Letter

Chlorines Are Not Evenly Substituted in Chlorinated Paraffins: A Predicted NMR Pattern Matching Framework for Isomeric Discrimination in Complex Contaminant Mixtures

Bo Yuan,* Daniel Henryk Lysak, Ronald Soong, Andrew Haddad, Arika Hisatsune, Arvin Moser, Sergey Golotvin, Dimitris Argyropoulos, Andre J. Simpson,* and Derek C. G. Muir*

Cite This: *Environ. Sci. Technol. Lett.* 2020, 7, 496−503

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Chlorinated paraffins (CPs) can be mixtures of nearly a half-million possible isomers. Despite the extensive use of CPs, their isomer composition and effects on the environment remain poorly understood. Here, we reveal the isomeric distributions of nine CP mixtures with single-chain lengths ($C_{14/15}$) and varying degrees of chlorination. The molar distribution of $C_nH_{2n+2-m}Cl_m$ in each mixture was determined using high-resolution mass spectrometry (MS). Next, the mixtures were analyzed by applying both one-dimensional ^1H, ^{13}C and two-dimensional nuclear magnetic resonance (NMR) spectroscopy. Due to substantially overlapping signals in the experimental NMR spectra, direct assignment of individual isomers was not possible. As such, a new NMR spectral matching approach that used massive NMR databases predicted by a neural network algorithm to provide the top 100 most likely structural matches was developed. The top 100 isomers appear to be an adequate representation of the overall mixture. Their modeled physicochemical and toxicity parameters agree with previous experimental results. Chlorines are not evenly distributed in any of the CP mixtures and show a general preference at the third carbon. The approach described here can play a key role in understanding of complex isomeric mixtures such as CPs that cannot be resolved by MS alone.
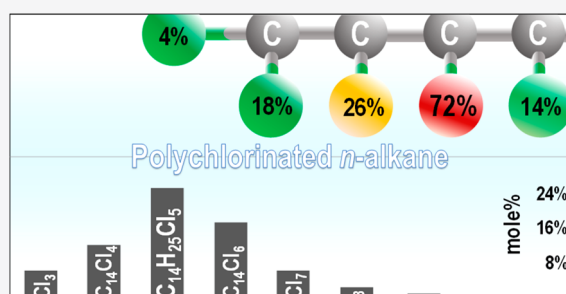
INTRODUCTION

The molecular structure of organic pollutants determines their physicochemical properties and toxicities, which are essential in investigating pollutant fate, transport, and environmental impacts.[1−3] Even small structural changes can greatly affect physicochemical properties. For instance, large variations in K_{OA} were observed between isomers of hexachlorocyclohexane,[4] resulting in different long-range transport pathways to the Arctic.[5] As another example, the metabolism of polybrominated diphenyl ethers differs among congeners with different bromine substitution patterns.[6] The specific molecular structure also determines the chemical activity. A recent study showed that halogen substitution neighboring the phenolic hydroxyl groups of bisphenol A analogues significantly enhanced their antimicrobial activity.[7] Despite the importance of the molecular structure of organic pollutants, extracting this information can be a great challenge for substances of unknown or variable composition, complex reaction products, and biological materials (UVCBs).[8] In particular, this challenge has existed for decades for the UVCBs known as chlorinated paraffins (CPs).[9]

CPs are high-production volume industrial products (>1 million metric tons/year) used as extreme pressure lubricant additives, flame retardants, and plasticizers.[10] They are usually synthesized under free radical chlorination with low positional selectivity.[9] Therefore, a typical CP product is an extremely complex mixture with potentially hundreds of thousands of possible isomers of polychlorinated *n*-alkanes (i.e., $C_nH_{2n+2-m}Cl_m$),[11] with manufacturers only specifying the carbon chain length range and degree of chlorination (% Cl, w/w). These specifications result in a rough classification of CP products: short-chain (SCCPs, C_{10-13}), medium-chain (MCCPs, C_{14-17}), and long-chain (LCCPs, $C_{\geq 18}$) CP mixtures.[12] Although all CPs have been identified as potentially hazardous substances since the early 1980s[13,14] and appear to be ubiquitous environmental contaminants,[15] SCCPs were only recently (2017)[16] identified and globally regulated as persistent organic pollutants (POPs), while MCCPs and LCCPs are still being used extensively.

The complexity of CPs is a serious obstacle for their analysis, research, and regulation.[9] Although column chromatography can resolve stereoisomers in a CP isomer standard,[17,18] no
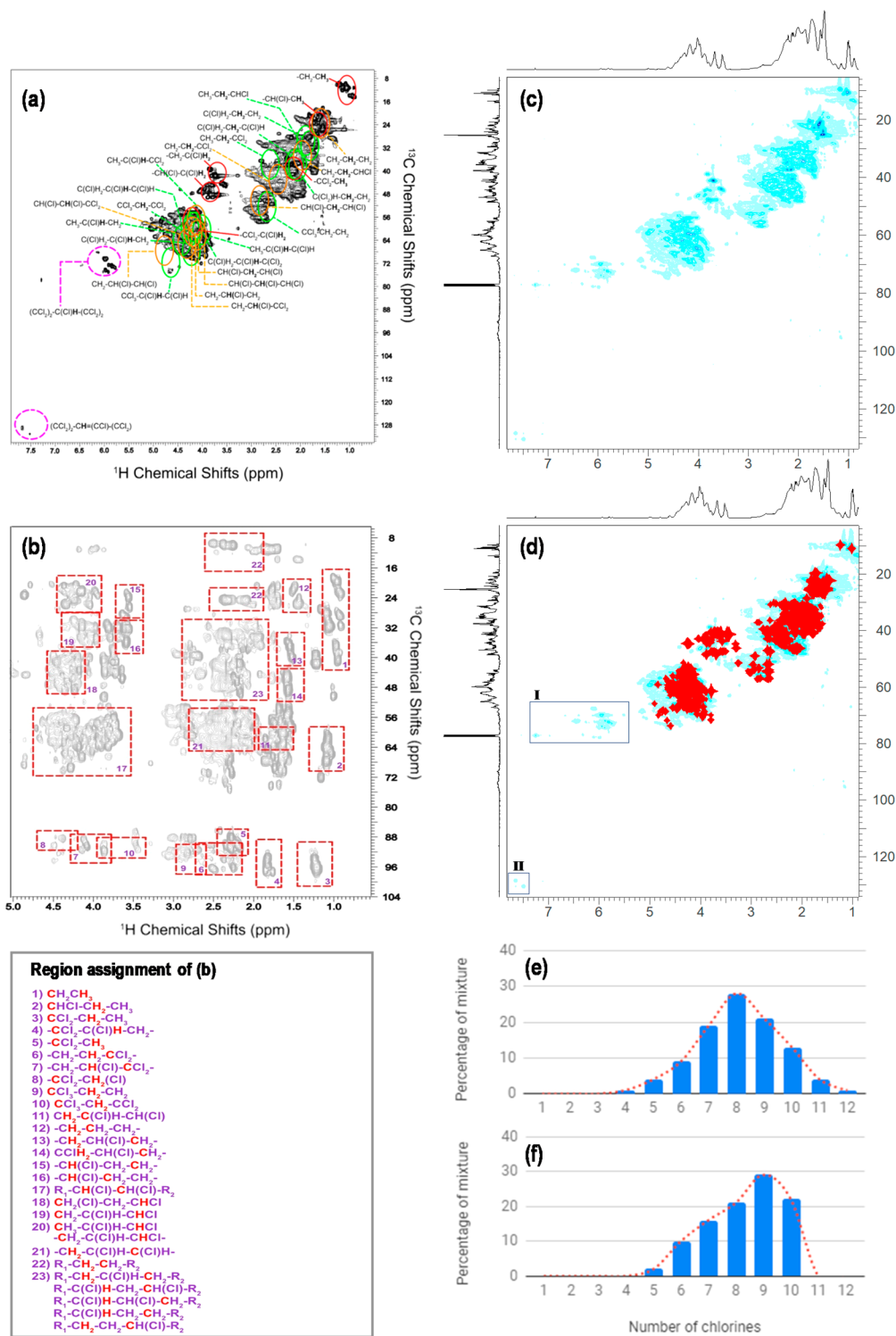
Figure 1. 2D NMR spectra and MS result of the C_{14} 60.14% Cl CP mixture. (a) $^1H-^{13}C$ HSQC spectrum and assignment of all regions based on both manual assignment and matches from predict databases. (b) $^{2,3}J$ HMBC $^1H-^{13}C$ spectrum and assignment of all regions. Comparison of (c) the HSQC spectrum and (d) the same HSQC spectrum with the top 100 isomer spectra overlaid. (e) Molar distribution of molecular formulas (i.e., congener groups) determined using MS. (f) Molar distribution of the top 100 isomers determined using NMR.

chromatographic technique has so far isolated individual isomers in a typical CP product.[19] In an optimal situation, CP isomers with a given number of carbons and chlorines can be resolved as a congener group (denoted as C_nCl_m) using high-resolution mass spectrometry (HRMS).[20] While mass spectrometry (MS) can provide valuable insights into the mass distribution of the mixture, the exact isomeric distribution can

be impossible to elucidate and other spectroscopic methods, such as nuclear magnetic resonance spectroscopy (NMR), are required.[21]

NMR is a powerful and non-invasive tool capable of providing detailed structural information about even the most complicated systems. The chemical shifts of ^1H and ^{13}C nuclei in one-dimensional (1D) NMR provide an indication of their chemical

environment and were used to identify the chlorine arrangement in substances such as pure CP isomers[22−24] but showed extensive overlap for the CP mixtures.[19,25] The peak capacity of a two-dimensional (2D) ^{1}H–^{13}C NMR spectrum is around 2 million, and the additional spectral dispersion versus 1D NMR helps reduce spectral overlap.[26,27] Sprengel et al. used 2D heteronuclear single-quantum coherence spectroscopy (HSQC) to assign nine substructures to single-chain length CPs synthesized by free radical chlorination using sulfuryl chloride.[19] However, the complexity and overlap of 2D NMR spectra made direct assignments of individual isomers impossible. Through the application of predicted 2D NMR to CPs, it is possible to profile the "spectral fingerprints" of different isomers and after databasing use the signatures for "best fit pattern matching" against the experimental 2D CP NMR. Such an approach was developed for identifying unknown organics in soil[28] but has yet be used to investigate complex isomeric environmental contaminant mixtures.

Herein, we significantly improve our knowledge of CP composition from the congener group level to actual isomeric discrimination. In this manner, multidimensional NMR was applied and vast 1D and 2D NMR databases were predicted using neural networks, which were then matched against the experimental data to provide an understanding of the most likely components present in CP mixtures. Multiple single-chain length CP mixtures were analyzed using both an established MS method[29] and the new NMR approach, resulting in identification of the isomers that are most likely to be present in each CP mixture. On the basis of the identified isomers, it became possible to predict the physicochemical properties and toxicities of the most likely CP mixture components.

■ MATERIALS AND METHODS

Chemicals. Single-chain length CP mixtures were used to reduce the complexity from a range of chain lengths.[19] Nine samples of CP mixtures were synthesized by Quimica del Cinca (Barcelona, Spain), a CP manufacturer, by chlorination of *n*-tetradecane and *n*-pentadecane. The mixtures were provided gratis by the Chlorinated Paraffin Industry Association (CPIA, Washington, DC). These consisted of five chlorinated *n*-tetradecanes (C_{14} CPs with degrees of chlorination of 39.75% Cl, 44.86% Cl, 50.07% Cl, 55.34% Cl, and 60.14% Cl) and four chlorinated *n*-pentadecanes (C_{15} CPs with degrees of chlorination of 40.30% Cl, 44.80% Cl, 50.00% Cl, and 54.75% Cl). For detailed synthesis of CPs and sample preparation, see the Supporting Information.

MS Spectrometry. Congener groups in the $C_{14/15}$ CP mixtures were measured using a direct injection dichloromethane-enhanced APCI-Q-Orbitrap MS[30] instrument (Q-Exactive, Thermal-Fisher Scientific) operating in negative full scan mode (m/z 300−2000) with an MS resolution of 120000 fwhm (full width at half-maximum). The molar compositions of congener groups in each CP mixture were quantified on the basis of an algorithm developed previously.[29] Each congener group was quantified on the basis of its own instrumental response factor, which thus reduces the level of response discrimination in CPs.[31] The quantification performance was evaluated by the CP mixture's degree of chlorination calculated on the basis of the congener group distribution.[32,33] For detailed instrumental settings and response factor calculation, see the Supporting Information and Figure S1.

NMR Spectroscopy. The NMR measurements consist of (i) 1D ^{1}H NMR, (ii) 1D ^{13}C NMR, (iii) 2D ^{13}C–^{1}H HSQC, and (iv) 2D ^{1}H–^{13}C heteronuclear multiple-bond coherence (HMBC). All NMR measurements were performed on a Bruker Advance III HD spectrometer operating at a ^{1}H frequency of 500.30 MHz fitted with a TCI Prodigy Cryoprobe. The assignment of the experimental NMR spectra was performed on the basis of matches against the predicted databases (1D ^{13}C, HSQC) and then cross referenced for consistency using experimental HMBC J couplings.[34] For detailed NMR settings and assignment information, see the Supporting Information.

Prediction and Matching of ^{13}C NMR Spectra. All possible CP isomers in the form of SMILES codes for the congener groups identified as major components by MS [$C_{14}Cl_{1−10}$ and $C_{15}Cl_{1−7}$ (Figure 1e and Figure S3)] were calculated. After identical structures had been removed (i.e., mirror images that were not stereoisomers), 410000 unique chemical structures remained. The ^{13}C NMR spectra were then predicted for all possible isomers using a neural network algorithm. These were arranged into one database containing the C_{14} CP isomers and one database containing the C_{15} CP isomers. Next, each ^{13}C NMR spectrum of the chlorinated mixtures was quantitatively compared against the predicted databases using the ACD/Labs similarity search algorithm (Figure S2a). The top 1000 matches, as ranked by hit quality index (HQI, eq S1), were retained. For validation, experimental and predicted data (^{13}C and ^{1}H) were compared for three polychlorinated alkane isomers: 1,2,3-trichloropropane, 1,1,1,2,2,3,3-heptachloropropane, and pentaerythrityl tetrachloride (Figure S3). Further details are provided in the Supporting Information.

Prediction and Matching of ^{1}H–^{13}C HSQC Spectra. CP structures were further filtered on the basis of the 2D NMR HSQC spectra. The HSQC spectra were predicted for the top 1000 isomers using a neural network algorithm. For matching, each of the 1000 isomers was compared against the CP mixtures (Figure S2b), yielding the top 100 isomers based on HQI (eq S2). The 100 isomers of each $C_{14/15}$ mixture were grouped on the basis of the molecular formulas to form the molar distribution in the mixture. The formula distributions identified independently by NMR and MS were compared on the basis of the coefficient of determination (R^2), with an R^2 of 1 indicating a perfect match and an R^2 of 0.5 a threshold of match.[35] Finally, physicochemical and toxicity parameters of the top 100 isomers were predicted using the Percepta Platform (ACD/Labs). Further details are provided in the Supporting Information.

■ RESULTS AND DISCUSSION

Molar Composition of Congener Groups. The MS results show that molar abundances of congener groups in each C_{14} or C_{15} CP mixture (Figure 1e and Figure S4) are the highest for congener groups with degrees of chlorination close to that of the mixture. For example, $C_{14}Cl_4$ (42.18% Cl) and $C_{14}Cl_3$ (35.25% Cl) are the most abundant congener groups in the C_{14} 39.75% Cl mixture, and $C_{15}Cl_6$ (50.75% Cl) is the most abundant congener group in the C_{15} 50.00% Cl mixture. This trend of congener group distribution in a CP mixture is consistent with previous studies.[29,33] The calculated degrees of chlorination of the $C_{14/15}$ CP mixtures agree with the manufacturer-reported values ($r^2 > 0.99$) with the absolute deviation ranging from −0.22% Cl to 0.06% Cl (Table S1). In addition, $C_{14}Cl_{15}$ was found in the C_{14} 60.14% Cl mixture. This confirms the presence of multi-chloro-substituted carbons in technical mixtures of CPs that have been reported in previous studies.[19,25,29]

C14Cl(40) (C1b = 0%, C14b = 0%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a
a	19%	16%	38%	8%	16%	20%	26%	26%	20%	16%	8%	38%	16%	19%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14c
b	0%	4%	0%	0%	1%	4%	4%	4%	4%	1%	0%	0%	4%	0%

C14Cl(45) (C1b = 6%, C14b = 6%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a
a	18%	21%	61%	17%	29%	40%	27%	27%	40%	29%	17%	61%	21%	18%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14c
b	6%	8%	3%	0%	4%	7%	11%	11%	7%	4%	0%	3%	8%	6%

C14Cl(50) (C1b = 4%, C14b = 4%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a
a	18%	26%	72%	14%	24%	45%	35%	35%	45%	24%	14%	72%	26%	18%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14c
b	4%	10%	4%	0%	1%	4%	2%	2%	4%	1%	0%	4%	10%	4%

C14Cl(55) (C1b = 4%, C14b = 4%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a
a	27%	45%	57%	23%	26%	26%	50%	50%	26%	26%	23%	57%	45%	27%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14c
b	4%	13%	5%	0%	4%	3%	8%	8%	3%	4%	0%	5%	13%	4%

C14Cl(60) (C1b = 10%, C14b = 10%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a
a	78%	55%	65%	33%	40%	45%	47%	47%	45%	40%	33%	65%	55%	78%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14c
b	10%	11%	3%	2%	5%	9%	9%	9%	9%	5%	2%	3%	11%	10%

C15Cl(40) (C1b = 0%, C15b = 0%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a	C15a
a	12%	23%	26%	4%	17%	27%	18%	11%	18%	27%	17%	4%	26%	23%	12%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14b	C15c
b	0%	8%	0%	0%	5%	11%	6%	0%	6%	11%	5%	0%	0%	8%	0%

C15Cl(45) (C1b = 0%, C15b = 0%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a	C15a
a	15%	17%	62%	8%	15%	25%	18%	44%	18%	25%	15%	8%	62%	17%	15%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14b	C15c
b	0%	6%	0%	0%	1%	2%	6%	11%	6%	2%	1%	0%	0%	6%	0%

C15Cl(50) (C1b = 0%, C15b = 0%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a	C15a
a	23%	25%	62%	10%	38%	36%	21%	43%	21%	36%	38%	10%	62%	25%	23%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14b	C15c
b	0%	1%	2%	1%	1%	1%	2%	1%	2%	1%	1%	1%	2%	1%	0%

C15Cl(55) (C1b = 3%, C15b = 3%)

	C1a	C2a	C3a	C4a	C5a	C6a	C7a	C8a	C9a	C10a	C11a	C12a	C13a	C14a	C15a
a	36%	50%	43%	26%	24%	31%	26%	32%	26%	31%	24%	26%	43%	50%	36%
	C1c	C2b	C3b	C4b	C5b	C6b	C7b	C8b	C9b	C10b	C11b	C12b	C13b	C14b	C15c
b	0%	8%	0%	0%	0%	1%	3%	6%	3%	1%	0%	0%	0%	8%	0%

Figure 2. Heat maps showing chlorine substitution at each carbon as a percentage for the top 100 CP isomers in the database of nine samples. The name of a CP mixture, e.g., C_{14} 39.75% Cl CPs, is shown as $C_{14}Cl(40)$. Note that stereoisomers are not considered. Therefore, both the mirrored carbons [e.g., C(1) and C(14) in a C_{14} CP mixture] and hydrogens a/b(/c) on a carbon atom are equivalent positions. The results are expressed as simple percentages, for example, for a specific position of the 10% displayed it means that in 10% of the top 100 isomers this position is substituted by chlorine.

Fragments of Functional Groups in CP Mixtures. The ^1H NMR spectra of C_{14} or C_{15} CP mixtures downfield shift with an increasing degree of chlorination due to deshielding from the chlorines (Figure S5). Such a trend was also reported by Sprengel et al.[19] However, the ^1H NMR spectra provide little detailed molecular information due to a lack of dispersion. In contrast, the ^{13}C NMR spectra provide better dispersion, allowing identification of discrete regions where major CP fragments occur, along with comparisons of their relative quantities (Table S2). With increasing degrees of chlorination, the quantities of [-CH$_2$-] and methyl groups in the CP mixtures decrease, while the quantities of chlorine-substituted groups increase. ^{13}C NMR also helped further confirm the presence of dichloro-substituted carbons, and a small quantity of trichloromethyl terminations[36] and unsaturated compounds[37,38] was also identified. However, 1D ^{13}C NMR still shows considerable overlap of fragment regions.

The 2D HSQC spectra improve the assignment by increasing spectral dispersion and providing one-bond H−C correlations.[19] Though spectral overlap is reduced in a 2D plot, due to the complex nature of the CPs, it is still highly complicated, and with assistance via pattern matching from the predicted databases (see below), the overlap from different functional moieties is extensive (Figure 1a). To support the HSQC

assignments, HMBC was used for the analysis of CPs. Upon inclusion of all of the $^{2,3}J$ $^{1}H-^{13}C$ correlations (i.e., carbons within two or three bonds of the selected proton resonance), the assignment of major fragment regions is less ambiguous (Figure 1b). While spectral assignment can provide an overview of the major structural components that are present, to gain information about individual isomers and their distributions, spectral matching of vast predicted NMR databases is required.

Prediction and Matching of NMR Data. The total number of possible isomers for C_{14} and C_{15} CPs is more than 410000, showcasing the extreme complexity of CP mixtures. This number is ~25 times larger than the number of 16271[11] calculated on the basis of ref 39. This is because the latter assumed no more than one chlorine atom bound to any carbon atom, while our calculation allowed for carbon atoms to be substituted by up to two chlorines. The number of possible isomers increases by geometric progression with an increase in chain length[39] and the number of substituted chlorines (Figure S6). Therefore, we limited the list of isomers within the ranges of $C_{14}Cl_{1-10}$ and $C_{15}Cl_{1-7}$.

It was not possible to generate $^{1}H-^{13}C$ HSQC spectra for all of the isomers, due to the computational expense of HSQC prediction. Instead, the ^{13}C NMR spectra of all of the 410000 isomers were predicted, and the top 1000 isomers matching the experimental ^{13}C NMR spectra were identified for each CP mixture (shown in Worksheets 10 and 11). Then $^{1}H-^{13}C$ HSQC spectra were predicted for the top 1000 isomers and compared with the spectra of all of the $C_{14/15}$ CPs, resulting in the top 100 isomers (Worksheets 1–9).

2D HSQC experiments tend not to be fully quantitative.[40] The peak search algorithm matches on the basis of patterns[28] in the 2D data rather than intensity. However, as the peak picking used is done both computationally and then by eye, both approaches will always bias the most intense and well-defined peaks in the data. As such, the results from the 2D mixture searches will favor the most abundant components that best match the chemical shift patterns in the experimental data.

Top 100 Isomers. The overlaid spectrum (an example is shown in Figure 1d) of the top 100 isomers (Worksheet 5) adequately represents the dominant signals (i.e., major components) in the experimental NMR spectrum of the C_{14} 60.14% Cl mixture (Figure 1c). Region I in the HSQC spectrum (Figure 1d) is not represented by the top 100 isomers, which arises from methine protons between two perchlorinated neighbors. While these compounds are included in the predicted databases, their very low abundance leads to negative discrimination by the ^{13}C matching algorithm (see the Supporting Information). It is also noteworthy that the simulations exclude unsaturated isomers [that were at very low concentrations of <1% (Table S2)], which explains the lack of signal around 6 ppm (^{1}H)/70 ppm (^{13}C), where unsaturated compounds resonate (region II in Figure 1d).

The top 100 isomers of all of the $C_{14/15}$ CP mixtures are generally comparable [median R^2 = 0.83; range of 0.30–0.93 (Table S1)] with the molecular formulas identified independently by MS (Figure 1e,f and Figure S4); they are thus an adequate representation of the overall mixture, although biased toward the most abundant components in the mixtures. The low R^2 (0.37) for the C_{14} 55.34% Cl mixture might due to this bias. Moreover, $C_{14}Cl_{>10}$ and $C_{15}Cl_{>7}$ were included in the MS results but not in the NMR results. This results in the least comparable congener group distribution in this for C_{15} 54.75% Cl (R^2 = 0.30) as the MS data show that $C_{15}Cl_{>7}$ mixtures contribute in an

amount of 33%. The degrees of chlorination calculated on the basis of the top 100 isomers show larger deviations (−6.99% Cl to 4.17% Cl) than the MS results but are comparable with the deviations in another study[33] that quantified congener groups.

Positional Selectivity of Chlorine Substitution. The top 100 isomers of each CP mixture were depicted as a substitution heat map (Figure 2) of the most likely position of chlorine on the alkane. The maps were generated by summing the number of times a chlorine occurred at a specific position across the top 100 isomers for each mixture (for details, see the Supporting Information). The heat maps clearly show that chlorine substitution at each carbon is not statistically evenly distributed in any CP mixture. With C_{14} 39.75% Cl (equivalent to $C_{14}H_{26.38}Cl_{3.62}$) as an example, if chlorines were evenly distributed, all 30 hydrogens of the tetradecane would have an identical substitution percentage of ~12%. Remarkably, the third carbon from an end of the chain, denoted as C(3), is the most reactive to chlorine in mixtures with low and intermediate degrees of chlorination (C_{14} 39.75–55.34% Cl and C_{15} 40.30–50.00% Cl). This may be explained by steric effects and the low bond energy at C(3). Gerson and Nyburg measured C−C bond lengths along n-tetracosane and n-hexacosane using X-ray crystallography; the C_{2-3} bond was the longest, while the C_{3-4} bond was the second longest.[41] These two C−C bonds also showed the highest probability of breaking in the fragmentation experiment of n-tetradecane and n-nonane under electron impact.[42] These studies suggest that C(3) is the most reactive position along the n-alkane under free radical chlorination. Meanwhile, C(4) in all CP mixtures (Figure 2) is least reactive to chlorine. Adjacent to a "hot spot" of chlorination [i.e., C(3)], C(4) may be influenced by both steric hindrance[43] and the electron cloud of the chlorine on C(3). Despite being adjacent to C(3), C(2) has a weaker steric effect from the methyl group [i.e., C(1)] and thus shows a higher substitution percentage than C(4) in all of the CP mixtures. The steric hindrance from C(3) is lower for carbons farther from C(3) [e.g., C(5)], facilitating chlorination of these carbons.

Among different CP mixtures, the maximum substitution percentage at the central carbon(s) [i.e., C(7) and C(8) of C_{14} CPs and C(8) of C_{15} CPs] was found in CP mixtures with intermediate degrees of chlorination. Intermediately chlorinated CP isomers may tend to reach the lowest-energy fold,[44,45] making the central carbon(s) experience the least steric hindrance. Chlorination is less preferable for primary carbons [i.e., C(1)] than secondary carbons,[43] which can be seen in CP mixtures with the lowest level of chlorination (Figure 2). However, with an increase in the degree of chlorination, the chlorine substitution percentage grows at C(1), with the percentage at C(1) being the highest in highly chlorinated CP mixtures (C_{14} 60.14% Cl). This may be because hydrogens on most secondary carbons are substituted with chlorine, which causes stronger steric hindrance compared with the primary carbon. Dichloro substitution is most likely at C(2) and the central carbon(s). The methyl group on C(2) has less steric hindrance to substitution of the second chlorine compared with the ethyl group on C(3).

Environmental Implications. The structures of CP isomers are crucial in predicting and elucidating physicochemical properties,[45,46] environmental behaviors,[47,48] and decomposition mechanisms[49−51] of complex CP mixtures comprised of these isomers. In the work presented here, the predicted physicochemical and toxicity parameters of the top 100 isomers are summarized in the Supporting Information for each CP

mixture. To compare our predictions with experimental data, we predicted parameters of each CP mixture (Table S3) on the basis of their isomeric composition, allowing comparison to experimentally available CP mixture data. The predicted LD_{50} parameters agree with previous findings that mammalian toxicity appears to decrease with chain length and increase with the degree of chlorination.[52] However, caution should be used in interpreting these predicted data as they are based on only C_{14} and C_{15} CPs, while the chain length range of CPs can be from C_{10} to C_{30} and even wider. Meanwhile, CPs with 50−55% Cl showed the minimum experimental log K_{OW} value among mixtures of the same chain length,[46] which is consistent with our prediction (Table S3). Quantitatively, the experimental log K_{OW} of C_{15} 50.4% Cl was 6.65,[46] while our prediction was 6.77 (C_{15} 50.00% Cl), which is 1 order of magnitude closer than a prediction of 7.78 based on assumed isomer compositions[45] (Table S4).

C_{14} and C_{15} CPs are major components of the commercial MCCP products.[53] They have frequently been found to have the dominant chain lengths of CPs in wildlife,[54−56] human,[11,57,58] sediment,[59,60] sewage sludge,[61] and indoor dust samples[62,63] worldwide. Our work addresses the urgent need for substantial knowledge of these mixtures and can extend to other chain lengths if individual mixtures are available.[64] However, the geometric growth of the number of possible isomers with an increasing chain length and chlorines may pose a demanding challenge for NMR prediction and databasing. This requires more computing resources and/or techniques working in tandem, e.g., separation and concentration of CPs prior to the NMR analysis. More importantly, our work demonstrates the usefulness of a predicted and databased NMR matching framework for a better understanding of complex environmental isomeric mixtures. These techniques may not only play a key role in elucidating individual isomers in CPs but also potentially be extended to any complex organic mixtures.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.estlett.0c00244.

Top 100 isomers (Excel Worksheets 1−9) and top 1000 isomers (Excel Worksheets 10 and 11) (XLSX)

Synthesis of CPs; MS experiment; NMR experiment; NMR data prediction, database creation, and spectral matching; comparison of NMR-MS results (Table S1); fragment quantities (Table S2); modeled physicochemical properties and toxicity parameters (Table S3); calculation of log K_{OW} (Table S4); C_nCl_m RF calculation (Figure S1); examples of spectral matching (Figure S2); NMR prediction validation (Figure S3); molar distributions (Figure S4); ^1H NMR spectra (Figure S5); and numbers of possible isomers (Figure S6) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Derek C. G. Muir − *Canada Centre for Inland Waters, Environment and Climate Change Canada, Burlington, ON, Canada L7S 1A1;* ● orcid.org/0000-0001-6631-9776; Phone: 1-905-319-6921; Email: derek.muir@canada.ca; Fax: 1-905-336-4699

Andre J. Simpson − *Environmental NMR Centre, University of Toronto, Toronto, ON, Canada M1C 1A4;* ● orcid.org/0000-0002-8247-5450; Email: andre.simpson@utoronto.ca

Bo Yuan − *Department of Environmental Science, Stockholm University, SE-10691 Stockholm, Sweden;* ● orcid.org/0000-0002-2043-8128; Email: bo.yuan@aces.su.se

Authors

Daniel Henryk Lysak − *Environmental NMR Centre, University of Toronto, Toronto, ON, Canada M1C 1A4*

Ronald Soong − *Environmental NMR Centre, University of Toronto, Toronto, ON, Canada M1C 1A4;* ● orcid.org/0000-0002-8223-9028

Andrew Haddad − *Environmental NMR Centre, University of Toronto, Toronto, ON, Canada M1C 1A4*

Arika Hisatsune − *Environmental NMR Centre, University of Toronto, Toronto, ON, Canada M1C 1A4*

Arvin Moser − *ACD/Labs, Toronto, ON, Canada M5C 1B5*

Sergey Golotvin − *ACD/Labs, Toronto, ON, Canada M5C 1B5*

Dimitris Argyropoulos − *ACD/Labs, Toronto, ON, Canada M5C 1B5*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.estlett.0c00244

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Selassie, C.; Verma, R. P. History of Quantitative Structure−Activity Relationships. In *Burger's Medicinal Chemistry and Drug Discovery*; Abraham, D. J., Ed.; Wiley: Hoboken, NJ, 2010.

(2) Willett, K. L.; Ulrich, E. M.; Hites, R. A. Differential Toxicity and Environmental Fates of Hexachlorocyclohexane Isomers. *Environ. Sci. Technol.* **1998**, *32*, 2197−2207.

(3) Sadik, O. A.; Witt, D. M. Peer reviewed: Monitoring endocrine-disrupting chemicals. *Environ. Sci. Technol.* **1999**, *33*, 368A−374A.

(4) Shoeib, M.; Harner, T. Using measured octanol-air partition coefficients to explain environmental partitioning of organochlorine pesticides. *Environ. Toxicol. Chem.* **2002**, *21*, 984−990.

(5) Li, Y. F.; Macdonald, R. W.; Jantunen, L. M. M.; Harner, T.; Bidleman, T. F.; Strachan, W. M. J. The transport of β-hexachlorocyclohexane to the western Arctic Ocean: a contrast to α-HCH. *Sci. Total Environ.* **2002**, *291*, 229−246.

(6) Stapleton, H. M.; Letcher, R. J.; Li, J.; Baker, J. E. Dietary accumulation and metabolism of polybrominated diphenyl ethers by

juvenile carp (Cyprinus carpio). *Environ. Toxicol. Chem.* **2004**, *23*, 1939−1946.

(7) Ji, F.; Wang, C.; Wang, H.; Liu, G.; Chen, B.; Hu, L.; Jiang, G.; Song, M.; Liang, Y. Tetrabromobisphenol A (TBBPA) exhibits specific antimicrobial activity against Gram-positive bacteria without detectable resistance. *Chem. Commun.* **2017**, *53*, 3512−3515.

(8) Scialli, A. R.; Guikema, A. J. REACH and reproductive and developmental toxicology: still questions. *Syst. Biol. Reprod. Med.* **2012**, *58*, 63−69.

(9) Tomy, G. T. Analysis of Chlorinated Paraffins in Environmental Matrices: The Ultimate Challenge for the Analytical Chemist. In *The handbook of environmental chemistry 10: Chlorinated paraffins*; de Boer, J., Ed.; Springer-Verlag: Berlin, 2010.

(10) De Boer, J.; El-Sayed Ali, T.; Fiedler, H.; Legler, J.; Muir, D. C.; Nikiforov, V. A.; Tomy, G. T.; Tsunemi, K. In *The handbook of environmental chemistry 10: Chlorinated paraffins*; de Boer, J., Ed.; Springer-Verlag: Berlin, 2010.

(11) Zhou, Y.; Yuan, B.; Nyberg, E.; Yin, G.; Bignert, A.; Glynn, A.; Odland, J. Ø.; Qiu, Y.; Sun, Y.; Wu, Y.; Xiao, Q.; Yin, D.; Zhu, Z.; Zhao, J.; Bergman, Å. Chlorinated Paraffins in Human Milk from Urban Sites in China, Sweden, and Norway. *Environ. Sci. Technol.* **2020**, *54*, 4356−4366.

(12) Gluge, J.; Schinkel, L.; Hungerbuhler, K.; Cariou, R.; Bogdal, C. Environmental Risks of Medium-Chain Chlorinated Paraffins (MCCPs): A Review. *Environ. Sci. Technol.* **2018**, *52*, 6743−6760.

(13) Campbell, I.; McConnell, G. Chlorinated Paraffins and the Environment 0.1. Environmental Occurrence. *Environ. Sci. Technol.* **1980**, *14*, 1209−1214.

(14) Jansson, B.; Andersson, R.; Asplund, L.; Litzen, K.; Nylund, K.; Sellströom, U.; Uvemo, U.-B.; Wahlberg, C.; Wideqvist, U.; Odsjö, T.; Olsson, M. Chlorinated and brominated persistent organic compounds in biological samples from the environment. *Environ. Toxicol. Chem.* **1993**, *12*, 1163−1174.

(15) van Mourik, L. M.; Gaus, C.; Leonards, P. E.; de Boer, J. Chlorinated paraffins in the environment: A review on their production, fate, levels and trends between 2010 and 2015. *Chemosphere* **2016**, *155*, 415−28.

(16) UNEP. UNEP/POPS/COP.8/SC-8/11. Listing of short-chain chlorinated paraffins. 2017. http://chm.pops.int/Portals/0/download. aspx?d=UNEP-POPS-COP.8-SC-8-11.English.pdf (accessed 2020-04-29).

(17) Frenzen, G.; Sippel, H.; Coelhan, M. The relative configuration of a stereoisomer of 1, 2, 5, 6, 9, 10-hexachlorodecane. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **1999**, *55*, IUC9800079−IUC9800079.

(18) van Mourik, L. M.; Lava, R.; O'Brien, J.; Leonards, P. E. G.; de Boer, J.; Ricci, M. The underlying challenges that arise when analysing short-chain chlorinated paraffins in environmental matrices. *J. Chromatogr. A* **2020**, *1610*, 460550.

(19) Sprengel, J.; Wiedmaier-Czerny, N.; Vetter, W. Characterization of single chain length chlorinated paraffin mixtures with nuclear magnetic resonance spectroscopy (NMR). *Chemosphere* **2019**, *228*, 762−768.

(20) Yuan, B.; Muir, D.; MacLeod, M. Methods for trace analysis of short-, medium-, and long-chain chlorinated paraffins: Critical review and recommendations. *Anal. Chim. Acta* **2019**, *1074*, 16−32.

(21) Fleming, I.; Williams, D. H. *Spectroscopic methods in organic chemistry*; McGraw-Hill: London, 1966.

(22) Vasil'eva, T. T.; Kruglova, N. V.; Dostovalova, V. I.; Freidlina, R. K. Rearrangement of polychloroalkyl radicals with 1,4-, 1,5- and 1,6-migration of hydrogen atom. *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **1978**, *27*, 1629−1634.

(23) Bessière, J. M.; Boutevin, B.; Taha, M.; Vial, F. 1H and 13C NMR study of vinyl chloride telomers with carbon tetrachloride. *Org. Magn. Reson.* **1984**, *22*, 792−793.

(24) Benedikt, G. M. NMR Spectroscopy of Poly(Vinyl Chloride) Defects. 1H-NMR Analysis of the 1,2-Dichloroethyl End Group to the Triad Level. *J. Macromol. Sci., Part A: Pure Appl.Chem.* **1992**, *29*, 85−98.

(25) Hüttig, J. *Determination of the "new" problem group chloroparaffins in sediments by HRGC-LRMS*; University of Basel, 2006.

(26) Hertkorn, N.; Ruecker, C.; Meringer, M.; Gugisch, R.; Frommberger, M.; Perdue, E. M.; Witt, M.; Schmitt-Kopplin, P. High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. *Anal. Bioanal. Chem.* **2007**, *389*, 1311−1327.

(27) Kelleher, B. P.; Simpson, A. J. Humic Substances in Soils: Are They Really Chemically Distinct? *Environ. Sci. Technol.* **2006**, *40*, 4605−4611.

(28) Simpson, A. J.; Lefebvre, B.; Moser, A.; Williams, A.; Larin, N.; Kvasha, M.; Kingery, W. L.; Kelleher, B. Identifying residues in natural organic matter through spectral prediction and pattern matching of 2D NMR datasets. *Magn. Reson. Chem.* **2004**, *42*, 14−22.

(29) Yuan, B.; Bogdal, C.; Berger, U.; MacLeod, M.; Gebbink, W. A.; Alsberg, T.; de Wit, C. A. Quantifying Short-Chain Chlorinated Paraffin Congener Groups. *Environ. Sci. Technol.* **2017**, *51*, 10633−10641.

(30) Yuan, B.; Tay, J. H.; Papadopoulou, E.; Haug, L. S.; Padilla-Sánchez, J. A.; de Wit, C. A. Complex Mixtures of Chlorinated Paraffins Found in Hand Wipes of a Norwegian Cohort. *Environ. Sci. Technol. Lett.* **2020**, *7*, 198−205.

(31) Krätschmer, K.; Schächtele, A. Interlaboratory studies on chlorinated paraffins: Evaluation of different methods for food matrices. *Chemosphere* **2019**, *234*, 252−259.

(32) Zencak, Z.; Reth, M.; Oehme, M. Dichloromethane-enhanced negative ion chemical ionization for the determination of polychlorinated n-alkanes. *Anal. Chem.* **2003**, *75*, 2487−92.

(33) Gao, Y.; Zhang, H.; Zou, L.; Wu, P.; Yu, Z.; Lu, X.; Chen, J. Quantification of Short-Chain Chlorinated Paraffins by Deuterodechlorination Combined with Gas Chromatography−Mass Spectrometry. *Environ. Sci. Technol.* **2016**, *50*, 3746−3753.

(34) Soong, R.; Pautler, B. G.; Moser, A.; Jenne, A.; Lysak, D. H.; Adamo, A.; Simpson, A. J. CASE (Computer-Assisted Structure Elucidation) Study for an Undergraduate Organic Chemistry Class. *J. Chem. Educ.* **2020**, *97*, 855−860.

(35) Bogdal, C.; Alsberg, T.; Diefenbacher, P. S.; MacLeod, M.; Berger, U. Fast quantification of chlorinated paraffins in environmental samples by direct injection high-resolution mass spectrometry with pattern deconvolution. *Anal. Chem.* **2015**, *87*, 2852−60.

(36) Yuan, B.; Alsberg, T.; Bogdal, C.; MacLeod, M.; Berger, U.; Gao, W.; Wang, Y.; de Wit, C. A. Deconvolution of Soft Ionization Mass Spectra of Chlorinated Paraffins To Resolve Congener Groups. *Anal. Chem.* **2016**, *88*, 8980−8.

(37) Schinkel, L.; Lehner, S.; Heeb, N. V.; Marchand, P.; Cariou, R.; McNeill, K.; Bogdal, C. Dealing with strong mass interferences of chlorinated paraffins and their transformation products: An analytical guide. *TrAC, Trends Anal. Chem.* **2018**, *106*, 116−124.

(38) Li, T.; Gao, S.; Ben, Y.; Zhang, H.; Kang, Q.; Wan, Y. Screening of Chlorinated Paraffins and Unsaturated Analogues in Commercial Mixtures: Confirmation of Their Occurrences in the Atmosphere. *Environ. Sci. Technol.* **2018**, *52*, 1862−1870.

(39) Tomy, G. T.; Stern, G. A. Analysis of C-14-C-17 polychloro-n-alkanes in environmental matrixes by accelerated solvent extraction-nigh-resolution gas chromatography/electron capture negative ion high-resolution mass spectrometry. *Anal. Chem.* **1999**, *71*, 4860−4865.

(40) Lane, D.; Skinner, T. E.; Gershenzon, N. I.; Bermel, W.; Soong, R.; Dutta Majumdar, R.; Liaghati Mobarhan, Y.; Schmidt, S.; Heumann, H.; Monette, M.; Simpson, M. J.; Simpson, A. J. Assessing the potential of quantitative 2D HSQC NMR in 13C enriched living organisms. *J. Biomol. NMR* **2019**, *73*, 31−42.

(41) Gerson, A. R.; Nyburg, S. Structures of two binary n-alkane solid solutions. *Acta Crystallogr., Sect. B: Struct. Sci.* **1994**, *50*, 252−256.

(42) Lavanchy, A.; Houriet, R.; Gäumann, T. The mass spectrometric fragmentation of n-alkanes. *Org. Mass Spectrom.* **1979**, *14*, 79−85.

(43) Leffler, J. E.; Grunwald, E. *Rates and equilibria of organic reactions: As treated by statistical, thermodynamic and extrathermodynamic methods*; Dover: New York, 2013.

(44) ChemTube3D. Polyvinyl Chloride Poly(chloroethene) PVC. https://www.chemtube3d.com/_pvcf/ (accessed 2020-02-13).

(45) Glüge, J.; Bogdal, C.; Scheringer, M.; Buser, A. M.; Hungerbuhler, K. Calculation of Physicochemical Properties for

Short- and Medium-Chain Chlorinated Paraffins. *J. Phys. Chem. Ref. Data* **2013**, *42*, 023103.

(46) Hilger, B.; Fromme, H.; Volkel, W.; Coelhan, M. Effects of Chain Length, Chlorination Degree, and Structure on the Octanol-Water Partition Coefficients of Polychlorinated n-Alkanes. *Environ. Sci. Technol.* **2011**, *45*, 2842−2849.

(47) Fisk, A. T.; Tomy, G. T.; Cymbalisty, C. D.; Muir, D. C. G. Dietary accumulation and quantitative structure-activity relationships for depuration and biotransformation of short (C10), medium (C14), and long (C18) carbon-chain polychlorinated alkanes by juvenile rainbow trout (Oncorhynchus mykiss). *Environ. Toxicol. Chem.* **2000**, *19*, 1508−1516.

(48) Diefenbacher, P. S.; Bogdal, C.; Gerecke, A. C.; Glüge, J.; Schmid, P.; Scheringer, M.; Hungerbühler, K. Short-Chain Chlorinated Paraffins in Zurich, Switzerland—Atmospheric Concentrations and Emissions. *Environ. Sci. Technol.* **2015**, *49*, 9778−9786.

(49) Xin, S.; Gao, W.; Wang, Y.; Jiang, G. Identification of the Released and Transformed Products during the Thermal Decomposition of a Highly Chlorinated Paraffin. *Environ. Sci. Technol.* **2018**, *52*, 10153−10162.

(50) Sun, Y.; Pan, W.; Lin, Y.; Fu, J.; Zhang, A. Chlorination pattern effect on thermodynamic parameters and environmental degradability for C10-SCCPs: Quantum chemical calculation based on virtual combinational library. *J. Environ. Sci.* **2016**, *39*, 184−197.

(51) Li, Y.; Hou, X.; Chen, W.; Liu, J.; Zhou, Q.; Schnoor, J. L.; Jiang, G. Carbon Chain Decomposition of Short Chain Chlorinated Paraffins Mediated by Pumpkin and Soybean Seedlings. *Environ. Sci. Technol.* **2019**, *53*, 6765−6772.

(52) Ali, T. E.; Legler, J. Overview of the Mammalian and Environmental Toxicity of Chlorinated Paraffins. In *The handbook of environmental chemistry 10: Chlorinated paraffins*; de Boer, J., Ed.; Springer-Verlag: Berlin, 2010.

(53) U.S. Environmental Protection Agency. TSCA new chemicals review program standard review risk assessment on medium-chain chlorinated paraffins (PMN P-14-0683, P-14-0684). 2015. https://www.epa.gov/sites/production/files/2015-12/documents/standard_review_risk_assessment_p-14-683-684_qualice_docket.pdf (accessed 2020-04-29).

(54) Du, X.; Yuan, B.; Zhou, Y.; Benskin, J. P.; Qiu, Y.; Yin, G.; Zhao, J. Short-, Medium-, and Long-Chain Chlorinated Paraffins in Wildlife from Paddy Fields in the Yangtze River Delta. *Environ. Sci. Technol.* **2018**, *52*, 1072−1080.

(55) Yuan, B.; Vorkamp, K.; Roos, A. M.; Faxneld, S.; Sonne, C.; Garbus, S. E.; Lind, Y.; Eulaers, I.; Hellström, P.; Dietz, R.; Persson, S.; Bossi, R.; de Wit, C. A. Accumulation of Short-, Medium-, and Long-Chain Chlorinated Paraffins in Marine and Terrestrial Animals from Scandinavia. *Environ. Sci. Technol.* **2019**, *53*, 3526−3537.

(56) Yuan, B.; Benskin, J. P.; Chen, C.-E. L.; Bergman, Å. Determination of Chlorinated Paraffins by Bromide-Anion Attachment Atmospheric-Pressure Chemical Ionization Mass Spectrometry. *Environ. Sci. Technol. Lett.* **2018**, *5*, 348−353.

(57) van Mourik, L. M.; Toms, L.-M. L.; He, C.; Banks, A.; Hobson, P.; Leonards, P. E. G.; de Boer, J.; Mueller, J. F. Evaluating age and temporal trends of chlorinated paraffins in pooled serum collected from males in Australia between 2004 and 2015. *Chemosphere* **2020**, *244*, 125574.

(58) Li, T.; Wan, Y.; Gao, S.; Wang, B.; Hu, J. High-Throughput Determination and Characterization of Short-, Medium-, and Long-Chain Chlorinated Paraffins in Human Blood. *Environ. Sci. Technol.* **2017**, *51*, 3346−3354.

(59) Zhang, C.; Chang, H.; Wang, H.; Zhu, Y.; Zhao, X.; He, Y.; Sun, F.; Wu, F. Spatial and Temporal Distributions of Short-, Medium-, and Long-Chain Chlorinated Paraffins in Sediment Cores from Nine Lakes in China. *Environ. Sci. Technol.* **2019**, *53*, 9462−9471.

(60) Yuan, B.; Brüchert, V.; Sobek, A.; de Wit, C. A. Temporal Trends of C8−C36 Chlorinated Paraffins in Swedish Coastal Sediment Cores over the Past 80 Years. *Environ. Sci. Technol.* **2017**, *51*, 14199−14208.

(61) Brandsma, S. H.; van Mourik, L.; O'Brien, J. W.; Eaglesham, G.; Leonards, P. E.; de Boer, J.; Gallen, C.; Mueller, J.; Gaus, C.; Bogdal, C.

Medium-Chain Chlorinated Paraffins (CPs) Dominate in Australian Sewage Sludge. *Environ. Sci. Technol.* **2017**, *51*, 3364−3372.

(62) Wong, F.; Suzuki, G.; Michinaka, C.; Yuan, B.; Takigami, H.; de Wit, C. A. Dioxin-like activities, halogenated flame retardants, organophosphate esters and chlorinated paraffins in dust from Australia, the United Kingdom, Canada. *Chemosphere* **2017**, *168*, 1248−1256.

(63) Brits, M.; de Boer, J.; Rohwer, E. R.; De Vos, J.; Weiss, J. M.; Brandsma, S. H. Short-, medium-, and long-chain chlorinated paraffins in South African indoor dust and cat hair. *Chemosphere* **2020**, *238*, 124643.

(64) Schinkel, L.; Bogdal, C.; Canonica, E.; Cariou, R.; Bleiner, D.; McNeill, K.; Heeb, N. V. Analysis of Medium-Chain and Long-Chain Chlorinated Paraffins: The Urgent Need for More Specific Analytical Standards. *Environ. Sci. Technol. Lett.* **2018**, *5*, 708−717.