. Khan K, Khan PM, Lavado G, Valsecchi C, Pasqualini J, Baderna D, Marzo M, Lombardo A, Roy K, Benfenati E (2019) (Q) SAR modeling of Daphnia magna and fish toxicities of biocides using 2D descriptors. Chemosphere 229:8–17
22. Khan K, Roy K, Benfenati E (2019) Ecotoxicological (Q)SAR modeling of endocrine disruptor chemicals. J Haz Mat 369:707–718
23. Khan K, Kar S, Sanderson H, Roy K, Leszczynski J (2018) Ecotoxicological assessment of pharmaceuticals using computational toxicology approaches: QSTR and interspecies QTTR modeling. In: Proceedings of MOL2-NET 2017, international conference on multidisciplinary sciences, 3rd edn. MDPI AG, Switzerland, Basel, p 1
24. Khan K, Kar S, Sanderson H, Roy K, Leszczynski J (2019) Ecotoxicological modeling, ranking and prioritization of pharmaceuticals using QSTR and i-QSTTR approaches: application of 2D and fragment based descriptors. Mol Inform, 38, article 1800078, http://dx.doi.org/10.1002/minf.201800078
25. Khan K, Benfenati E, Roy K (2019) Consensus (Q)SAR modeling of toxicity of pharmaceuticals to different aquatic organisms: ranking and prioritization of the DrugBank database compounds. Ecotox Environ Safe 168:287–297
26. Khan K, Roy K (2017) Ecotoxicological modelling of cosmetics for aquatic organisms: a QSTR approach. SAR (Q)SAR Environ Res 28:567–594
27. De P, Kar S, Roy K, Leszczynski J (2018) Second generation periodic table-based descriptors to encode toxicity of metal oxide nanoparticles to multiple species: QSTR modeling for exploration of toxicity mechanisms. Environ Sci Nano 5:2742–2760
28. Braakhuis HM, Kloet SK, Kezic S, Kuper F, Park MV, Bellmann S, van der Zande M, Le Gac S, Krystek P, Peters RJ (2015) Progress and future of in vitro models to study translocation of nanoparticles. Arch Toxicol 89:1469–1495
29. ECHA European Chemicals Agency (2012) Guidance on registration. Version 2.0. Guidance for the implementation of REACH

30. Netzeva T, Pavan M, Worth A (2007) Review of data sources, (Q)SARs and integrated testing strategies for aquatic toxicity. European Communities, Luxembourg
31. Roy K, Ambure P, Kar S, Ojha PK (2018) Is it possible to improve the quality of predictions from an "intelligent" use of multiple (Q)SAR/QSPR/QSTR models? J Chemom 32:e2992
32. Roy K, Ambure P, Kar S (2018) How precise are our quantitative structure-activity relationship derived predictions for new query chemicals? ACS Omega 3:11392–11406
33. Enslein K, Gombar VK (1997) TOPKAT 5.0 and modulation of toxicity. Mutat Res-Fund Mol M 379:S14–S14
34. Plošnik A, Zupan J, Vračko M (2015) Evaluation of toxic endpoints for a set of cosmetic ingredients with CAESAR models. Chemosphere 120:492–499
35. De Vaugelade S, Nicol E, Vujovic S, Bourcier S, Pirnay S, Bouchonnet S (2018) Ultraviolet-visible phototransformation of dehydroacetic acid – structural characterization of photoproducts and global ecotoxicity. Rapid Commun Mass Spectrom 32:862–870
36. Fendinger NJ, McAvoy DC, Eckhoff WS, Price BB (1997) Environmental occurrence of polydimethylsiloxane. Env Sci Technol 31:1555–1563
37. Khan PM, Roy K (2018) QSPR modelling for prediction of glass transition temperature of diverse polymers. SAR (Q)SAR Environ Res 29:935–956
38. Roy K, Kar S, Das RN (2015) A primer on (Q)SAR/QSPR modeling: fundamental concepts. Springer, UK. https://www.rsc.org/journals-books-databases/about-journals/environmental-science-nano/
39. Roy K, Kar S, Das RN (2015) Statistical methods in (Q)SAR/QSPR. In: A primer on (Q)SAR/QSPR modeling. Springer, NY, USA, pp 37–59
40. Mauri A, Consonni V, Pavan M, Todeschini R, software D (2006) An easy approach to molecular descriptor calculations. Match 56:237–248
41. Kuz'min VE, Artemenko AG, Polischuk PG, Muratov EN, Hromov AI, Liahovskiy AV, Andronati SA, Makan SY (2005) Hierarchic system of (Q)SAR models (1D–4D) on the base of simplex representation of molecular structure. J Mol Model 11:457–467
42. Alvadesc (2019) https://www.alvascience.com/alvadesc/
43. Yap CW (2011) PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J Comput Chem 32:1466–1474
44. Kennard RW, Stone LA (1969) Computer aided design of experiments. Technometrics 11:137–148
45. Golmohammadi H, Dashtbozorgi Z, Acree WE Jr (2012) Quantitative structure-activity relationship prediction of blood-to-brain partitioning behavior using support vector machine. Eur J Pharm Sci 47:421–429
46. Zhang Q, Couloigner I (2005) A new and efficient k-medoid algorithm for spatial clustering. In: International conference on computational science and its applications. Springer, NY, USA, pp 181–189
47. Roy K, Ambure P (2016) The "double cross-validation" software tool for MLR (Q)SAR model development. Chemom Intell Lab Syst 159:108–126
48. De P, Aher RB, Roy K (2018) Chemometric modeling of larvicidal activity of plant derived compounds against zika virus vector Aedes aegypti: application of ETA indices. RSC Adv 8:4662–4670
49. Roy K, Das RN, Ambure P, Aher RB (2016) Be aware of error measures. Further studies on validation of predictive (Q)SAR models. Chemom Intell Lab Syst 152:18–33
50. Langham AA, Khandelia H, Schuster B, Waring AJ, Lehrer RI, Kaznessis YN (2008) Correlation between simulated physicochemical properties and hemolycity of protegrin-like antimicrobial peptides: predicting experimental toxicity. Peptides 29:1085–1093
51. Khan PM, Roy K (2019) Consensus QSPR modelling for the prediction of cellular response and fibrinogen adsorption to the surface of polymeric biomaterials. SAR (Q)SAR Environ Res 30:363–382
52. Kholodovych V, Smith JR, Knight D, Abramson S, Kohn J, Welsh WJ (2004) Accurate predictions of cellular response using QSPR: a feasibility test of rational design of polymeric biomaterials. Polymer 45:7367–7379
53. J.R. Smith, D. Knight, J. Kohn, K. Rasheed, N. Weber, S. Abramson (2003) Using non-linear regression to predict bioresponse in a combinatorial library of biodegradable polymers, MRS Online Proc Libr, vol 804, Cambridge, UK
54. Smith JR, Knight D, Kohn J, Rasheed K, Weber N, Kholodovych V, Welsh WJ (2004) Using surrogate modeling in the prediction of fibrinogen adsorption onto polymer surfaces. J Chem Inform Comput Sci 44:1088–1097

55. Khan PM, Rasulev B, Roy K (2018) QSPR modeling of the refractive index for diverse polymers using 2D descriptors. ACS Omega 3:13374–13386
56. R Duchowicz P, C Comelli N, V Ortiz E, A Castro E (2012) (Q)SAR study for carcinogenicity in a large set of organic compounds. Current Drug Saf 7:282–288
57. Talevi A, L Bellera C, Di Ianni M, R Duchowicz P, E Bruno-Blanch L, A Castro E (2012) An integrated drug development approach applying topological descriptors. Curr Comput Aided Drug Des 8:172–181
58. Gajewicz A, Jagiello K, Cronin MTD, Leszczynski J, Puzyn T (2017) Addressing a bottle neck for regulation of nanomaterials: quantitative read-across (Nano-QRA) algorithm for cases when only limited data is available. Environ Sci Nano 4:346–358