

## Supporting Information

**Non-ignorable dipeptides as nitrogenous precursors in chlor(am)ination:**

**Tyrosyl-*L*-phenylalanine generates higher toxicity to mammalian cells than**

**aromatic monomeric amino acids**

Yi-Yi Hu<sup>1</sup>, Yao Lu<sup>2</sup>, Yu-En Guo<sup>1</sup>, Hua Li<sup>3</sup>, Ying Chen<sup>1</sup>, Min Liu<sup>1</sup>, Wei Hu<sup>4</sup>, Qian-Yuan Wu<sup>5</sup>, Ye

Du (✉)<sup>1</sup>

1 College of Architecture and Environment, Sichuan University, Chengdu 610000, China

2 Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University,  
Hong Kong 999077, China; The Hong Kong Polytechnic University Shenzhen Research Institute,  
Shenzhen 518057, China

3 Core Research Facilities, Southern University of Science and Technology, Shenzhen 518055,  
China

4 Department of Civil and Environmental Engineering, University of Alberta, Edmonton AB T6G  
1H9, Canada

5 Key Laboratory of Microorganism Application and Risk Control of Shenzhen, Guangdong  
Provincial Engineering Research Center for Urban Water Recycling and Environmental Safety,  
Institute of Environment and Ecology, Tsinghua Shenzhen International Graduate School,  
Tsinghua University, Shenzhen 518055, China

---

✉ Corresponding author  
E-mail: duyeah@scu.edu.cn

## Contents

Text S1 Sources and purities of the reagents .....	4
Text S2 Pretreatment, storage and parameter analysis of wastewater samples .....	5
Text S3 Solid-phase extraction .....	5
Text S4 Cell culture .....	6
Text S5 Cytotoxicity assay and cytotoxicity equivalent quantification.....	6
Text S6 DNA DSB assay and genotoxicity equivalent quantification.....	8
Text S7 Liquid-liquid extraction for typical byproducts analysis.....	10
Text S8 Determination of total organic chlorine.....	10
Text S9 Detection method of UPLC-HRMS .....	11
Text S10 UPLC-HRMS data analysis using Compound Discoverer.....	12
Text S11 Structural and quantitative analysis .....	13
Text S12 Statistical Analysis.....	14
Table S1 Water quality parameters of samples used in this study* .....	15
Table S2 The method detection limits (MDLs) of typical DBPs in this study .....	16
Table S3 MS information of tentative byproducts in samples after chlorination .....	17
Table S4 MS information of tentative byproducts in samples after chloramination ...	18
Table S5 The method detection limits (MDLs) of seven confirmed DBPs in this study .....	19
Table S6 Formation yields ( $\mu\text{g-DBP/mg-C}$ ) of confirmed DBPs from precursors of Phe, Tyr, and Tyr-Phe during chlorination and chloramination .....	20
Figure S1 Workflow for selective qualitative and quantitative analysis of DBPs .....	21
Figure S2 Concentration-effect curves of cytotoxicity of after chlorination.....	22
Figure S3 Concentration-effect curves of cytotoxicity of Tyr, Phe and Tyr-Phe after chloramination .....	23
Figure S4 Concentration-effect curves of genotoxicity of Tyr, Phe and Tyr-Phe after chlorination .....	24
Figure S5 Genotoxicity changes of Tyr, Phe and Tyr-Phe after chloramination.....	25
Figure S6 Venn diagrams showing the distribution of DBPs formulas .....	26

Figure S7 Number and total peak area of DBPs after chlorination and chloramination .....	27
Figure S8 Van Krevelen diagram of byproducts after chloramination .....	28
Figure S9 The molecular distribution of chlorinated DBPs after chloramination .....	29
Figure S10 Molecular composition distribution and characteristics of halogen-containing components after chloramination.....	30
Figure S11 Verification of byproducts generated by chlorination and chloramination	31
Figure S12 Structures of tentative byproducts formed after chlorination and chloramination of Tyr-Phe .....	32
Figure S13 Structure inference of tentative byproducts generated by chlorination and chloramination .....	33
Figure S14 Detection of eleven kinds of byproducts after chlorination and chloramination .....	34
Figure S15 The cytotoxicity changes of three secondary effluents after chlorination.	35
Figure S16 The cytotoxicity changes of three secondary effluents after chloramination .....	36
Figure S17 Contribution of the confirmed byproducts to the overall cytotoxicity of secondary effluents .....	37
References.....	38

## Text S1 Sources and purities of the reagents

Sodium hypochlorite (NaClO, effective chlorine  $\geq 5.5\%$ ), ammonium chloride ( $\text{NH}_4\text{Cl}$ ,  $\geq 99.5\%$ ) were purchased from KESHI, China. Sodium hydroxide (NaOH,  $\geq 99\%$ ), sulfuric acid ( $\text{H}_2\text{SO}_4$ ), 2-Chloroacetamide ( $\geq 99\%$ ), dichloroacetamide ( $\geq 99\%$ ), 2,5-dichlorohydroquinone ( $\geq 95\%$ ), 2,4-dichlorophenol ( $\geq 99.5\%$ ), 3-chloro-5-hydroxybenzotrile ( $\geq 95\%$ ), 4-chlorophenol ( $\geq 99\%$ ), 3,5-dichloro-4-hydroxybenzotrile ( $\geq 97\%$ ), Chloroacetic acid ( $\geq 99\%$ ), L-Aspartic acid ( $\geq 99\%$ ), L-Serine ( $\geq 99\%$ ), L-Tryptophan ( $\geq 99\%$ ), L-Phenylalanine ( $\geq 99\%$ ), L-Tyrosine ( $\geq 99\%$ ) were purchased from Aladdin, China. Tyr-Phe ( $\geq 98\%$ ), Trp-Phe ( $\geq 98\%$ ), Trp-Trp ( $\geq 98\%$ ), Asp-Phe ( $\geq 98\%$ ), Gly-Tyr ( $\geq 98\%$ ), Gly-Phe ( $\geq 98\%$ ) were purchased from Hanhong, China. 4-chloro-*d*-phenylalanine ( $\geq 95\%$ ), Trichloroacetaldehyde ( $\geq 99\%$ ), Chloroacetone ( $\geq 95\%$ ), 1,1-Dichloro-2-propanone ( $\geq 95\%$ ), 1,1,1-Trichloroacetone ( $\geq 95\%$ ), Dichloroacetonitrile ( $\geq 98\%$ ) were purchased from Macklin, China. Chloroform ( $\geq 99.9\%$ ) was purchased from Knowles, China.

For the toxicity assays, dimethyl sulfoxide (DMSO,  $>99.7\%$ ), fetal bovine serum (FBS), Dulbecco's modified Eagle's medium/nutrient mixture F-12 (DMEM/F-12, 1:1) and penicillin–streptomycin was purchased from Thermo Fisher Scientific (USA). Cell Counting Kit-8 (CCK-8, MeilunBio, China) was obtained from Dojindo Molecular Technologies (Japan). Paraformaldehyde, Hoechst 33342 and 4-nitroquinoline N-oxide (4-NQO,  $\geq 98\%$ ) were purchased from Sigma-Aldrich (USA). Triton X-100 and Albumin Bovine V were purchased from Solarbio (China). Phospho-histone H2AX (S139) rabbit antibody and antirabbit IgG Fab2 Alexa Fluor

647 molecular probes were obtained from Cell Signaling Technology (USA).

### **Text S2 Pretreatment, storage and parameter analysis of wastewater samples**

Before conducting parameter measurements and formal experiments, wastewater samples were filtered through cellulose acetate membranes (0.45  $\mu\text{m}$ ) and stored at 4  $^{\circ}\text{C}$ . The dissolved organic carbon (DOC) and pH were measured using a total organic carbon (TOC) analyzer (TOC-L, Shimadzu, Japan) and a pH meter (PB-10, Sartorius Design, Germany). UV-1800 trophotometer (Shimadzu, Japan) was used to record the UV absorbance at 254 nm. The concentration of ammonia nitrogen ( $\text{NH}_4^+$ ) were measured using a N4 UV-Vis Spectrophotometer (INESA, China).

### **Text S3 Solid-phase extraction**

Before solid-phase extraction, methanol and ultrapure water were sequentially passed through the suction line to remove residual organic substances in the line. The SPE column (Oasis HLB, 6 mL, Waters, USA) was activated with 10 mL of methanol, and then 10 mL of ultrapure water was passed through the column to wash away the methanol. Before extraction, the pH of water samples was adjusted to  $2.00 \pm 0.02$  with 2 mol/L  $\text{H}_2\text{SO}_4$ , and the flow rate of the water sample during extraction was controlled to be no more than 5 mL/min. After the water samples were completely extracted, the suction was continued until the SPE column was completely dry. Then, the organic substances adsorbed on the SPE column were eluted successively with 5 mL of methanol, 2 mL of acetone and 2 mL of dichloromethane into a clean glass

tube. The eluted samples were dried under high-purity nitrogen gas. The dried samples were stored in a -20 °C refrigerator for further use.

#### **Text S4 Cell culture**

The Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12, 1:1) supplemented with double antibiotics and fetal bovine serum was used as the culture medium for Chinese hamster ovary cells (CHO-K1, ZhongQiaoXinZhou Biotechnology, China). The cells were cultured in a 37 °C incubator with 5% CO<sub>2</sub> and subcultured every 48 hours.

#### **Text S5 Cytotoxicity assay and cytotoxicity equivalent quantification**

For cytotoxicity tests,  $1 \times 10^4$  cells/well were seeded in sterile 96-well plates (Nunclon 167008, Thermo Fisher Scientific, USA) and cultured for 12 hours to allow cell adhesion (Wu et al., 2020). The organic substances dried under nitrogen were dissolved in DMEM/F12 containing 0.5% DMSO (v:v). The dissolved samples were prepared into six concentration gradients. DMEM/F12 containing 0.5% DMSO was used as the negative control, while phenol dissolved in 0.5% DMSO was prepared at six known concentration gradients as a reference compound. All the 96-well plates for toxicity assay were covered with the sterilized seal films to avoid the cross contamination.

After 48 hours of exposure, the culture medium was removed, and 100 μL of PBS was added to each well to wash away the residual medium. Then, 100 μL of

CCK-8 detection solution (CCK-8: DMEM/F12 = 1:9) was added to each well. The plates were incubated in an incubator for 2 hours, and the absorbance at 450 nm of each well was measured using a microplate reader (BioTek, China). Each test was repeated 3-6 times.

The cell viability (CV) of different samples and the positive control was calculated using Equation (1):

$$CV=(A_s-A_b)/(A_n-A_b) \quad (1)$$

where,

$A_s$ ——Absorbance of the sample or positive control at 450 nm.

$A_b$ ——Absorbance of the blank control (only the Cell Counting Kit-8 reagent dissolved in DMEM/F12 medium).

$A_n$ ——Absorbance of the negative control at 450 nm.

Regression analysis was performed on the CV values of each sample and phenol at different concentrations using the BiDoseResp function to fit the data and establish a concentration-response curve. Based on the regression analysis, the concentration at which cell viability is 50% was defined as the lethal concentration 50% ( $LC_{50}$ ). The cytotoxic equivalent of the samples was calculated using Equation (2):

$$CE = LC_{p,50}/LC_{s,50} \quad (2)$$

where,

CE——Cytotoxic equivalent, expressed in mg-Phenol/L.

$LC_{p,50}$ —— $LC_{50}$  of the reference compound phenol, in mg/L.

$LC_{s,50}$ —— $LC_{50}$  of the sample, in concentrated multiples.

The toxicity contribution of the compound is calculated based on its concentration and

LC<sub>50</sub> value using Equation (3):

$$TC = (c / LC_{S,50}) \times LC_{p,50} / CE \quad (3)$$

where,

TC: toxicity contribution.

c: concentration of each DBP.

LC<sub>S,50</sub>——LC<sub>50</sub> of the sample, in concentrated multiples.

LC<sub>p,50</sub>——LC<sub>50</sub> of the reference compound phenol, in mg/L.

CE——Cytotoxic equivalent, expressed in mg-Phenol/L.

### **Text S6 DNA DSB assay and genotoxicity equivalent quantification**

To perform the genotoxicity assay,  $7 \times 10^3$  cells were seeded in each well of sterile 96-well plate (Phenoplate 6055302, Perkinelmer, USA), pre-cultured for 12 h, and exposed to organic byproducts dissolved in DMEM/F12 containing 0.5% DMSO at 6-8 concentrations. DMEM/F12 containing 0.5% DMSO was used as the negative control, while 4-nitroquinoline N-oxide (4-NQO) dissolved in DMEM/F12 containing 0.5% DMSO was used as a reference compound. All the 96-well plates for toxicity assay were covered with the sterilized seal films to avoid the cross contamination. The 4-8 replicates of each test were performed.

After exposure for 24 h, cells were fixed with paraformaldehyde, permeated with Triton-100, and blocked with bovine serum albumin. Cells were incubated with the primary antibody phospho-histone H2AX and then stained with the secondary

antibody Alexa Fluor® 647 conjugate together with Hoechst 33258. After the staining, images were captured using a High Content Analysis system (Operaphenix Plus, PerkinElmer, USA) with a 40× objective lens. The pH2AX foci were obtained from the CY5 channel and nucleus DNA was obtained from the DAPI channel. In each well of the plate, the number of pH2AX per cell was calculated by the total number of pH2AX foci over the number of nucleus.

The relative induction rate (IR), which represents the ratio of the number of pH2AX foci per cell for the sample or positive control relative to the negative control, was calculated using the Equation (4) and (5):

$$P = N_p/N_c \quad (4)$$

$$IR = P_s/P_n \quad (5)$$

where,

P—Number of pH2AX foci per cell.

$N_p$ —Total number of pH2AX foci in the field of view.

$N_c$ —Total number of cell nuclei in the field of view.

IR—Relative induction rate.

$P_s$ —Number of pH2AX foci per cell for the sample or positive control.

$P_n$ —Number of pH2AX foci per cell for the negative control.

The IR for each sample and 4-NQO at different concentrations were analyzed using regression analysis with the BiDoseResp function to fit the data and establish a concentration-effect curve. Based on this curve, the concentration corresponding to an IR of 1.5 was identified and defined as  $IR_{1.5}$ . The genotoxicity equivalent (GE) of the

samples was then calculated using the following formula (6):

$$GE = IR_{4\text{-NQO},1.5}/IR_{S,1.5} \text{ (6)}$$

where,

GE—Genotoxicity equivalent,  $\mu\text{g-4-NQO/L}$ .

$IR_{4\text{-NQO},1.5}$ —Concentration of 4-NQO corresponding to  $IR = 1.5$ ,  $\mu\text{g/L}$ .

$IR_{S,1.5}$ —Concentration of the sample corresponding to  $IR = 1.5$ .

### **Text S7 Liquid-liquid extraction for typical byproducts analysis**

30 mL of the water sample to be tested was placed in a 40 mL brown glass bottle, and 3 mL of the internal standard stock solution (MTBE containing 100  $\mu\text{g/L}$  1,2-dibromopropane). Then, 6 g of  $\text{Na}_2\text{SO}_4$  was added and the bottle was vigorously shaken immediately to dissolve most of the contents. The bottle was then rotated and shaken for 15 minutes, followed by a 15-minute rest. 1.5 mL of the organic phase was transferred to a gas chromatography vial and left to stand for 10 minutes. After ensuring there was no water present, it was used for gas chromatography detection (GC-ECD, 9720Plus, Fuli Instruments, China).

### **Text S8 Determination of total organic chlorine**

40 mL of the sample was acidified to  $\text{pH} = 2 \pm 0.02$  with 1 mol/L nitric acid. Two activated carbon adsorption columns were connected in series and washed successively with  $\text{KNO}_3$  (5000  $\text{mg-NO}_3^-/\text{L}$ ) and ultrapure water under a nitrogen flow. Then, the activated carbon columns were placed in a furnace and heated at 950  $^\circ\text{C}$  for

10 min. The combustion gases were blown into the KOH absorption solution through oxygen. Finally, the chloride ion concentrations in the absorption solution were determined using an ion chromatograph (Dionex Aquion, Thermo Fisher Scientific, USA) in combination with a Dionex IonPac™ AS19 chromatographic column (Dionex Aquion, Thermo Fisher Scientific, USA). 2,4,6-trichlorophenol was used as the standard recovery substance for organic chlorine, the recovery rate was measured as 87% in this study. Each sample was tested three times.

### **Text S9 Detection method of UPLC-HRMS**

The dried organic matter obtained from the aforementioned SPE procedure was reconstituted in a methanol/ultrapure water (1:1, v/v) mixture. An analytical column (hypersil gold C18 column 2.1 × 100 mm, 1.9 μm particle size, Thermo Fisher Scientific, USA) was operated in gradient mode. The column was used to separate 10 μL samples with ultrapure water (phase A) and acetonitrile (phase B) as mobile phases, and the flow rate was 0.3 mL/min. The gradient was as follow: phase B was increased to 95% in 8 min from 5% at initial, it was returned back to 5% and held for 3 min.

Mass spectrometric (MS) detection was achieved using a heated electrospray ionization (HESI) source in positive ion mode, and a scan range of mass to charge ratios ( $m/z$ ) from 50–650. The ion source settings were: 30 psi sheath gas (nitrogen), 10 psi aux gas (nitrogen), 320°C gas temperature, and 3 kV capillary voltage. The HESI source generated protonated molecular ions ( $[M+H]^+$ ). The MS-data were

recorded in centroid mode using a date-dependent MS/MS (dd-MS<sup>2</sup>) method. The resolution of 70,000 was set for full scan and 17,500 was set for dd-MS<sup>2</sup> method separately. The AGC target and maximum IT were  $1 \times 10^5$  ions capacity and 50 ms in dd-MS<sup>2</sup> settings. The isolation window was 4 *m/z*. The TopN (N, the number of top most abundant ions for fragmentation) were set to 10. The normalized collision energy (NCE) was set as 10, 25, 40 eV, respectively. Dynamic exclusion time was set to be 10 s.

### **Text S10 UPLC-HRMS data analysis using Compound Discoverer**

Compound Discoverer (CD) 3.1 software (Thermo Scientific, USA) was used to streamline the workflow by integrating blank subtraction, peak picking, molecular formula generation, isotopic pattern comparison, fragmentation pattern assignment, and database searching into a single platform, Peak picking was performed with a mass tolerance of 5 ppm, an intensity tolerance of 30%, a signal-to-noise (S/N) threshold of 3, and a minimum peak intensity of 10,000. Feature merging and grouping were carried out with a mass tolerance of 5 ppm and a retention time (RT) tolerance of 0.1 min. Formula prediction was conducted using a mass tolerance of 5 ppm, a maximum ring double bond equivalents (RDBE) of 40, and a maximum H/C ratio of 3.5. Pattern matching was performed with an intensity tolerance of 30%, an intensity threshold of 0.1%, and an S/N threshold of 3. Fragment matching was also conducted with a mass tolerance of 5 ppm and an S/N threshold of 3. This approach enabled the visualization of the molecular formula, retention time, peak intensity, and

presumptive identity for each peak.

### **Text S11 Structural and quantitative analysis**

High-response chromatographic peaks from the chlorination and chloramination of amino acids and dipeptides were analyzed based on ion fragment losses in the mass spectrum to determine their molecular structures. Compounds exhibiting retention time differences of less than 0.1 minutes between the standard and treated samples, along with matching mass spectrometry fragmentation patterns, were categorized as "confirmed DBPs". Those without available standards were preliminarily identified by comparing their MS/MS spectra with database entries and categorized as "tentative DBPs" if they exhibited consistent retention times and spectral features ([Schymanski et al., 2014](#)).

The confirmed and tentative DBPs were compared with analytical results from actual secondary effluents after chlorination and chloramination to confirm their presence in real wastewater. Quantitative analysis of confirmed DBPs, as well as amino acids and dipeptides, was performed using mass spectrometry standard curves derived from known concentrations of reference standards. Procedural blanks and recovery samples were included in each batch to monitor operational errors and contamination risks. The average recovery rate of the DBPs at 1 ng/mL was 72%–100%, and each sample was injected three times.

## **Text S12 Statistical Analysis**

One-way analysis of variance (ANOVA) was used to evaluate whether there were significant differences between the toxicity results of samples. If a significant F value ( $p < 0.05$ ) was obtained, to determine the lowest toxic effect concentration, a least significance difference (LSD) multiple comparison test with the samples was further performed. The  $LC_{50}$  values were calculated using the concentration–effect curves, and were used to calculate the cytotoxicity equivalents (mg-Phenol/L). The pH2AX induction ratios of samples at different concentration factors with cell viability higher than 70% were used to obtain concentration–effect curves for genotoxicity by regression analysis, and the 1.5-fold induction rate ( $IR_{1.5}$ ) values were calculated using the concentration–effect curves. The  $IR_{1.5}$  values were used to calculate the genotoxicity equivalents ( $\mu\text{g-4-NQO/L}$ ). One-way ANOVA was also performed to determine the significant differences among the toxicity equivalents.

**Table S1** Water quality parameters of samples used in this study\*

samples	TOC (mg/L)	UV <sub>254</sub>	pH	NH <sub>4</sub> <sup>+</sup> (mg-N/L)
SE1	4.1±0.2	0.121±0.001	7.60±0.03	0.12±0.01
SE2	4.5±0.2	0.132±0.001	6.98±0.02	0.15±0.01
SE3	3.4±0.5	0.102±0.001	6.71±0.02	0.09±0.01

\*: Calculated as the average of three independent tests

**Table S2** The method detection limits (MDLs) of typical DBPs in this study

Typical DBPs	MDL ( $\mu\text{g/L}$ )
Trichloromethane	0.02
Chloroacetone	0.006
Dichloroacetonitrile	0.04
Trichloroacetaldehyde	0.01

**Table S3** MS information of tentative byproducts in samples after chlorination

Chemical	Formula	SE1		SE2		SE3	
		RT(min)	m/z	RT(min)	m/z	RT(min)	m/z
seclazone	C <sub>10</sub> H <sub>8</sub> ClNO <sub>3</sub>	10.66	224.0717 189.1027	10.60	224.0720; 189.1028	n.d.	n.d.
3-Chloro-7-nitrobenzofuran-2-carboxylic acid	C <sub>9</sub> H <sub>4</sub> ClNO <sub>5</sub>	n.d.	n.d.	9.24	239.9708; 195.9805	n.d.	n.d.

Notes: “n.d.” means not detected.

**Table S4** MS information of tentative byproducts in samples after chloramination

Chemical	Formula	SE1		SE2		SE3	
		RT (min)	m/z	RT (min)	m/z	RT (min)	m/z
2-(4-Chlorophenyl)-5-methyl-1,3-oxazole-4-carboxylic acid	C <sub>10</sub> H <sub>10</sub> ClNO <sub>3</sub>	n.d.	n.d.	9.09	226.0277; 182.0978	9.13	226.0278; 182.0972
(S)-2-((2-Chlorophenyl) formamido) propanoic acid	C <sub>11</sub> H <sub>8</sub> ClNO <sub>3</sub>	n.d.	n.d.	n.d.	n.d.	11.8 7	236.0121; 192.0263

Notes: "n.d." means not detected.

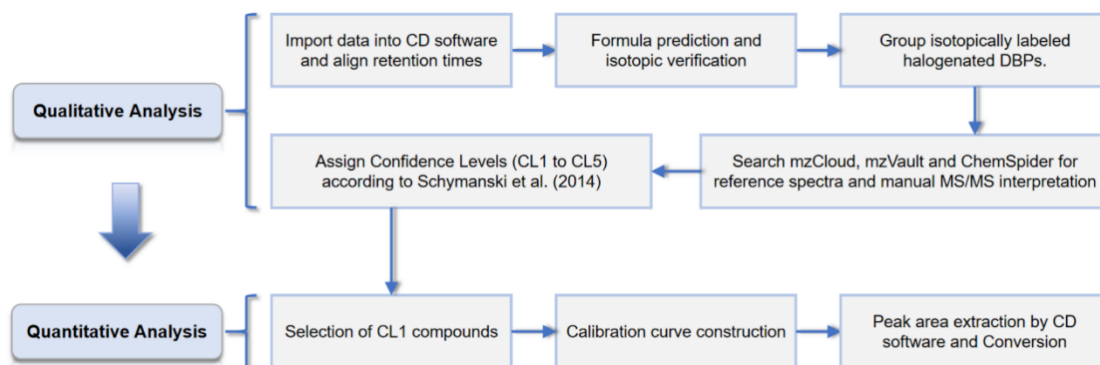
**Table S5** The method detection limits (MDLs) of seven confirmed DBPs in this study

Confirmed DBPs	MDL ( $\mu\text{g/L}$ )
4-chlorophenol	0.1
2,4-dichlorophenol	0.2
2,4,6-trichlorophenol	0.07
4-chloro- <i>d</i> -phenylalanine	0.1
3,5-dichloro-4-hydroxybenzotrile	0.04
3-chloro-5-hydroxybenzotrile	0.03
2,5-dichlorohydroquinone	0.08

Table S6 Formation yields ( $\mu\text{g-DBP/mg-C}$ ) of confirmed DBPs from precursors of Phe, Tyr, and Tyr-Phe during chlorination and chloramination

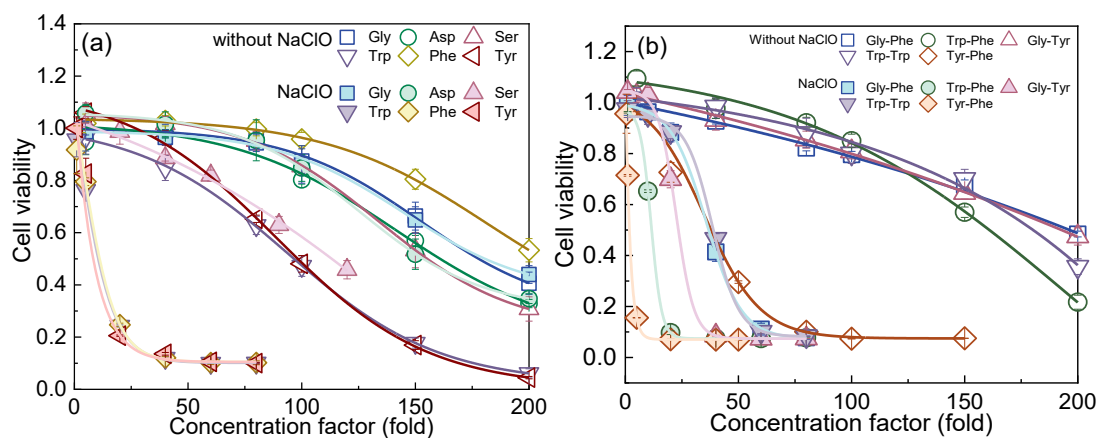
confirmed DBPs	Chlorination			Chloramination		
	Phe	Tyr	Tyr-Phe	Phe	Tyr	Tyr-Phe
4-chlorophenol	n.d.	31.2	68.6	n.d.	16.4	36.8
2,4-dichlorophenol	n.d.	n.d.	n.d.	11.4	n.d.	n.d.
2,4,6-trichlorophenol	8.9	11.4	n.d.	n.d.	n.d.	n.d.
4-chloro- <i>D</i> -phenylalanine	n.d.	n.d.	30.5	n.d.	n.d.	25.8
3,5-dichloro-4-hydroxybenzoxitrile	16.3	n.d.	16.1	7.9	16.8	15.9
3-chloro-5-hydroxybenzoxitrile	n.d.	n.d.	9.1	n.d.	n.d.	n.d.
2,5-dichlorohydroquinone	n.d.	27.8	n.d.	6.2	15.6	23.6

Notes: "n.d." means not detected.



**Fig. S1** Workflow for selective qualitative and quantitative analysis of DBPs.

Reprinted from [Guo et al., 2025](#), *Water Research*, Copyright (2025), with permission from Elsevier.



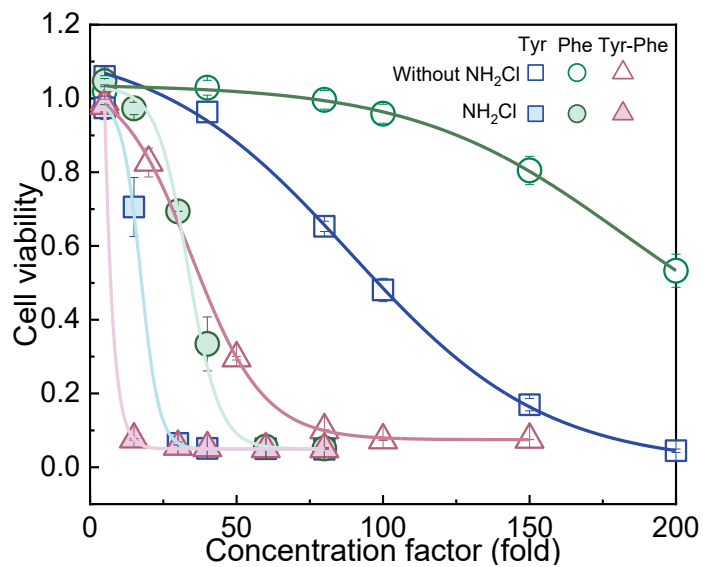
**Fig. S2 Concentration-effect curves of cytotoxicity of after chlorination:**

**(a)** Amino acids; **(b)** Dipeptides. (Gly: Glycine; Asp: Aspartic acid; Ser: serine;

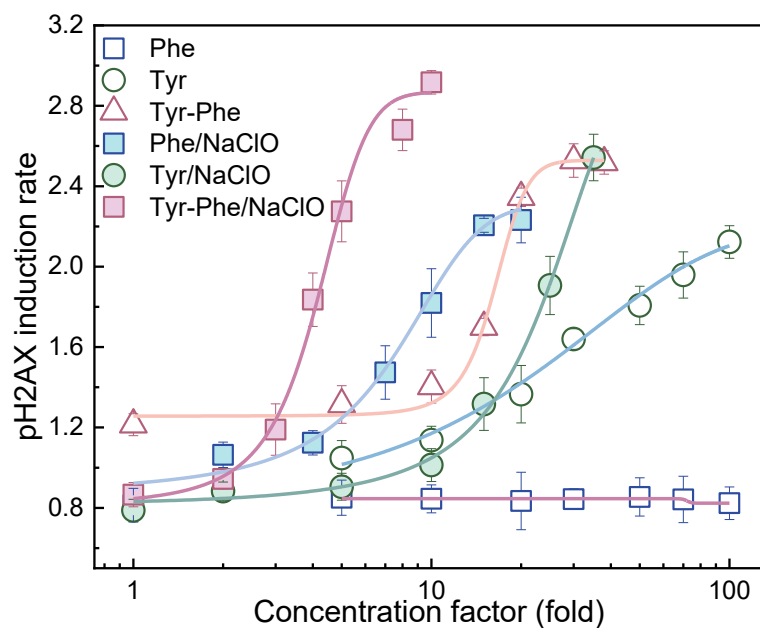
Trp: Tryptophan; Phe: Phenylalanine; Tyr: Tyrosine; Gly-Phe: glycyl-L-

phenylalanine; Trp-Phe: Tryptophylphenylalanine; Gly-Tyr: Glycyltyrosine; Tyr-

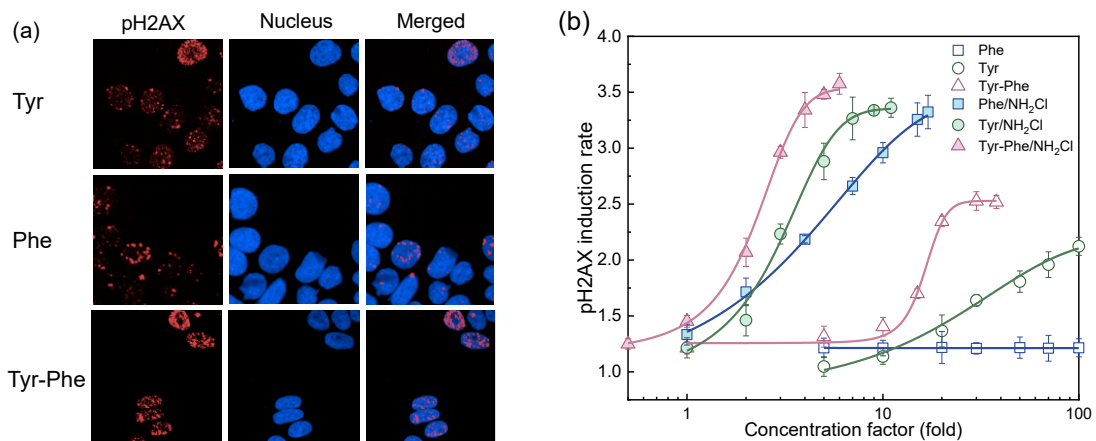
Phe: Tyrosyl-L-phenylalanine; Trp-Trp: Tryptophylltryptophan).



**Fig. S3 Concentration-effect curves of cytotoxicity of Tyr, Phe and Tyr-Phe after chloramination** (Phe: Phenylalanine; Tyr: Tyrosine; Tyr-Phe: Tyrosyl-*L*-phenylalanine).

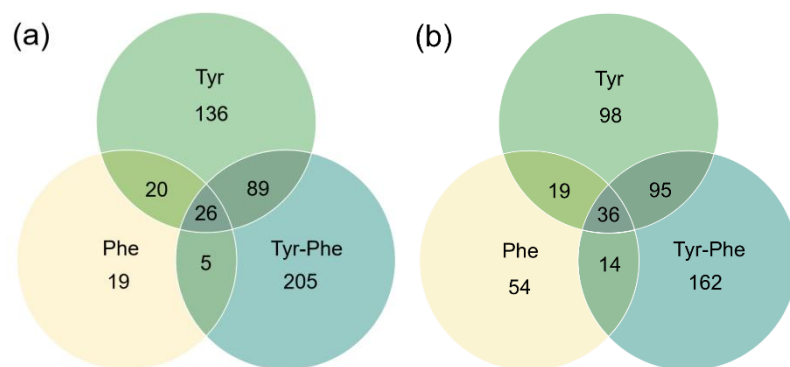


**Fig. S4 Concentration-effect curves of genotoxicity of Tyr, Phe and Tyr-Phe after chlorination** (Phe: Phenylalanine; Tyr: Tyrosine; Tyr-Phe: Tyrosyl-L-phenylalanine).

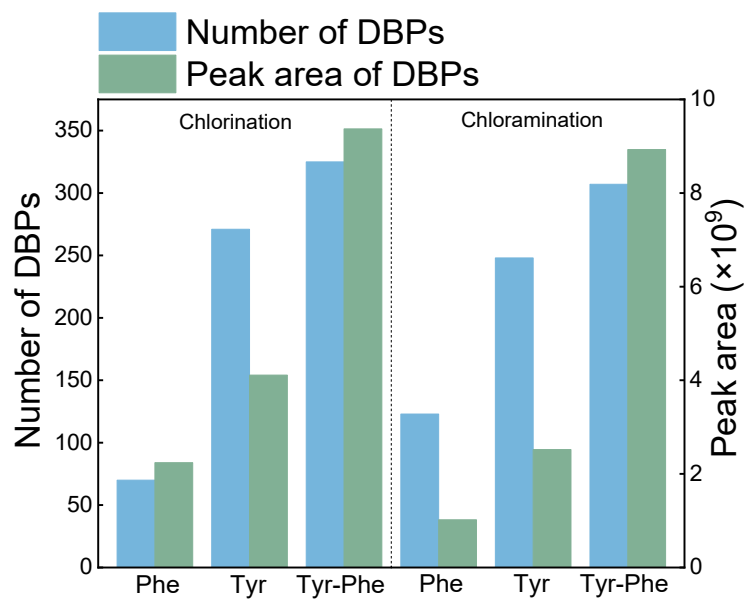


**Fig. S5 Genotoxicity changes of Tyr, Phe and Tyr-Phe after**

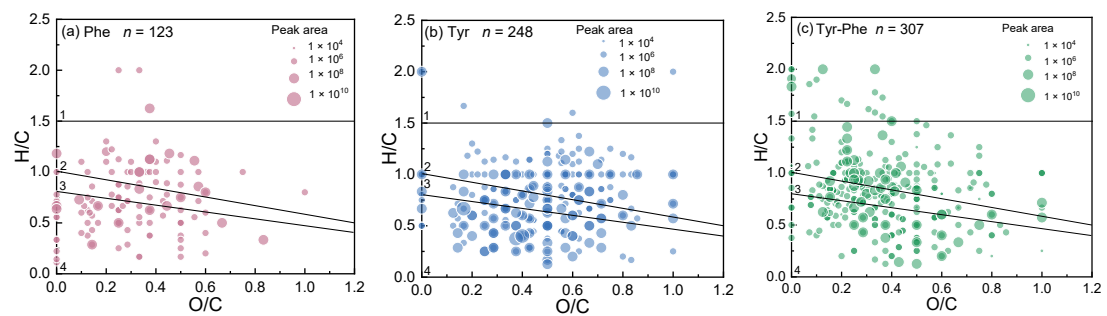
**chloramination:** (a) Immunofluorescence staining of nuclei (blue) and phosphorylated H2AX (pH2AX) foci (red) of Tyr, Phe and Tyr-Phe exposed to chloramination; (b) Concentration-effect curves of genotoxicity of Tyr, Phe and Tyr-Phe after chloramination (Phe: Phenylalanine; Tyr: Tyrosine; Tyr-Phe: Tyrosyl-*L*-phenylalanine).



**Fig. S6 Venn diagrams showing the distribution of DBPs formulas: (a) chlorination; (b) chloramination. (Phe: Phenylalanine; Tyr: Tyrosine; Tyr-Phe: Tyrosyl-*L*-phenylalanine).**



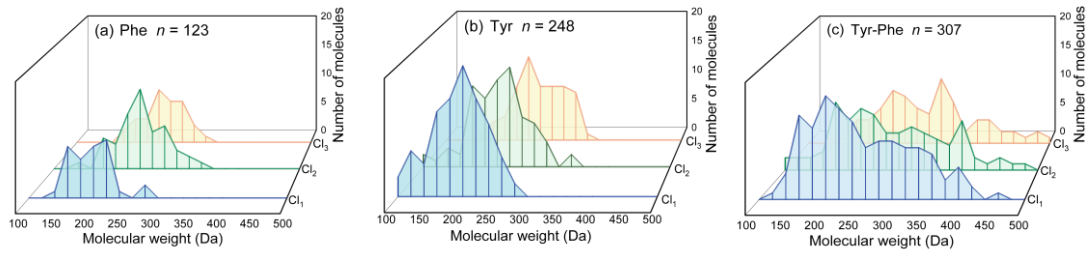
**Fig. S7 Number and total peak area of DBPs after chlorination and chloramination** (Phe: Phenylalanine; Tyr: Tyrosine; Tyr-Phe: Tyrosyl-L-phenylalanine).



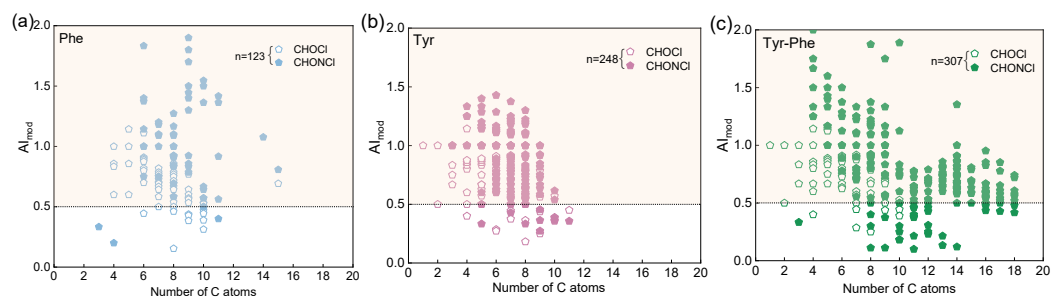
**Fig. S8 Van Krevelen diagram of byproducts after chloramination: (a)**

**Phe; (b) Tyr; (c) Tyr-Phe.** (Phe: Phenylalanine; Tyr: Tyrosine; Tyr-Phe:

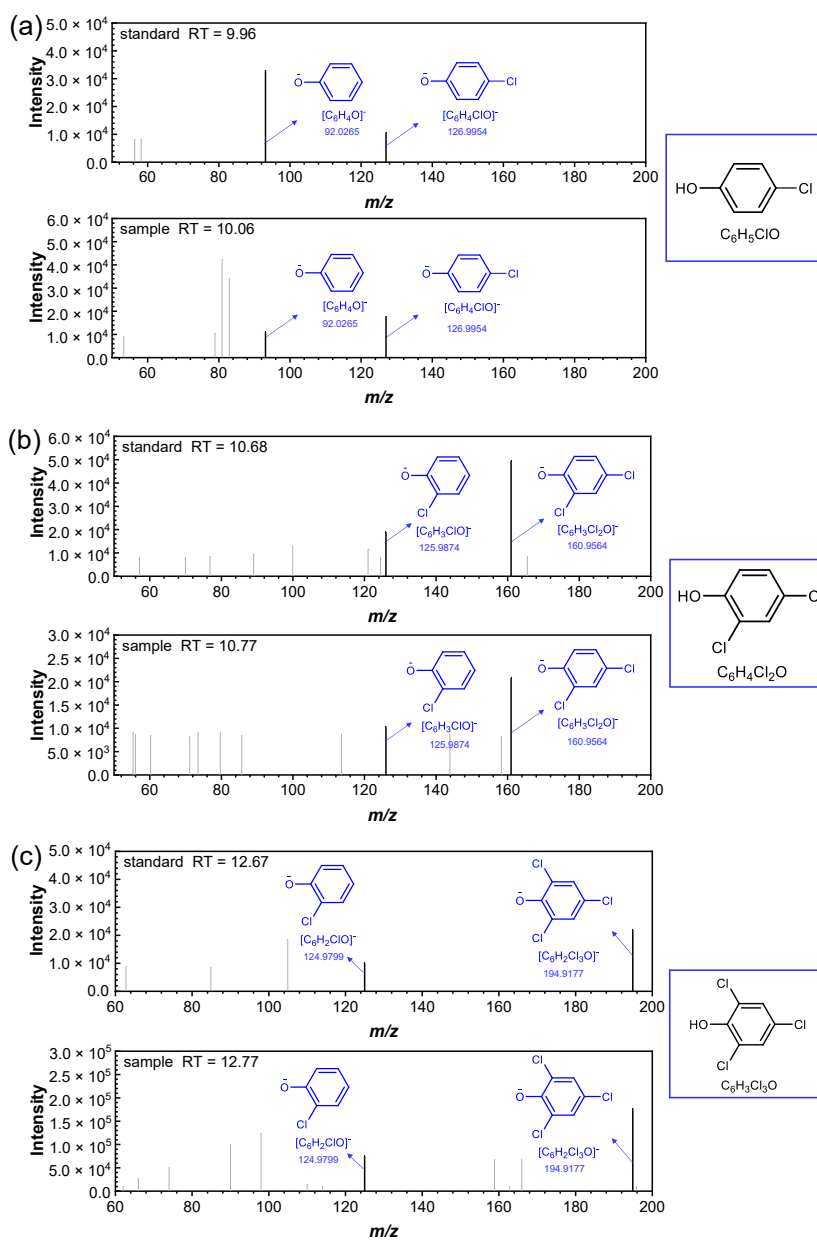
Tyrosyl-*L*-phenylalanine).



**Fig. S9 The molecular distribution of chlorinated DBPs after chloramination: (a) Phe; (b) Tyr; (c) Tyr-Phe. (Phe: Phenylalanine; Tyr: Tyrosine; Tyr-Phe: Tyrosyl-*L*-phenylalanine).**

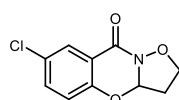


**Fig. S10 Molecular composition distribution and characteristics of halogen-containing components after chloramination: (a) Phe; (b) Tyr; (c) Tyr-Phe. (Phe: Phenylalanine; Tyr: Tyrosine; Tyr-Phe: Tyrosyl-L-phenylalanine).**

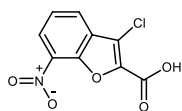


**Fig. S11 Verification of byproducts generated by chlorination and chloramination: (a) 4-Chlorophenol; (b) 2,4-Dichlorophenol; (c) 2,4,6-Trichlorophenol.**

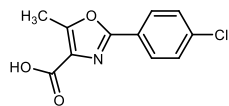
Tentative DBPs



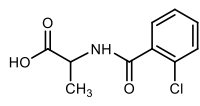
seclazone



3-Chloro-7-nitrobenzofuran-2-carboxylic acid

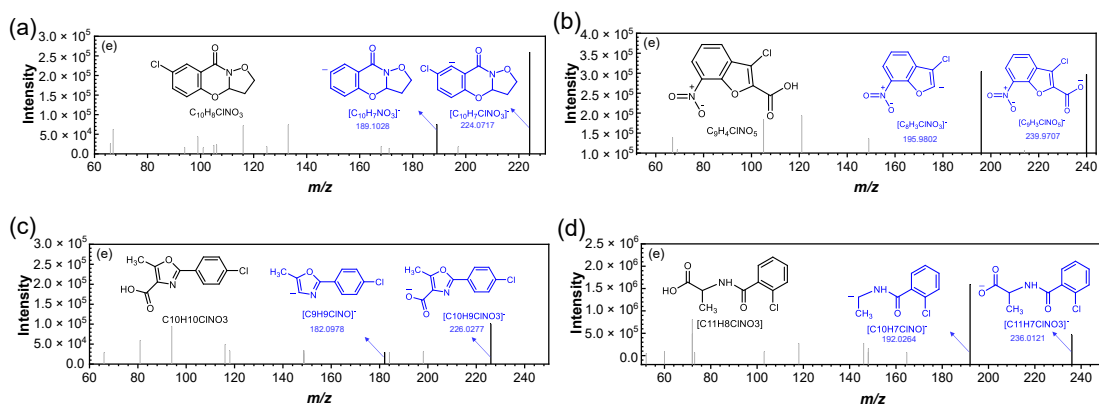


2-(4-Chlorophenyl)-5-methyl-1,3-oxazole-4-carboxylic acid

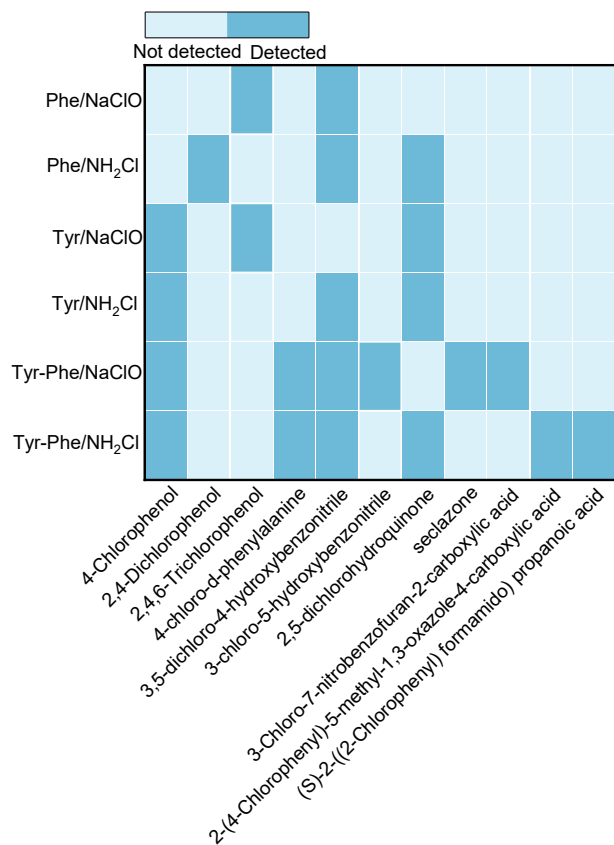


(S)-2-((2-Chlorophenyl)formamido)propanoic acid

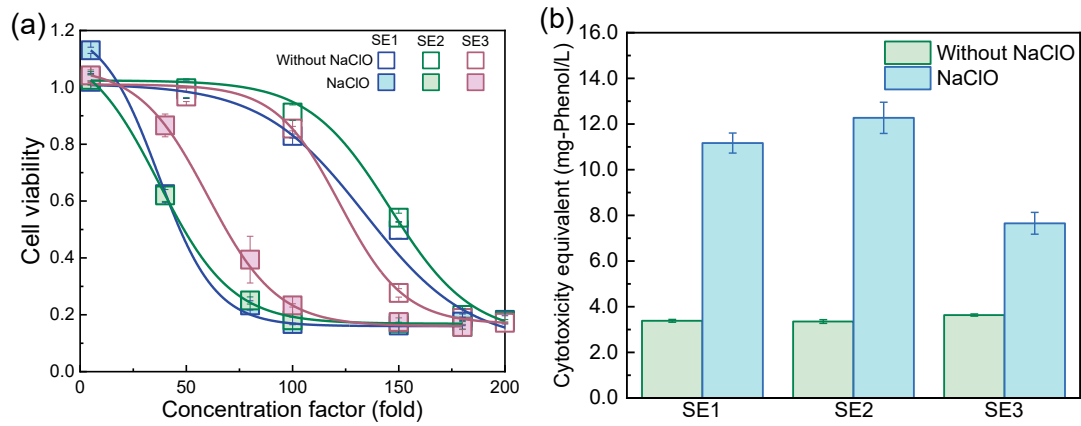
**Fig. S12** Structures of tentative byproducts formed after chlorination and chloramination of Tyr-Phe



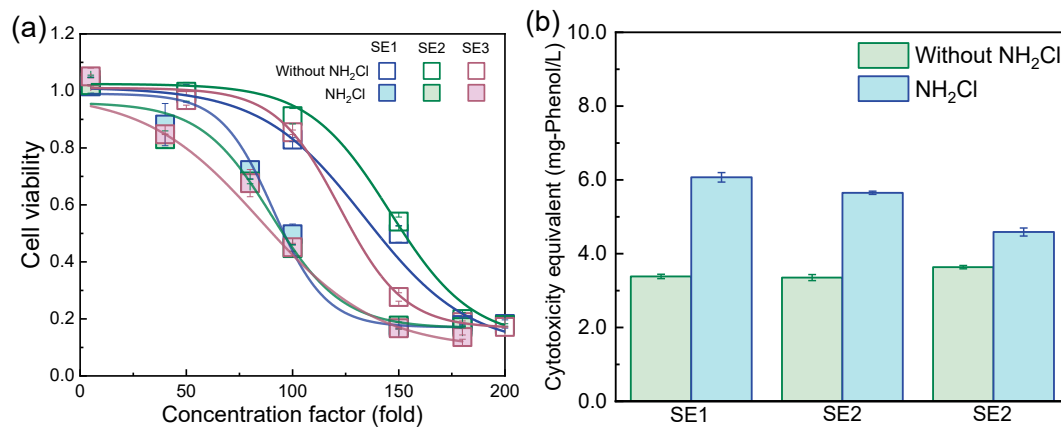
**Fig. S13 Structure inference of tentative byproducts generated by chlorination and chloramination: (a) seclazone; (b) 3-Chloro-7-nitrobenzofuran-2-carboxylic acid; (c) 2-(4-Chlorophenyl)-5-methyl-1,3-oxazole-4-carboxylic acid; (d) (S)-2-((2-Chlorophenyl) formamido) propanoic acid.**



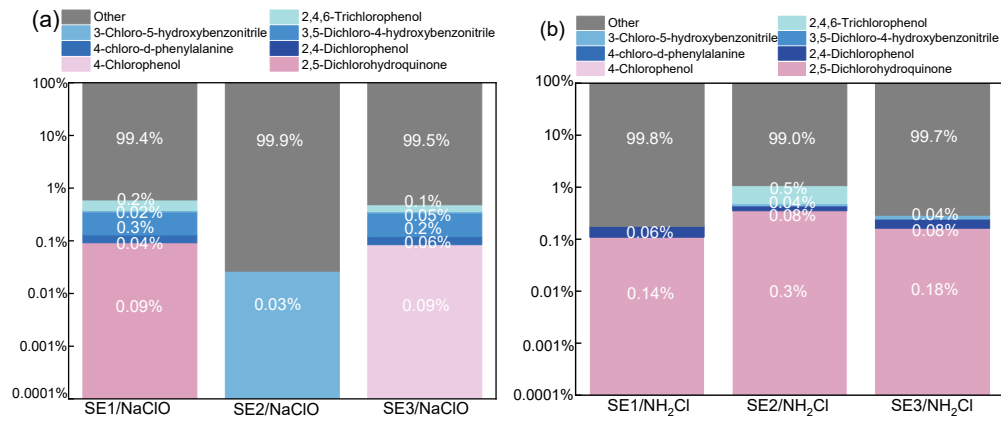
**Fig. S14** Detection of eleven kinds of byproducts after chlorination and chloramination



**Fig. S15 The cytotoxicity changes of three secondary effluents after chlorination: (a) Concentration-effect curves of cytotoxicity of SE1, SE2 and SE3; (b) Cytotoxicity equivalents of SE1, SE2 and SE3.**



**Fig. S16 The cytotoxicity changes of three secondary effluents after chloramination: (a) Concentration-effect curves of cytotoxicity of SE1, SE2 and SE3; (b) Cytotoxicity equivalents of SE1, SE2 and SE3.**



**Fig. S17 Contribution of the confirmed byproducts to the overall cytotoxicity of secondary effluents: (a) Chlorination; (b) Chloramination.**

## References

- Guo, S., Hu, H., Wang, P., Guo, Y., Lu, Y., Du, Y., 2025. Structural and mechanistic insights into halogenated disinfection byproducts of tannic acid with toxicity analysis during chloramination. *Water Res.*, 124219.
- Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* 48(4), 2097-2098.
- Wu, Q., Yang, L., Zhang, X., Wang, W., Lu, Y., Du, Y., Lu, Y., Hu, H., 2020. Ammonia-mediated bromate inhibition during ozonation promotes the toxicity due to organic byproduct transformation. *Environ. Sci. Technol.* 54(14), 8926-8937.