Using model-based screening to help discover unknown environmental contaminants

Michael S. McLachlan*, Amelie Kierkegaard, Michael Radke, Anna Sobek, Anna Malmvärn, Tomas Alsberg

Department of Applied Environmental Science (ITM), Stockholm University, Stockholm, Sweden

Jon A. Arnot, Trevor N. Brown†, Frank Wania

Department of Physical and Environmental Sciences, University of Toronto Scarborough, Toronto, Ontario, Canada

†Current address: Department of Analytical Environmental Chemistry, UFZ-Helmholtz Centre for Environmental Research, Leipzig, Germany

Knut Breivik

Norwegian Institute for Air Research, Kjeller, Norway; Department of Chemistry, University of Oslo, Oslo, Norway

Shihe Xu

Health and Environmental Sciences, Dow Corning Corporation, Midland, Michigan, United States

*corresponding author: Department of Applied Environmental Science (ITM), Stockholm University, SE-106 91 Stockholm, Sweden, tel.: +46 8 674 7228; email: michael.mclachlan@itm.su.se

Abstract

Of the tens of thousands of chemicals in use, only a small fraction has been analyzed in environmental samples. To effectively identify environmental contaminants, methods to prioritize chemicals for analytical method development are required. We used a high throughput model of chemical emissions, fate, and bioaccumulation to identify chemicals likely to have high concentrations in specific environmental media, and we prioritized these for target analysis. This model-based screening was applied to 215 organosilicon chemicals culled from industrial chemical production statistics. The model-based screening prioritized several recognized organosilicon contaminants and generated hypotheses leading to the selection of three chemicals that have not previously been

31   identified as potential environmental contaminants for target analysis. Trace analytical

32   methods were developed and the chemicals were analyzed in air, sewage sludge, and

33   sediment. All three substances were found to be environmental contaminants. Phenyl-

34   tris(trimethylsiloxy)silane was present in all samples analyzed, with concentrations of ~50

35   pg m^{-3} in Stockholm air and ~0.5 ng g^{-1} dw in sediment from the Stockholm archipelago.

36   Tris(trifluoropropyl)trimethyl-cyclotrisiloxane and tetrakis(trifluoropropyl)tetramethyl-

37   cyclotetrasiloxane were found in sediments from Lake Mjøsa at ~1 ng g^{-1} dw. The

38   discovery of three novel environmental contaminants shows that models can be useful for

39   prioritizing chemicals for exploratory assessment.

40   **Introduction**

41   Thousands of organic chemicals are being used in industrial and household applications,

42   and every year several hundred additional chemicals are introduced into the market.

43   Various regulatory programs seek to protect human health and the environment from the

44   risks that may be posed by chemicals (e.g., REACH).[1] Risk is a function of exposure;

45   however, biomonitoring and monitoring data are generally lacking for exposure and risk

46   characterization for humans and the environment for the vast majority chemicals.[2,3] Our

47   knowledge of the contamination status of the environment is also limited because only a

48   small fraction of these chemicals has been analyzed in environmental samples.. Due to

49   the high cost of developing and applying methods for the trace analysis of chemicals in

50   complex environmental matrices, it is not feasible to measure all chemicals in use.

51   Methods for prioritizing chemicals for target analysis are required and these methods

52   need to be evaluated.

53   Environmental contaminants have frequently been discovered by analytical chemists who

54   observe unknown signals during the trace analysis of environmental matrices.[4-7] However,

55   this approach is biased towards identifying contaminants which are similar to known

56   contaminants. Non-target screening is a less biased method for contaminant discovery;

57   typically it is based on gas or liquid chromatography coupled to mass spectrometry. It has

58   been applied with success.[8-9] It is currently limited by difficulties in data processing and in

59   assigning specific molecular structures to individual peaks[10]. An alternative approach that

60   has been developed over the last 20 years is effect directed analysis.[11] Starting from an

61   adverse effect observed for a specific ecotoxicological endpoint, extract fractionation is

62   used to isolate and identify the causative agent. This approach is particularly useful

63   because it focuses on contaminants of ecotoxicological relevance, but it has proven

64   challenging to implement[12] and is limited to the proper choice of endpoints. A common

65   feature of the existing methods to identify new environmental contaminants is that they

66   are limited to identifying the contaminant; they do not provide mechanistic insight into

67   possible source/receptor relationships that could help in determining causes of the

68   contamination or developing control strategies.

69   Detecting and identifying a previously undiscovered contaminant in the environment and

70   then working backwards to infer its source(s) and modes of transport has many

71   challenges, as discussed above. An alternative is the reverse approach, starting with

72   chemicals of interest and working forwards to identify potential sources and modes of

73   transport and – eventually – confirming their presence as a contaminant in the

74   environment. This approach is also challenging, as predicting the sources and modes of

75   transport of a chemical requires a quantitative understanding of the chemical and

76   environmental properties involved.

77   Despite these difficulties the theoretical screening of large numbers of chemicals to

78   identify those with the greatest potential to be environmental exposure hazards has

79   proven useful.[13-14] Initially this approach was applied to estimate a limited number of

80   screening criteria such as persistence and bioaccumulation.[15-17] More recently, it has been

81 extended to screening chemicals for the magnitude of potential environmental

82 contamination and organism exposure.[18-19] This was made possible by the development of

83 linked environmental fate and bioaccumulation models, together with tools that can

84 predict physical chemical properties for large numbers of substances. The next step that

85 made screening for actual, rather than potential, exposure possible was the development

86 of high throughput methods to screen chemicals for emissions.[20] Using these methods, a

87 list of 12,619 discrete organic chemicals was screened for far-field exposure, yielding

88 predicted concentrations in humans that ranged over 17 orders of magnitude.[21]

89 In this work we explore the hypothesis that fate and exposure model-based screening

90 methods can identify contaminants of concern and identify unknown environmental

91 contaminants. We use high throughput models to identify chemicals that are likely to

92 have high concentrations in specific environmental media including humans. The

93 chemicals selected with this model-based screening are fed into a targeted analytical

94 screening which concludes with the target analysis of selected compounds in

95 environmental samples. The procedure is applied to a list of organosilicon chemicals.

96 **Methods**

97 <u>Model-based screening</u>

98 The model-based screening was conducted with the Risk Assessment IDentification And

99 Ranking (RAIDAR) version 2.0 model.[22] RAIDAR couples a regional-scale multimedia

100 environmental fate and transport mass balance model with bioaccumulation mass

101 balance models to calculate exposures and potential risks to humans and various aquatic

102 and terrestrial ecological receptors. RAIDAR is a steady state model of the "Unit World"

103 type (i.e., there is no spatial resolution of the environmental compartments air, water, soil

104 and sediment), which is a generic model format that has proven particularly useful for

105 screening purposes.[23] The model can be used to simulate multimedia, multipathway far-

106   field human exposures and a version of the model has been programmed in Excel™ to

107   facilitate high throughput calculations for large numbers of chemicals.[21]

108   RAIDAR requires user input of information on chemical emissions and chemical

109   properties. To estimate emissions for large numbers of chemicals, a recently developed

110   rapid throughput method was employed.[20] The method is based upon the European

111   Union Technical Guidance Document and uses information on chemical quantities in

112   commerce (production and/or import rates), chemical function (use patterns) and

113   physical chemical properties to estimate emissions to air, soil and water for five stages of

114   the chemical life-cycle. The chemical properties required by the model (molar mass,

115   octanol-water partition coefficient, Henry's law constant, degradation half-lives in air,

116   water, soil and sediment, and primary biotransformation half–lives in vertebrates) were

117   either based on measurements or predictions obtained from the U.S. Environmental

118   Protection Agency's Estimation Program Interface Suite (EPI Suite™) software program.[24]

119   The integration of the chemical property estimation procedures, emissions scenarios, and

120   the RAIDAR model to screen for chemical exposure is described in detail elsewhere.[21]

121   The chemicals to be screened were taken from a database of 12,619 organic chemicals.

122   This list included industrial chemicals with reported production and/or import in Europe,

123   the United States, Canada, Japan, and OECD countries. Further details of this list are given

124   elsewhere.[21] To provide a manageable subset for targeted chemical screening, it was

125   decided to focus on organosilicon chemicals. This selection was made because there was

126   little reported research on organosilicon chemicals as environmental contaminants, which

127   made this chemical group particularly suitable for evaluating a screening procedure for

128   identifying unknown contaminants. All chemicals that contained silicon were extracted

129   from the database. This yielded 287 chemicals. They were then screened and 72

130   substances were removed that were not discrete chemicals (e.g., technical mixtures),

131   were inorganic (i.e., had been mistakenly classified as organic in the original database), or

132   were quaternary amines (which are well outside of the domain of the property prediction

133   software used, see below). This left 215 distinct organic chemicals that formed the basis

134   for the model-based screening (see Table S1 in the Supporting Information).

135   The screening was performed using a stepwise exclusion procedure. First the steady state

136   concentrations of the chemicals were predicted using the RAIDAR model, and the

137   chemicals were ranked according to their median concentrations in different

138   environmental media and biota. The first filter was to set lower limits for the predicted

139   concentrations in the environment. The purpose of this filter was to exclude chemicals

140   that were likely to be below the detection limits of an eventual targeted analytical

141   screening, and for which screening would thus be pointless. Air and sediment were the

142   matrices selected for the targeted analytical screening (for motivation see below), and the

143   thresholds were set to 10^{-13} moles m^{-3} and 10^{-13} moles cm^{-3}, respectively. The second filter

144   was based on predicted human exposure. The purpose of this filter was to restrict the list

145   to chemicals that higher organisms were exposed to, as such chemicals are most likely to

146   pose an environmental risk. Humans were used as a surrogate for higher organisms in this

147   context. The chemicals were ranked according to their predicted concentration in humans

148   (highest concentration = rank 1), and chemicals with a human concentration rank >65

149   were excluded. The third filter was to exclude known environmental contaminants.

150   The Hydrowin model in EPI Suite™ provides useful information for hydrolysis reactions

151   including "rules of thumb" and half-lives when entering chemicals one at a time; however,

152   in "batch mode" for high throughput screening many chemicals cannot be treated by the

153   model, i.e., it stops running when it encounters a structure for which it cannot estimate a

154   hydrolysis rate. Therefore, we did not use Hydrowin and did not include hydrolysis half-

155   lives as RAIDAR model input for initial screening. However, some organosilicon chemicals

156   are readily hydrolyzed,[25] whereby the rates are highly variable, depending on three major

157   factors. First, for an organosilicon compound with a Si-X bond (where X=a hydrolysable

158   functional group), the rate of hydrolysis will depend on the polarity of the Si-X bond. For

159   example, the hydrolysis rates generally follow the order: organohalosilanes >

160   carboxysilanes = oximinosilanes > aminosilanes = alkoxysilanes. In addition, hydrolysis

161   rates increase with increase of the number of Si-X bonds per Si atom. Finally, the size of

162   the non-hydrolyzable functional group attached to the Si atom also influences the

163   hydrolysis rates. Bulky functional groups such as phenyl and t-butyl usually dramatically

164   reduce the hydrolysis rates relative to a smaller group such as methyl, while a small group

165   such as H may increase the hydrolysis rates. Based on the above considerations, three

166   bins were created to distinguish the organosilicon compounds according to their

167   hydrolytical reactivity. Those chemicals in the most reactive bin were excluded.

168   We emphasize that the stepwise exclusion procedure was intended to prioritize chemicals

169   for target analytical screening. This procedure can be structured in different manners,

170   depending on the properties of potential contaminants that one wishes to prioritize.

171   <u>Targeted analytical screening</u>

172   The chemicals identified in the model-based screening procedure were passed on to the

173   targeted analytical screening stage. The first step in this stage was to select a sub-set of

174   the chemicals for method development in order to focus the available analytical

175   resources. Exclusion criteria for this step included lack of availability of standards and

176   structural similarity to known organosilicon contaminants (a conscious decision was made

177   to focus on those chemicals with structures that were distinctly different from known

178   organosilicon contaminants).

179   The second step was to test whether the chemicals could be detected using the available

180   analytical instrumentation. Standard solutions were prepared in n-hexane/

181   dichloromethane and in one case in ethyl acetate. These solutions were then analyzed

182   using GC/MS with electron ionization employing both splitless and on-column injection. If

183   the chemical was not detected it was analyzed using negative chemical ionization and/or

184   positive chemical ionization. For several chemicals for which none of these techniques

185   was successful, LC/MS was tested. Standard solutions for LC/MS were prepared in

186   methanol and acetonitrile. The fragmentation patterns using electrospray ionization were

187   investigated with direct infusion QTOF-MS, and UHPLC/MS-MS methods were developed

188   for the screening of environmental samples. More details on the analytical parameters

189   used are provided in Table S2.

190   The third step was to exclude chemicals that were not stable in solution. If the chemical

191   concentration in a standard solution (in organic solvent) decreased by an order of

192   magnitude or more during storage for one month in the refrigerator, then it was

193   considered unlikely that the chemical was sufficiently persistent in the environment to be

194   a detectable contaminant, and it was excluded from further consideration.

195   The fourth step was target analysis of environmental samples. A tiered approach was

196   employed in selecting samples for analysis. The guiding principle was to begin with

197   environmental samples where the levels were expected to be highest, and – if the

198   chemicals were found there – to proceed to samples in which lower levels were expected.

199   Air and sewage sludge / sediment were chosen as the primary matrices for screening: air

200   because the modeling showed that many organosilicon chemicals are volatile and

201   distribute preferentially into air; sewage sludge / sediment because we expected many

202   organosilicons to be emitted to water (this is the case for a significant fraction of the

203   known organosilicon contaminants) and the modeling showed that their hydrophobic

204   properties would lead to their sequestration to sediment.

205   Initially air samples were collected at the Bromma wastewater treatment plant (WWTP) in

206   Stockholm. The plant serves 300 000 population equivalents and the secondary activated

207   sludge treatment is housed in an enclosed building. Aeration of municipal wastewater in

208   an urban environment with limited air dilution combines different factors (proximity to

209   sources, efficient mass transfer, limited dilution) that can contribute to elevated

210   contaminant concentrations in air. The samples were collected on 15-25 mL cartridges

211   containing 80-140 mg of Isolute ENV+ (Biotage AB, Uppsala, Sweden) following the

212   procedure described elsewhere.[26] For the second tier, extracts of air samples that had

213   been collected for another study in Zürich during a period with a strong thermal

214   inversion[27] were also subjected to GC analysis. In that study cyclic volatile methylsiloxanes

215   were analyzed, and the sample collection and extraction procedures were the same as

216   used here. In a third tier effort, air samples were collected on the Stockholm University

217   campus, which is located within the city of Stockholm, using the same method as

218   employed at the Bromma WWTP. Sampling was done on two occasions, and in each case

219   duplicate samples were collected in parallel. Details about the collection of the air

220   samples are provided in Table S3.

221   Sewage sludge was the matrix used in the first tier of the screening for hydrophobic

222   organosilicons. Samples of digested and dewatered sludge were collected from the

223   Bromma and Henriksdal WWTPs in Stockholm during 2013. In the second tier, surface

224   sediment was analyzed, beginning with 4 samples collected during September 2011 on a

225   transect from Stockholm Harbor, which is the recipient for Henriksdal and Bromma WWTP

226   effluents for about 1 million person equivalents, through the inner Stockholm

227   Archipelago. Sediment samples collected for other studies from two Swedish lakes[28] and

228   Lakes Mjøsa and Femunden in Southern Norway[29] were also analyzed. Details of the

229   sewage sludge and sediment samples are provided in Table S4.

230 Target analysis was conducted for three chemicals; GC was used for one chemical (phenyl-

231 tris(trimethylsiloxy)silane (M3TPh)), and LC for the other two chemicals

232 (tris(trifluoropropyl)trimethyl-cyclotrisiloxane (D3F), and

233 tetrakis(trifluoropropyl)tetramethyl-cyclotetrasiloxane (D4F)). For the air samples, the

234 ENV+ cartridges were extracted with 3-10 fractions of 300 μL of dichloromethane (air

235 from the WWTP for GC analysis), 1.3 mL of n-hexane (air from the university for GC

236 analysis) or 3 x 300 μL acetonitrile (for LC analysis). Each fraction was analyzed individually

237 without cleanup or concentration of the extract.

238 For the analysis of sludge and sediment, a wet sample of 10 g (dewatered sludge) or 25 g

239 (centrifuged sediment) was extracted with acetone (20-25 mL) and dichloromethane (1.5

240 mL). The acetone quantity was optimized to produce one single liquid phase, the water

241 included. The samples were ultrasonicated for 2 x 15 min. After centrifugation the liquid

242 phase was decanted and the sample was re-extracted with 2 mL (sludge) or 10 mL

243 (sediment) acetone together with 5 mL dichloromethane. The extracts were combined

244 and dichloromethane was carefully added to the single liquid phase until a

245 dichloromethane/acetone phase formed that was distinct from the water phase. The

246 organic phase was transferred into 250 mL flasks and cleaned up using a purge and trap

247 technique described elsewhere.[29] The trap consisted of an ENV+ cartridge that was eluted

248 with 3 fractions of 300 μL of acetonitrile for the analysis of D3F and D4F. For the analysis

249 of M3TPh the combined acetonitrile fractions were quantitatively transferred to n-hexane

250 (1+1 mL) by the addition of 1.5 mL of water.

251 M3TPh was analyzed by GC/MS(EI) while D3F and D4F were analyzed by UHPLC/MS-MS

252 with electrospray ionization. Quantification was accomplished using standard addition to

253 duplicate samples. Standard addition was done into the centrifuged sediment with a

254   syringe after the addition of a 1.5 mL layer of dichloromethane. The instrumental

255   methods are summarized in Table S5.

256   **Results and discussion**

257   <u>Model-based screening</u>

258   The results of the model-based screening are summarized in the TOC art. RAIDAR

259   predicted concentrations that ranged from 10^{-7} - 10^{-19} mol m^{-3} in air, 10^{-11} - 10^{-19} mol cm^{-3}

260   in sediment, and 10^{-4} - 10^{-16} mol m^{-3} (or mmol kg^{-1} assuming a whole body density of 1000

261   kg/m^3) in humans for the 215 organosilicon chemicals (see Table S1). Of these, 52

262   chemicals exceeded the threshold concentration for air and 24 the threshold

263   concentration for sediment in the first filter. There were 13 substances common to these

264   two groups, leaving 63 substances after the environmental concentration filter.

265   The human exposure filter resulted in the elimination of a further 24 chemicals. Most of

266   these were substances with high predicted concentrations in air; only one of the

267   eliminated chemicals had a high predicted concentration in sediment. This suggests that

268   human exposure is correlated more strongly with concentration in sediment than with

269   concentration in air. This can be explained by the fact that chemicals with high

270   concentrations in sediment often possess a tendency to partition to organic matter. This

271   tendency also fosters bioaccumulation (i.e., chemicals which partition into sediment will

272   also partition into biota).

273   Seven of the chemicals that passed the second filter were known environmental

274   contaminants, namely hexamethyldisiloxane (L2), octamethyltrisiloxane (L3),

275   decamethyltetrasiloxane (L4), hexamethylcyclotrisiloxane (D3),

276   octamethylcyclotetrasiloxane (D4), decamethylcyclopentasiloxane (D5), and

277   dodecamethylcyclohexasiloxane (D6). D4, D5 and D6, three known organosilicon

278   contaminants that have been subjected to particularly intensive regulatory assessment,[30-

279   32] ranked in the top four based on concentration rankings for both predicted sediment

280   and human concentrations. The presence of these known environmental contaminants at

281   the top of the screened list provides some confidence about the relevance and

282   effectiveness of the screening procedure for identifying chemicals of concern.

283   Of the 32 remaining chemicals, 19 were estimated to be rapidly hydrolyzed. They were

284   removed from the chemical list. A further 7 were classified as being moderately

285   susceptible to hydrolysis, but were retained. The chemical list after this filter is provided

286   in Table S6.

287   The model-based screening procedure reduced the original list of 215 discrete chemicals

288   to 13 chemicals that were carried through to targeted analytical screening. Redundancy in

289   the screening procedure can be assessed using the selectiveness of each filter stage; if a

290   filter removes only a small fraction of the chemicals passing through it, it may be possible

291   to omit it from the screening procedure without having a large impact on the screening

292   final outcome. The initial concentration-based filter was the most selective, removing 71%

293   of the chemicals, followed by the reactivity filter, which removed 59% of the chemicals

294   that passed through it (TOC art). The human exposure filters was also quite effective,

295   removing 38% of the chemicals passing through it. The known contaminant filter was the

296   least selective (18%), but the organosilicons were selected for this evaluation because

297   there are few known organosilicon contaminants. Other chemical groups could have a

298   larger fraction in this category, and then this filter could be more important. In conclusion,

299   each filter stage in the screening procedure was useful; none was redundant.

300   <u>Targeted analytical screening</u>

301   Of the 13 chemicals identified by the model-based screening, eight were selected for

302   analytical method development. The other five were not selected for a number of reasons

303   including lack of commercial availability at the time of method development, and

304   structural similarity to known organosilicon contaminants.[26,33]

305   Detection methods were successfully developed for five of the chemicals selected for

306   targeted analytical screening (see Table S6). It is possible that a more extensive effort

307   could have produced detection methods for the remaining three chemicals. However, it is

308   also possible that the negative result was due to the instability of the chemicals in

309   solution; all of the remaining three chemicals were classified as moderately susceptible to

310   hydrolysis in the three bin categorization system.

311   The stability test showed that one of the five detectable chemicals was unstable in

312   solution (see Table S6). This chemical was classified as moderately susceptible to

313   hydrolysis. All of the other four chemicals were classified as less susceptible to hydrolysis.

314   The four detectable and stable chemicals identified by the screening procedure are shown

315   in Figure 1. Environmental samples were screened for three of these chemicals that were

316   compatible with our tried and proven purge and trap clean-up method for cyclic volatile

317   methylsiloxanes, namely phenyl-tris(trimethylsiloxy)silane (M3TPh) (CAS# 2116849),

318   tris(trifluoropropyl)trimethyl-cyclotrisiloxane (D3F) (CAS# 2374143), and

319   tetrakis(trifluoropropyl)tetramethyl-cyclotetrasiloxane (D4F) (CAS# 429674). The fourth

320   chemical tetraphenyl-tetramethyl-trisiloxane (MDMPh4) (CAS# 3982829)was much less

321   volatile and displayed poor recovery in the purge and trap clean-up step.

322   *M3TPh*

323   M3TPh was found in several matrices. The concentrations were sufficiently high and the

324   analytical method was sufficiently specific that most features of the M3TPh mass

325   spectrum were observed in full scan spectra of environmental samples. This is illustrated

326   in Figure S1, which compares the mass spectra in a standard, a sewage sludge sample, and

327   a sediment sample.

328   The M3TPh concentrations measured in air are given in Table 1. The concentrations were

329   highest in the WWTP, at 5.5 ng m^{-3}. They were a little more than an order of magnitude

330   lower in Zürich during the inversion event and a further order of magnitude lower on the

331   Stockholm University campus. There was good agreement between the parallel samples

332   collected in Stockholm (difference <5%). The compound was not detected in the field

333   blanks from the Zürich campaign. The LOQ (5 times the noise) was 0.02 ng m^{-3}, which is

334   12-26 times lower than the concentrations in the Zürich samples. This indicates that the

335   levels in these samples were not due to contamination in the field or the laboratory. In

336   summary, the air sampling provided strong evidence that M3TPh is an atmospheric

337   contaminant.

338   M3TPh was found in all sewage sludge and sediment samples (Table 1). As with the air

339   samples, the parallel analyses of Henriksdal sludge and Beckholmen sediment gave similar

340   concentrations (difference <15%). The mean recovery of the method was 78% (CV 8%,

341   n=7) for the sediment samples and 54% for sewage sludge. The good reproducibility and

342   recovery provides confidence in the quality of the analytical method. No field blanks were

343   available for the sediment samples. The least contaminated sediments provide an upper

344   estimate for the field blank. The M3TPh contents of the Hjulstafjärden and Femunden

345   samples were the lowest measured (0.25 and 0.45 ng sample^{-1}, respectively).

346   The M3TPh concentrations in sewage sludge from the two Stockholm WWTPs were ~500

347   ng g^{-1} dw (Table 1). The amount of M3TPh in the procedural blank was >3 orders of

348   magnitude lower than in the samples. This indicated that M3TPh is likely being released to

349   the aquatic environment, and led to the analysis of sediment samples.

350   Sediment collected at Beckholmen in Stockholm Harbor close to the outfalls of the

351   Henriksdal and Bromma WWTPs contained 20 ng g^{-1} dw of M3TPh. This was ~25 times

352   lower than the concentrations measured in the sewage sludge. The M3TPh quantity in the

353   sample exceeded the upper estimate of the M3TPh quantity in the field blank by >2

354   orders of magnitude, indicating that M3TPh was present in the sediment as an

355   environmental contaminant.

356   A clear M3TPh concentration gradient was observed from the WWTPs along the transect

357   through the inner Stockholm archipelago out towards the Baltic Sea. Moving from

358   Beckholmen towards the open Baltic Sea, the M3TPh concentration decreased to 3.2 ng g^{-}

359   ^{1} dw at Fisksätra and further to 0.38 and 0.46 ng g^{-1} dw at the two outer stations,

360   Gustavsberg and Baggensfjärden. These observations indicate that M3TPh is being

361   emitted from the Stockholm WWTPs, resulting in contamination of the recipient and a

362   contamination plume that extends out towards the Baltic Sea.

363   Sediments were also analyzed from two lakes in Sweden and two lakes in Norway. In each

364   country, one of the lakes had much higher concentrations than the other (see Table 1). In

365   each case, the lake with the higher M3TPh concentration (Östersjön in Sweden and Mjøsa

366   in Norway) was significantly impacted by WWTP effluent and had been shown to be

367   contaminated with D5, a known organosilicon contaminant originating from WWTPs,

368   while the lake with the lower M3TPh concentration (Hjulstafjärden in Sweden and

369   Femunden in Norway) was much less impacted by WWTP effluent and had been shown to

370   be much less contaminated with D5.[28-29] This is further evidence that M3TPh is an

371   environmental contaminant in aquatic ecosystems that originates from WWTPs.

372   *D3F*

373   This chemical presented interesting analytical challenges. MS/MS analysis of standard

374   solutions prepared in acetonitrile and analyzed using direct infusion in an acetonitrile :

375   water (1:1, v:v) mixture produced ions corresponding to $(M+17)^-$. Our interpretation was

376   that D3F was ionized after a hydrolysis reaction resulting in breakage of the ring and

377   formation of a linear siloxanediol (D3FOH) through the addition of water (see Figure S2).

378   When the same standard was chromatographed in the reversed phase column using a

379   water/acetonitrile gradient as the mobile phase, two well separated peaks with identical

380   MS/MS transitions were obtained. We judged the early eluting peak to be D3FOH formed

381   in the standard prior to analysis, while the late eluting peak was D3FOH formed via

382   hydrolysis of D3F directly in the source. This was verified by adding water to a standard

383   solution of D3F in acetonitrile. As the fraction of water increased, the size of the early

384   eluting peak increased at the expense of a decreased late eluting peak (see Figure S3),

385   presumably as a result of greater hydrolysis of the D3F in the standard solution prior to

386   analysis. Moreover, when the standard in acetonitrile was subsequently analyzed using

387   isocratic elution with acetonitrile, the size of the later eluting peak was an order of

388   magnitude lower, which we attribute to less hydrolysis of D3F in the ion source in the

389   near absence of water. Scanning for parents of a neutral loss of 20 (many fragments in the

390   mass spectrum of the diol originates from the repeated loss of HF) indicated that D3FOH

391   was the major hydrolysis product formed. D3F was quantified using the sum of the areas

392   of the two peaks. This could result in an overestimation of the D3F concentrations, as

393   some fraction of the D3FOH present in the environmental samples may also have been

394   extracted. Since no standard was available for D3FOH, no response factor could be

395   determined; it was assumed to be equal to that of D3F. This adds uncertainty to the

396   quantification, but the D3F peak contributed on average 30% to the sum of the areas of

397   the two peaks.

398   In air, D3F was only analyzed in the samples collected at Stockholm University (both 2012

399   and 2013). Traces of D3FOH were found (~3-5 pg m^{-3}), but similar quantities were also

400 found in the procedural blanks (~3 pg m^{-3}). Thus no evidence was found that D3F is an

401 atmospheric contaminant.

402 The results of the targeted analytical screening of D3F + D3FOH in solid matrices are

403 summarized in Table 2. D3F + D3FOH were found in sewage sludge at levels ranging from

404 70-130 pg g^{-1} dw. This led to their analysis in sediment. In contrast to M3TPh, many

405 sediment samples contained D3F + D3FOH concentrations at the low end of the observed

406 range. Since there was no field blank, the sample with the lowest concentration

407 (Femunden) was used as a field blank, and samples with concentrations less than 3 times

408 this concentration were considered below the LOD. The levels in the sediment samples

409 collected in Stockholm Harbor ranged from <LOD-70 pg g^{-1} dw. No gradient in

410 concentrations was observed moving away from the harbor, but this could be due to the

411 proximity of the measured levels to the LOD or the low method precision arising from the

412 analytical challenges discussed above. These chemicals were detected in one of 2

413 sediment samples from Swedish lakes impacted by WWTPs (60 pg g^{-1} dw). Much higher

414 levels of D3F + D3FOH (1600 pg g^{-1} dw) were present in Lake Mjøsa sediment. This lake

415 has also been found to contain comparatively high levels of D5.[29] Taken together, the

416 results indicate that D3F + D3FOH are contaminants in some aquatic systems. Given the

417 evidence for relatively rapid hydrolysis of D3F in aqueous solution in the laboratory, its

418 presence in sediment is consistent with some portion of D3F being emitted in sorbed

419 form.

420 *D4F*

421 D4F presented analytical challenges similar to those encountered for D3F. The hydrolysis

422 product D4FOH was formed in the ion source, in the UPLC column, and in solution, which

423 hampered quantification. However, the ratio of the D4FOH:D4F signals was in general

424   smaller than the ratio of the D3FOH:D3F signals, indicating that D4F was less susceptible

425   to hydrolysis. On average D4F accounted for 95% of the sum of the D4F and D4FOH signal.

426   Like D3F + D3FOH, D4F + D4FOH were not found in the air samples from Stockholm

427   University at levels markedly above the method blank (~2 pg m^{-3}).

428   The results of the targeted analytical screening of D4F + D4FOH in sewage sludge and

429   sediment are summarized in Table 2. The LOD was 90 pg g^{-1} dw based on the sample with

430   the lowest concentration as described above. D4F + D4FOH were not detected in most of

431   the sediment samples. However, they were clearly present in the two sewage sludge

432   samples and two of the sediment samples. Again, comparatively high levels were

433   detected in Lake Mjøsa (1800 pg g^{-1} dw). This indicates that D4F + D4FOH are also

434   contaminants in some aquatic systems.

435   <u>Evaluation of the overall screening strategy</u>

436   The combination of model-based screening and targeted analytical screening proved to be

437   effective. Three unknown environmental contaminants were discovered. All of the

438   chemicals that were subjected to targeted analytical screening of environmental samples

439   based on the model-based screening were found. This high success rate was certainly to

440   some degree fortuitous and cannot be expected in all cases. However, it does suggest that

441   model-based screening combined with analytical pre-screening is useful in guiding the

442   search for unknown environmental contaminants.

443   This case study clearly illustrates that high throughput model-based screening methods

444   for chemical fate and exposure may result in a substantial number of false positives for

445   chemicals that are subject to hydrolysis. Environmentally relevant chemical hydrolysis

446   databases and revised high-throughput QSPR tools to predict hydrolysis rates are needed

447   to improve model-based screening and prioritization methods. To overcome this, we

448   developed a set of rules for organosilicon substances. However, this approach cannot be

449   readily applied to a broader spectrum of chemicals.

450   To obtain more insight into the performance of the model-based screening, the measured

451   concentrations of the target analytes were compared with the concentrations predicted

452   by the model. For M3TPh the median predicted concentration in air was 2 pg m^{-3} (3×10^{-3} -

453   2×10^{3} was the 95% uncertainty interval predicted by RAIDAR). This compares with the

454   measured concentrations at Stockholm University (the sampling site believed to best

455   represent regional background) of 44-72 pg m^{-3}. The model median under-predicted the

456   concentration in air by a factor of ~25. For sediment, the predicted median concentration

457   was 0.4 ng g^{-1} dw (5×10^{-4} - 3×10^{2}). This agrees well with the concentrations measured at

458   the seaward stations in the Stockholm archipelago (0.38 and 0.46 ng g^{-1} dw) and in the

459   remote lake Femunden (0.14 ng g^{-1} dw), while it is higher than measured in Hjulstafjärden

460   (0.015 ng g^{-1} dw).

461   For D3F and D4F the predicted median concentrations in air were 56 (4×10^{-2} - 7×10^{4}) and

462   0.4 (2×10^{-3} - 81) pg m^{-3}, respectively. Neither was detected in the air at a LOD of ~10 pg m^{-}

463   ^{3}. In sediment the predicted median concentrations were 36000 (40 - 3×10^{7}) pg g^{-1} dw and

464   700 (6 - 8×10^{4}) pg g^{-1} dw, respectively. For D3F the median prediction overestimates the

465   concentrations measured in most of the sediment samples (6-65 pg g^{-1} dw) by about 3

466   orders of magnitude (1 order of magnitude for Lake Mjøsa); however, some of the

467   measurements are within the predicted range including uncertainty. For D4F the

468   measured values in sediment from the Stockholm archipelago and Swedish lakes (30-300

469   pg g^{-1} dw) were similar to the median model prediction. The model's overestimation of

470   the D3F concentration in sediment is in part attributable to the fact that the model did

471   not consider hydrolysis. The analytical method development showed that D3F is

472   particularly susceptible to hydrolysis during analysis. Thus hydrolysis will likely be a

473   significant loss process in the aquatic environment.

474   The uncertainty intervals for the model predicted concentrations of these three chemicals

475   in air and sediment were approximately ±3 orders of magnitude and were largely driven

476   by uncertainty in emission rate estimates (see Arnot et al.[21] for a discussion of the

477   estimation of the uncertainty interval). The measured concentrations lie well within the

478   uncertainty intervals of the predictions. Nevertheless, the high model uncertainty can

479   have consequences for the results of the model-based screening. Underestimation of

480   environmental concentrations can lead to the screening out of chemicals that have high

481   concentrations in the environment (false negatives), while overestimation can lead to the

482   screening in of chemicals that have non-detectable concentrations in the environment

483   (false positives). In this work no false positives were identified, whereby it cannot be ruled

484   out that there were false positives after the air, sediment, and human exposure filters

485   that were then eliminated by subsequent filters in the screening process. It is possible

486   that the screening process did result in false negatives; however, we did not evaluate the

487   system for this possibility. Consequently, we do not conclude that chemicals screened out

488   in the model-based filters are not environmental contaminants. The fact that the well-

489   known organosilicon environmental contaminants were screened in and ranked highly

490   provides confidence in the capacity of the screening system to provide viable hypotheses

491   for targeted testing.

492   This work illustrates the potential of model-based screening to assist in identifying

493   unknown environmental contaminants. In addition, using emissions-based mechanistic

494   models at the early stages of contaminant discovery means that additional information

495   that they provide, for instance on sources, major chemical transport pathways and

496   responses to changes in emissions, is available to guide the further exploration of the

497   contaminant issue and the eventual development of effective chemical management. We

498   expect that model-based screening will become more useful as model uncertainty is

499   reduced. Currently, the emissions estimates are the dominant source of model

500   uncertainty.[21] Furthermore, the lack of high throughput methods to estimate hydrolysis

501   proved to be a particularly serious shortcoming in screening organosilicons.

502   Finally, we note that identifying an unknown environmental contaminant is only the first

503   step. Much more work is required to assess a contaminant's sources, environmental

504   behavior and toxicity before it can be established whether it is an environmental concern.

505   **Supporting Information**

506   Three figures (mass spectra of M3TPh, the structural formula of D3FOH and D4FOH, and

507   chromatograms showing the elution of D3F standard solutions containing different

508   fractions of water) and six tables (list of the 215 discrete organosilicon chemicals,

509   analytical conditions employed when screening for detectability, details of air sampling,

510   details of sediment sampling, instrumental methods, and list of chemicals after the

511   reactivity filter). This material is available free of charge via the Internet at

512   http://pubs.acs.org.

520 Stockholm sediment samples, and Cajsa Wahlberg (Stockholm Vatten) for providing

521 access to the Bromma WWTP for air sampling.

522 **References**

523 1. European Parliament, Regulation (ec) 1907/2006 of the european parliament and

524 of the council of 18 december 2006 concerning the registration, evaluation,

525 authorisation and restriction of chemicals (REACH), establishing a European

526 chemicals agency, amending directive 1999/45/ec and repealing council

527 regulation (eec) 793/93 and commission regulation (ec) 1488/94 as well as council

528 directive 76/769/eec and commission directives 91/155/eec, 93/67/eec,

529 93/105/ec and 2000/21/ec. *Official Journal of the European Union*. 30.12.2006. L

530 396/1- 849. **2006**.

531 2. Egeghy, P. P.; Judson, R.; Gangwal, S.; Mosher, S.; Smith, D.; Vail, J.; Cohen Hubal,

532 E. A. The exposure data landscape for manufactured chemicals. *Sci. Total Environ.*

533 **2012,** *414*, 159-166.

534 3. Egeghy, P. P.; Vallero, D. A.; Cohen Hubal, E. A. Exposure-based prioritization of

535 chemicals for risk assessment. *Environ. Sci. Policy* **2011**, *14*, 950-964.

536 4. Adolfsson-Erici, M.; Pettersson, M.; Parkkonen, J.; Sturve, J. Triclosan, a commonly

537 used bactericide found in human milk and in the aquatic environment in Sweden.

538 *Chemosphere* **2002**, *46*, 1485-1489.

539 5. Kierkegaard, A.; Björklund, J.; Fridén, U. Identification of the flame retardant

540 decabromodiphenyl ethane in the environment. *Environ. Sci. Technol.* **2004**, *38*,

541 3247-3253.

542 6. Hoh, E.; Zhu, L.; Hites, R. A. Dechlorane Plus, a chlorinated flame retardant, in the

543 Great Lakes. *Environ. Sci. Technol.* **2006**, *40*, 1184 -1189.

544   7. Hoh, E.; Zhu, L.; Hites, R. A. Novel flame retardants, 1,2-bis(2,4,6-

545   tribromophenoxy)ethane and 2,3,4,5,6-pentabromoethylbenzene, in United

546   States' environmental samples. *Environ. Sci. Technol.* **2005**, *39*, 2472 -2477.

547   8. Hernández, F.; Portolés, T.; Pitarch, E.; López, F. J. Target and nontarget screening

548   of organic micropollutants in water by solid-phase microextraction combined with

549   gas chromatography/high-resolution time-of-flight mass spectrometry. *Anal.*

550   *Chem.* **2007**, *79*, 9494-9504.

551   9. Hoh, E.; Dodder, N. G.; Lehotay, S. J.; Pangallo, K. C.; Reddy, C. M.; Maruya, K. A.

552   Nontargeted comprehensive two-dimensional gas chromatography/time-of-flight

553   mass spectrometry method and software for inventorying persistent and

554   bioaccumulative contaminants in marine environments. *Environ. Sci. Technol.*

555   **2012**, *46*, 8001-8008.

556   10. Krauss, M.; Singer, H.; Hollender, J. Lc–high resolution ms in environmental

557   analysis: From target screening to the identification of unknowns. *Anal. Bioanal.*

558   *Chem.* **2010,** *397*, 943-951.

559   11. Brack, W. Effect-directed analysis: a promising tool for the identification of

560   organic toxicants in complex mixtures? *Anal. Bioanal. Chem.* **2003**, *377*, 397-407.

561   12. Schymanski, E. L.; Bataineh, M.; Goss, K.-U.; Brack, W. Integrated analytical and

562   computer tools for structure elucidation in effect-directed analysis. *TrAC-Trend.*

563   *Anal. Chem.* **2009,** *28*, 550-561.

564   13. Rodan, B. D.; Pennington, D. W.; Eckley, N.; Boethling, R. S. Screening for

565   persistent organic pollutants: Techniques to provide a scientific basis for POPs

566   criteria in international negotiations. *Environ. Sci. Technol.* **1999**, *33*, 3482-3488.

567   14. Muir, D. C. G.; Howard, P. H. Are there other persistent organic pollutants? A

568   challenge for environmental chemists. *Environ. Sci. Technol.* **2006**, *40*, 7157–7166.

15. Walker, J. D.; Carslon, L. QSARS for identifying and prioritizing substances with persistence and bioconcentration potential. *SAR QSAR Environ. Res.* **2002**, *13*, 713-725.

16. Wiandt, S.; Poremski, H. J. Selection and prioritisation procedure of hazardous substances for the marine environment within SPAR/DYNAMEC. *Ecotoxicology* **2002**, *11*, 393-399.

17. Öberg, T. Virtual screening for environmental pollutants: Structure-activity relationships applied to a database of industrial chemicals. *Environ. Toxicol. Chem.* **2006**, *25*, 1178-1183.

18. Czub, G.; Wania, F.; McLachlan, M. S. Combining long range transport and bioaccumulation considerations to identify potential Arctic contaminants. *Environ. Sci. Technol.* **2008**, *42*, 3704–3709.

19. Brown, T. N.; Wania, F. Screening chemicals for the potential to be persistent organic pollutants: a case study of Arctic contaminants. *Environ. Sci. Technol.* **2008**, *42*, 5202-5209.

20. Breivik, K.; Arnot, J. A.; Brown, T. N.; McLachlan, M. S.; Wania, F. Screening organic chemicals in commerce for emissions in the context of environmental and human exposure. *J. Environ. Monit.* **2012**, *14*, 2028-2037.

21. Arnot, J. A.; Brown, T. N.; Wania, F.; Breivik, K.; McLachlan, M. S. Prioritizing chemicals and data requirements for screening-level exposure and risk assessment. Environ. Health Persp. **2012**, *120*, 1565-1570.

22. Arnot, J. A.; Mackay, D. Policies for chemical hazard and risk priority setting: can persistence, bioaccumulation, toxicity and quantity information be combined? *Environ. Sci. Technol.* **2008**, *42*, 648–4654.

593   23. Mackay, D.; Di Guardo, A.; Paterson, S.; Cowan, C. Evaluating the environmental

594   fate of a variety of types of chemicals using the EQC model. *Environ.Toxicol.*

595   *Chem.* **1996**, *15*, 1627-1637.

596   24. U.S. EPA (U.S. Environmental Protection Agency). 2011. Estimation Programs

597   Interface (EPI) Suite™ for Microsoft® Windows, ver. 4.1., Part Released October,

598   2011. Washington, DC:U.S. EPA.

599   25. Noll, W. Chemistry and Technology of Silicones. Academic Press: Orlando, 1968.

600   26. Kierkegaard, A.; McLachlan, M. S. Determination of linear and cyclic volatile

601   methylsiloxanes in air at a regional background site in Sweden. *Atmos. Environ.*

602   **2013** *80*, 322-329.

603   27. Buser, A. M.; Kierkegaard, A.; Bogdal, C.; MacLeod, M.; Scheringer, M.;

604   Hungerbühler, K. Concentrations in ambient air and emissions of cyclic volatile

605   methylsiloxanes in Zurich, Switzerland. *Environ. Sci. Technol.* **2013**, *47*, 7045-7051.

606   28. Kierkegaard, A.; Bignert, A.; McLachlan, M. S. Bioaccumulation of

607   decamethylcyclopentasiloxane in perch in Swedish lakes. *Chemosphere* **2013**, *93*,

608   789-793.

609   29. Borgå, K.; Fjeld, E.; Kierkegaard, A.; McLachlan, M. S. Consistency in trophic

610   magnification factors of cyclic methyl siloxanes in pelagic freshwater food webs

611   leading to brown trout. *Environ. Sci. Technol.* **2013**, *47*, 14394-14402.

612   30. Brooke, D. N.; Crookes, M. J.; Gray, D.; Robertson, S., Risk Assessment Report:

613   Decamethylcyclopentasilicone; Environment Agency of England and Wales:

614   Bristol, U.K., 2009. http://cdn.environment-agency.gov.uk/scho0309bpqx-e-e.pdf

615   (accessed Jan. 4, 2014).

616   31. Brooke, D. N.; Crookes, M. J.; Gray, D.; Robertson, S., Risk Assessment Report:

617   Octamethylcyclotetrasilicone; Environment Agency of England and Wales: Bristol,

618   U.K., 2009. http://cdn.environment-agency.gov.uk/scho0309bpqz-e-e.pdf

619   (accessed Jan. 4, 2014).

620   32. Brooke, D. N.; Crookes, M. J.; Gray, D.; Robertson, S., Risk Assessment Report:

621   Dodecamethylcyclohexasilicone; Environment Agency of England and Wales:

622   Bristol, U.K., 2009. http://cdn.environment-agency.gov.uk/scho0309bpqy-e-e.pdf

623   (accessed Jan. 4, 2014).

624   33. Graiver, D.; Farminer, K. W.; Narayan, R. A review of the fate and effects of

625   silicones in the environment. *J. Polym. Environ.* **2003**, *11*, 129-136.

626

627 **Tables**

628 Table 1: Levels of phenyl-tris(trimethylsiloxy)silane (M3TPh) in environmental samples

Air	ng m^{-3}	ng sample^{-1}
629		
630 Bromma WWTP	7.4	550
631 Zürich Kaserne 17 field blank		<0.2[a]
632 Zürich Kaserne 17 A	0.21	1.9
633 Zürich Kaserne 24 field blank		<0.2[a]
634 Zürich Kaserne 24 A	0.49	4.0
635 Stockholm 111201A[b,c]	0.069	4.1
636 Stockholm 111201B[b,c]	0.072	4.8
637 Stockholm 111124A[b,c]	0.046	3.2
638 Stockholm 111124B[b,c]	0.044	2.9
639		
Sewage sludge	ng g^{-1} dw	ng sample^{-1}
641 Procedural blank		0.21
642 Bromma 2013	470	1400
643 Henriksdal A 2013[c]	490	1100
644 Henriksdal B 2013[c]	530	1200
645		
Sediment, transect Stockholm archipelago	ng g^{-1} dw	ng sample^{-1}
647 Beckholmen A[c]	21	69
648 Beckholmen B[c]	18	81
649 Fisksätra	3.2	11
650 Gustavsberg	0.38	4.9
651 Baggensfjärden	0.46	3.9
652		
Sediment, Swedish lakes	ng g^{-1} dw	ng sample^{-1}
654 Östersjön	1.2	12
655 Hjulstafjärden	0.015	0.25
656		
Sediment, Norwegian lakes	ng g^{-1} dw	ng sample^{-1}
658 Mjøsa Gjøvik	1.7	13
659 Femunden	0.11	0.45

660

661 [a] No peak was observed in any of the blanks. This quantity was estimated based on 5
662 times the noise.

663 [b] The number represents the sampling date (year-month-day).

664 [c] A and B denote parallel samples (air) or parallel analysis of the same sample (sludge and
665 sediment).

666

667 Table 2: Levels of tris(trifluoropropyl)trimethyl-cyclotrisiloxane (D3F) and
668 tetrakis(trifluoropropyl)tetramethyl-cyclotetrasiloxane (D4F) in environmental samples

	D3F+ D3FOH[#]	D4F+ D4FOH[#]
	pg g^{-1} dw	pg g^{-1} dw
Sewage sludge		
Henriksdal 2013	130	600
Bromma 2013	70	130
Sediment, transect Stockholm archipelago		
Beckholmen	19	300
Fisksätra	70	80
Gustavsberg	10	30
Baggensfjärden	30	50
Sediment, Swedish lakes		
Östersjön	10	30
Runn	-	40
Hjulstafjärden	60	30
Sediment, Norwegian lakes		
Femunden	6	70
Mjøsa (Gjøvik)1600	1600	1800

690 [#] D3F was quantified as the sum of the areas of the D3F peak and the D3FOH peak in the
691 chromatogram. D4F was quantified in an analogous manner.
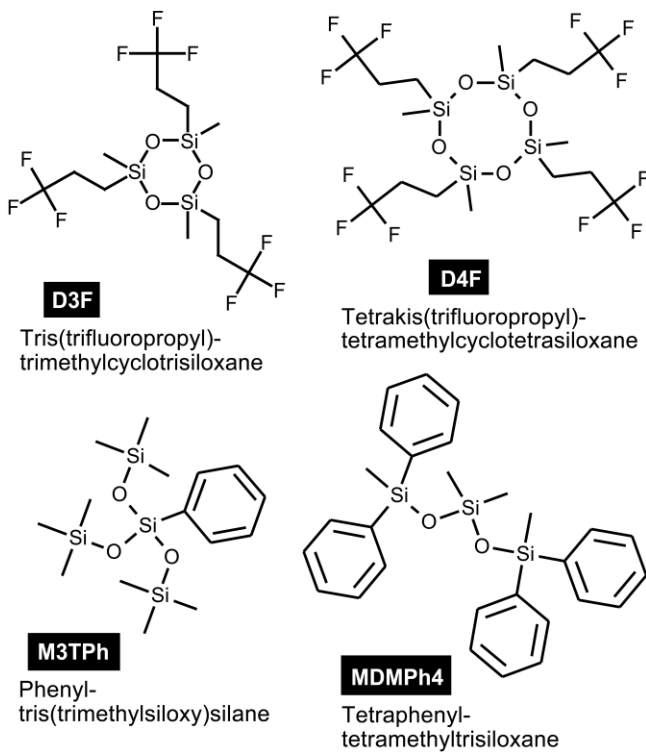
692

693 **TOC Art**
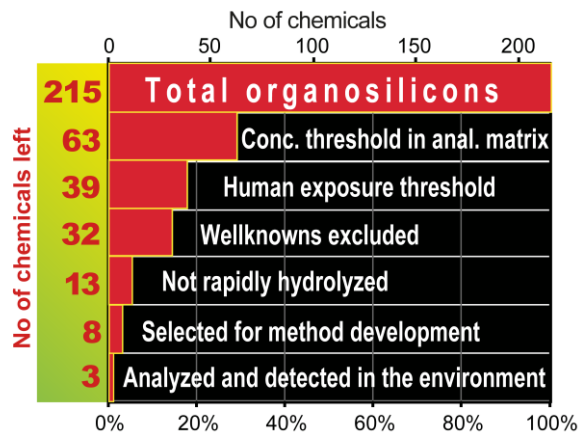
694

695 **Figures**

696

697

698 Figure 1: Substances identified for targeted analytical screening

699