Chlorine Demands of Amino Acids and Amino Sugars in Water

Gebhard B. LUILO*, Koji KOSAKA**, Mari ASAMI**

*Department of Chemistry and Chemical Biology, University of New Mexico, MSC03 2060, 1-University of New Mexico, Albuquerque, NM 87131-0001, USA
**Department of Environmental Health, National Institute of Public Health, 2-3-6 Minami, Wako, Saitama 351-0197, Japan

ABSTRACT
Nitrogenous compounds are ubiquitous in drinking water sources and they increase the chlorine demand of water, shift the breaking point, and also cause strong odors. The present study was performed to investigate the chlorine demands of selected amino acids and amino sugars in water. The results indicated that the chlorine demands of the precursors were in the order: aromatic amino acids (except histidine) > S-amino acids > non-S-amino acids and amino sugars. Aromatic amino acids were expected to have the highest chlorine demand due to chlorine substitution in both the aliphatic and aromatic parts of these molecules. However, the chlorine demand of histidine was threefold lower than those reported previously, which may be attributed to the experimental conditions or shorter contact time used in this study. The chlorine demands of most compounds did not show marked differences at contact times of 15 - 96 h, although some of the nitrogenous organic compounds showed an increasing trend in chlorine demand with time when our 24 h study was compared with previous studies conducted at 72 and 96 h. Chlorine demand in 24 h showed a good correlation with predicted data. Kinetic studies are required to understand how fast the precursors can react with chlorine in typical water treatment contact times and chlorine doses. The information presented here will be useful in controlling disinfection byproducts.

Keywords: amino acids, amino sugars, chlorine demand, QSPR model

INTRODUCTION
Microbial contamination of water is not simply a problem of the past, and the morbidity and mortality rates associated with waterborne diseases are highly related to poverty or the lack of access to sanitary and safe potable water supply systems (Galal-Gorchev, 1996; Hermer, 1999; Ballester and Sunyer, 2000). Epidemiological data have shown that cases, outbreaks, and mortality rates of cholera are higher in developing countries than in developed countries (Swerdlow et al., 1997; WHO, 1998; Andersson and Bohan, 2001; Lee et al., 2002; Hutin et al., 2003; Griffith et al., 2006). Thus, potable water must be routinely disinfected using chlorine (chlorine gas and hypochlorites) or alternative disinfectants to eliminate pathogens (USEPA, 1999). While developed countries can afford both chlorine and alternative disinfectants, such as ozone or chlorine dioxide, chlorine is the only reliable, affordable, and convenient disinfectant in many countries (Galal-Gorchev, 1996; Chaidou et al., 1999; WQHC, 2002). The advantages of chlorine over alternative disinfectants are that it has a wide biocidal spectrum, does not require high levels of technical skill to use, and has low investment and operation costs (Galal-Gorchev, 1996; PNL, 1998; USEPA, 1999; ACC 2008).

While chlorine has played a major role in reducing waterborne diseases worldwide (Galal-Gorchev, 1996; Lee et al., 2002), it has posed a new challenge to potable water
supply authorities because chlorine has high oxidizing power (Larson and Weber, 1994). It reacts with traces of different classes of naturally occurring and anthropogenic organic compounds, which are collectively called dissolved organic matter, to form chlorinated organic compounds (Crittenden et al., 2005). This is because water treatment processes do not completely remove components of dissolved organic matter (Ribas et al., 1991; Volk et al., 2002; Volk et al., 2005), and dissolved organic matter is a complex mixture of various natural and anthropogenic organic molecules of unknown structures (Clesceri et al., 1998; Peuravuori and Pihlaja, 2007).

Nitrogenous organic compounds are one of the classes of dissolved organic matter present in water, and the most abundant forms are amino acids, amino sugars, nucleic acids, and proteins derived from natural and anthropogenic sources (NRC, 1987). Although water treatment can remove most nitrogenous organic compounds, it may still leave trace amounts of organic compounds at the point of chlorination (Pietsch et al., 2001; Dotson and Westerhoff, 2009). For example, the reaction of chlorine with amino acids produces trihalomethanes and haloacetic acids (Hureiki et al., 1994; Bull et al., 2006; Hong et al., 2009) as well as odorous chloroaldimines (Brosillon et al., 2009). Although all nitrogenous organic compounds have reduced nitrogen (amine, amide, imine) with a lone pair of electrons in their structure, they are not equally reactive toward chlorine. The reactivity of organic compounds with chlorine at a given pH may be influenced by a number of factors, such as the number and oxidation states of amine groups, number and relative positions of OH and NH₂ in aromatic rings, and number of aliphatic sulfur atoms (Hureiki et al., 1994; Bull et al., 2006; Arnold et al., 2008; Hong et al., 2009; Luilo and Cabaniss, 2010). Some studies have indicated that activated aromatic amino acids had higher chlorine demands than sulfur-containing amino acids (hereafter denoted as S-amino acids), followed by non-sulfur-containing amino acids (hereafter denoted as non-S-amino acids) after an incubation period of 72 - 96 h at pH 7 - 8 (Hureiki et al., 1994; Hong et al., 2009). However, the typical potable water treatment contact time is much less than 72 h. Thus, the chlorine demands reported over such a long contact time scale represent those that occur both at the treatment plant and in the distribution system due to the reaction of residual chlorine with traces of amino acids. Another study investigated the chlorine demands of five amino acids with a contact time of 15 h at pH 7, and their results were comparable to those determined in other studies performed over longer time scales (de Laat et al., 1982; Hureiki et al., 1994; Hong et al., 2009). These observations suggest that these amino acids are mostly depleted between 15 and 72 h. There are gaps in chlorine demand data for all amino acids and amino sugars at contact times of 15, 24 and 48 h. There is also a lack of kinetic data important for estimating chlorine dose sufficient to sustain chlorine residuals in distribution systems. As it is still difficult to deduce these figures from the literature due to the lack of consistency between studies, our experiments obtained data under consistent conditions to allow comparison, especially with regard to contact time.

This study was performed to determine the chlorine consumptions of amino acids and amino sugars as representative nitrogenous organic compounds in water with a contact time of 24 h. The results were compared to those obtained using other contact times to determine the impact of time on chlorine demands, to predict which amino acids are most likely to be depleted at water treatment plants and which will be depleted in storage tanks or distribution systems.
MATERIALS AND METHODS
Model compounds and reagents
Pure reagents, nitrogenous organic compounds, and ammonium chloride were used in this study for the bench-scale chlorination experiments. All reagents used were of analytical grade unless otherwise noted. The 20 amino acids were purchased from Wako Pure Chemicals (Osaka, Japan); three amino sugars and ammonium chloride were purchased from Nacalai Tesque (Kyoto, Japan) and Sigma-Aldrich (St. Louis, MO, USA). Ultrapure water purified with Gradient A10 (Millipore, Bedford, MA, USA) was used for all experiments. Phosphate buffer was prepared with potassium dihydrogen phosphate and disodium hydrogen phosphate obtained from Wako Pure Chemicals (Osaka, Japan). Hypochlorite (10% - 15%) used as the chlorine source was obtained from Sigma-Aldrich (St. Louis, MO, USA).

Chlorination and chlorine demand
The solutions of model nitrogenous organic compounds (5 µM) and ammonia (5 µM) in glass flasks were buffered at pH 7 using phosphate buffer solution (1 mM), and one buffer solution was used as a control. Each buffered solution was dosed with a small amount of hypochlorite solution while stirring using a magnetic stirrer. The control was subjected to the same reaction conditions except no nitrogenous organic compound or ammonia was present. The sample solutions were added to the flasks to minimize the headspace. After mixing for about 1 min, all flasks were placed in the dark at 20°C for 24 h without mixing. Chlorine residual was set at 1.0 ± 0.2 mg Cl₂/L after 24 h. To obtain the samples at 1.0 ± 0.2 mg Cl₂/L after 24 h, several flasks dosed with different amounts of chlorine were prepared for each compound. That is, chlorine doses were different among the nitrogenous organic compounds and ammonia solutions (i.e., 1.1 - 5.3 mg Cl₂/L). The chlorine residuals for control, nitrogenous organic compounds, and ammonia solutions at 24 h were determined using the standard \( N, N \)-diethyl-p-phenylenediamine (DPD) titration method (Eaton et al., 2005).

Prediction of chlorine demand
Models are valuable tools for validating experimental data. Although there are many empirical models for predicting chlorine demand and disinfection byproduct formation in drinking water, there has been only one previous report of a Quantitative Structure-Property Relationship (QSPR) model (Luilo and Cabaniss, 2010). In this study, the QSPR was used as a model because the structure of a compound is the determining factor in chlorination reaction at optimal contact time, pH, and temperature. Here, the QSPR included eight constitutional descriptors mentioned below.

The QSPR model is summarized by Equation 1. This QSPR was calibrated using 159 compounds and was validated using 42 external data. The QSPR had coefficient of determination of calibration \( R_c^2 \) equal to 0.86 and a residual standard deviation (SDE) of 1.24 mol-Cl₂/mol-Cp (Luilo and Cabaniss 2010). The brief description of each descriptor in the model is described below.

\[
Cl_2 \text{ demand} = 7.61 \times RAI + 1.16 \times ArOH + 3.00 \times ACN + 1.23 \times CI + 2.37 \times AS + 1.01 \times O:C + 0.49 \times ArORact - 0.72 \times ArORnact \quad \text{.... (Eq. 1)}
\]

The ring activation index (RAI) is the ratio of the sum of strong aromatic ring activators...
(NH\textsubscript{2} and OH) to the number of rings. Carboxyl index (CI) is motivated by the observation that carbonyl compounds undergo chlorine substitution reaction via keto-enol tautomerization and that \(\beta\)-dicarbonyl compounds (e.g., 3-oxopentanedioic acid), consume more chlorine than simple ketones (e.g., 2-propanone). The symbol ArOH is the number of phenolic groups in aromatic ring, AS is the number aliphatic sulphur, ACN is the number of aliphatic carbon bonded to amines, O : C is the ratio of atomic oxygen to carbon, ArOR\text{act} is the number of ring-activating alkoxy groups in aromatic ring, i.e., alkoxy groups (weak ring activators) attached to aromatic rings that have no strong ring activators (NH\textsubscript{2} and OH) on them; and ArOR\text{nact} is the number of non-activating groups in aromatic ring, i.e., alkoxy groups attached to aromatic rings that have strong ring activators (NH\textsubscript{2} and OH) on them. A value of zero was given to a molecule that did not have any one of the descriptors. The predictive power of the QSPR model for chlorine demand was evaluated using external validation data (any sets of data that were not used in calibration of the model). The \(q^2\) (coefficient of determination of validation) can be calculated using Equation 3 below (Golbraikh and Tropsha, 2002; Tropsha \textit{et al.}, 2003) and is calculated differently from \(R_c^2\), which can easily be calculated manually using Equation 2. However, most statistical software can generate \(R_c^2\) and residual SDE (how far each data point is from the best fitted line or model) when performing simple or multiple linear regression between response and one variable or a list variables.

\[
R_c^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}
\]  
(Eq. 2)

where, \(y_i\) and \(\hat{y}_i\) are experimental and predicted values respectively and \(\bar{y}\) is the mean of experimental data in calibration dataset. On the other hand the \(q^2\) for external validation, \(q^2_{\text{ext}}\), was calculated using Equation 3 (Golbraikh and Tropsha, 2002; Tropsha \textit{et al.}, 2003).

\[
q^2_{\text{ext}} = 1 - \frac{\sum_{i=1}^{n_{\text{ext}}}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{\text{ext}}}(y_i - \bar{y}_{\text{ext}})^2}
\]  
(Eq. 3)

where \(y_i\) and \(\hat{y}_i\) are experimental and predicted Cl\textsubscript{2} demands (from external validation data) respectively; \(\bar{y}_{\text{ext}}\) is the average of the experimental Cl\textsubscript{2} demands in the entire calibration data (Golbraikh and Tropsha, 2002; Tropsha \textit{et al.}, 2003). Since the \(q^2\) may be influenced by the outliers in training and external data sets (Eq. 2), Golbraikh and Tropsha (2002) recommended using slope, \(k\) and \(R^2\) of regression line obtainable from the plot of predicted values against the experimental values (normal regression) or vice versa (reverse regression) in addition to \(q^2\). The symbols, \(k_i\) and \(R_i^2\) are slope and \(R^2\), respectively, and are obtained from the regression plot of predicted Cl\textsubscript{2} demand and experimental Cl\textsubscript{2} demand with \(y\)-intercept. Whereas \(k_o\) and \(R_o^2\) are slope and \(R^2\), respectively, when the \(y\)-intercept in the regression in the same plot is set to zero (regression through origin). The essence of this test is that if the QSAR/QSPR model is close to perfection, \(k_i\) and \(k_o\), and \(R_i^2\) and \(R_o^2\) should not be significantly different from
each other and both should be closer to 1. Thus, the QSAR/QSPR model is regarded to have high predictive power if it meets the following criteria: $R_c^2 > 0.6$, $q^2 > 0.5$; $(R_i^2 - R_o^2)/R_i^2 < 0.1$; $0.85 \leq k \leq 1.15$ (Golbraikh and Tropsha, 2002; Tropsha et al., 2003).

The root mean square error (RMSE) and model bias deviation (MBD) are also useful in checking the predictive power of the model and were computed using Equations 4 and 5,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \quad \text{……………………………………………..…… (Eq. 4)}$$

$$MBD = \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)}{\sum_{i=1}^{N}y_i} \times 100\% \quad \text{……………………………………………..……. (Eq. 5)}$$

where $y_i$ and $\hat{y}_i$ are experimental and predicted Cl$_2$ demand and $N$ is the total number of observations. The RMSE of external validation is expected to be closer to the residual (SDE) of model calibration for a model with high predictive power. The MBD provides a qualitative diagnosis of the predictive power of the model. A MBD of zero indicates that the model has no prediction bias. However, a negative MBD indicates that the model predicts lower than the experimental value and positive MBD indicates that the model predicts higher than the experimental value. However, the magnitude of MBD does not necessarily indicate how many data were biased higher or lower than expected because one data point (compound) that is either over-predicted or under-predicted may drive total residuals (numerator in the Equation 5) up or down, respectively.

Therefore, in this work the eight descriptors represented in Equation 1 were calculated from the structure of each amino acid and amino sugar in accordance with the procedure described in the literature (Luilo and Cabaniss, 2010). A value of zero was given to any of the eight descriptors if a molecule lacked functional groups needed to calculate it. The descriptors for each molecule were substituted in the QSPR (Equation 1) to estimate the chlorine demand while the predictive power of the QSPR on the model compounds were determined in accordance with the methods described in literature (Golbraikh and Tropsha, 2002; Tropsha et al., 2003).

RESULTS AND DISCUSSION

Chlorine demand of nitrogenous compounds

The chlorine demands (mg Cl$_2$/L) for 20 amino acids, 3 amino sugars, ammonia and ultrapure water (blank) were determined from the difference between chlorine dose and chlorine residual. The chlorine demand of each model compound was corrected by subtracting the chlorine demand of the blank (0.1 mg Cl$_2$/L). The net chlorine demand of each compound was further transformed into a unit of mole of chlorine per mole of compound (mol-Cl$_2$/mol-Cp) (Table 1). The results showed that tyrosine and tryptophan had chlorine demands of 10.99 mol-Cl$_2$/mol-Cp and 12.54 mol-Cl$_2$/mol-Cp, respectively, whereas that of histidine was 4.54 mol-Cl$_2$/mol-Cp.
Table 1 - Chlorine demands of amino acids and amino sugars.

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Cl₂ demand (mg/L)</th>
<th>Compound (µmol/L)</th>
<th>MW (g/mol)</th>
<th>Cl₂ demand (µmol/L)</th>
<th>Cl₂ demand (mol/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycine</td>
<td>1.65</td>
<td>5</td>
<td>75.07</td>
<td>23.24</td>
<td>4.65</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.90</td>
<td>5</td>
<td>89.09</td>
<td>12.68</td>
<td>2.54</td>
</tr>
<tr>
<td>Valine</td>
<td>0.85</td>
<td>5</td>
<td>117.15</td>
<td>11.97</td>
<td>2.39</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>0.90</td>
<td>5</td>
<td>131.17</td>
<td>12.68</td>
<td>2.54</td>
</tr>
<tr>
<td>Leucine</td>
<td>0.90</td>
<td>5</td>
<td>131.17</td>
<td>12.68</td>
<td>2.54</td>
</tr>
<tr>
<td>Serine</td>
<td>1.20</td>
<td>5</td>
<td>105.09</td>
<td>16.90</td>
<td>3.38</td>
</tr>
<tr>
<td>Threonine</td>
<td>1.40</td>
<td>5</td>
<td>76.12</td>
<td>19.72</td>
<td>3.94</td>
</tr>
<tr>
<td>Methionine</td>
<td>2.35</td>
<td>5</td>
<td>149.21</td>
<td>33.10</td>
<td>6.62</td>
</tr>
<tr>
<td>Cysteine</td>
<td>2.65</td>
<td>5</td>
<td>121.16</td>
<td>37.32</td>
<td>7.46</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>1.40</td>
<td>5</td>
<td>133.10</td>
<td>19.72</td>
<td>3.94</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>0.95</td>
<td>5</td>
<td>147.13</td>
<td>13.38</td>
<td>2.68</td>
</tr>
<tr>
<td>Lysine</td>
<td>1.60</td>
<td>5</td>
<td>146.19</td>
<td>22.54</td>
<td>4.51</td>
</tr>
<tr>
<td>Arginine</td>
<td>2.00</td>
<td>5</td>
<td>174.20</td>
<td>28.17</td>
<td>5.63</td>
</tr>
<tr>
<td>Histidine</td>
<td>1.50</td>
<td>5</td>
<td>155.15</td>
<td>21.13</td>
<td>4.23</td>
</tr>
<tr>
<td>Asparagine</td>
<td>2.05</td>
<td>5</td>
<td>132.12</td>
<td>28.87</td>
<td>5.77</td>
</tr>
<tr>
<td>Glutamine</td>
<td>1.30</td>
<td>5</td>
<td>146.14</td>
<td>18.31</td>
<td>3.66</td>
</tr>
<tr>
<td>Proline</td>
<td>1.20</td>
<td>5</td>
<td>115.13</td>
<td>16.90</td>
<td>3.38</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>4.45</td>
<td>5</td>
<td>204.23</td>
<td>62.68</td>
<td>12.54</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>1.00</td>
<td>5</td>
<td>165.19</td>
<td>14.08</td>
<td>2.82</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>3.90</td>
<td>5</td>
<td>181.19</td>
<td>54.93</td>
<td>10.99</td>
</tr>
<tr>
<td>Amino sugar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucosamine</td>
<td>1.60</td>
<td>5</td>
<td>179.19</td>
<td>22.54</td>
<td>4.51</td>
</tr>
<tr>
<td>Galactosamine</td>
<td>1.89</td>
<td>5</td>
<td>179.19</td>
<td>26.62</td>
<td>5.32</td>
</tr>
<tr>
<td>Mannosamine</td>
<td>1.50</td>
<td>5</td>
<td>179.19</td>
<td>21.13</td>
<td>4.23</td>
</tr>
<tr>
<td>Ammonia</td>
<td>0.85</td>
<td>5</td>
<td>17.03</td>
<td>11.97</td>
<td>2.39</td>
</tr>
</tbody>
</table>

*Amino acids are listed from simple to complex.

The low chlorine demand for histidine determined in the present study relative to previous 72 h and 96 h studies may be attributable to the shorter contact time (Hureiki et al., 1994; Hong et al., 2009). However, the chlorine demand for tyrosine in the present study, 10.99 mol-Cl₂/mol-Cp, was similar to 11.40 mol-Cl₂/mol-Cp reported in a previous 15-h study (de Laat et al., 1982). As these two data for tyrosine were not significantly different from each other despite the 9 h difference, it is likely that most of the tyrosine reacts in less than 15 h. However, contact time longer than 24 h may increase the chlorine demand of tyrosine. Thus, the chlorine demand of tyrosine in the present study was lower than 13.40 and 13.20 mol-Cl₂/mol-Cp reported previously (Hureiki et al., 1994; Hong et al., 2009). There were no significant differences in chlorine demands between 72 and 96 h, indicating that tyrosine is depleted in less than 72 h. Figure 1 shows a comparison of the chlorine demands in the present study with those reported in the literature.
The high chlorine demand of aromatic amino acids is generally related to the structure because each has an aromatic ring with an alanyl substituent. The contribution of the alanyl substituent is expected to be around 3 mol-Cl$_2$/mol-Cp and the rest should be contributed by the aromatic ring through electrophilic substitution reaction. Tyrosine has a phenol group that activates the ring favoring electrophilic substitution reaction. The chlorine demand of the 4-hydroxyphenyl ring is expected to contribute about 9 mol-Cl$_2$/mol-Cp based on the results of previous studies for phenol or 4-hydroxytoluene (de Laat et al., 1982; Gallard and von Gunten, 2002). Similarly, tryptophan is expected to undergo electrophilic substitution reaction in the aromatic heterocyclic indole moiety, and its contribution is expected to be around 10 - 13 mol-Cl$_2$/mol-Cp depending on the contact time (de Laat et al., 1982; Hureiki et al., 1994; Hong et al., 2009).

The chlorine demand of histidine, which has imidazole and alanyl moieties, was 4.23 mol-Cl$_2$/mol-Cp. Our study also showed that alanine had a chlorine demand of 2.5 mol-Cl$_2$/mol-Cp. If the alanyl substituent in the histidine is taken into account, the imidazole moiety contributed only 1.3 mol-Cl$_2$/mol-Cp. This contribution is lower than expected from the electrophilic substitution in histidine, which has two endocyclic amines. Other studies indicated that the chlorine demands of histidine at 72 and 96 h were 12.0 and 9.9 mol-Cl$_2$/mol-Cp, respectively (de Laat et al., 1982; Hureiki et al., 1994; Hong et al., 2009). However, it is difficult to clarify the reason why the 96-h study had lower chlorine demand than the 72 h study. If the chlorine demands of alanine at 72 and 96 h were taken into account in each study, the imidazole moiety is expected to contribute about 6 and 9 mol-Cl$_2$/mol-Cp, respectively. This suggests that there is either an error in our data or that histidine undergoes electrophilic substitution very slowly in the first 24 h as studies with longer contact times indicated higher chlorine demands than observed in our experiments (Hureiki et al., 1994; Hong et al., 2009).

S-amino acids had intermediate chlorine demands, with values of 7.46 and 6.62 mol-Cl$_2$/mol-Cp for cysteine and methionine, respectively. These results are not markedly different from those reported in previous studies performed with contact times.
of 72 and 96 h (Hureiki et al., 1994; Hong et al., 2009). S-amino acids have two basic groups (sulfur and amine) at which chlorine substitution takes place. Aside from these two common amino acids, water may be contaminated with traces of cystine, thiamine, and biotin, all of which have sulfur in their structures. These contaminants and any other sulfur-containing organic molecules may contribute to higher chlorine demand in water treatment plants. Another group of amino acids with relatively high chlorine demand closer to those of the S-amino acids are arginine and asparagine. These amino acids possess amide and imine groups that can undergo chlorine substitution, and they have chlorine demands of 5.63 and 5.77 mol-Cl$_2$/mol-Cp, respectively. However, this study showed that chlorine demand for arginine was about 3 units lower than those reported in previous studies with longer reaction times (Hureiki et al., 1994; Hong et al., 2009). It was found from the structure that there are 3 amines and 1 imine groups in arginine. However, 2 amines and 1 imine that are connected to the carbon-6 may not be as basic as the α-amine group due to resonance stabilization in the former. Thus, the reactions of these groups in arginine (or its byproducts) with chlorine is slow and it may most likely be incomplete in the 24 h contact time used in this study. On the other hand, the chlorine demand of asparagine was comparable to or higher than those described in previous studies (Hureiki et al., 1994; Hong et al., 2009). Similarly, the chlorine demand of glutamine was similar to those in previous studies. Thus, the reactions of these compounds or their byproducts with chlorine were completed in less than 24 h.

The chlorine demand of asparagine was larger than that of glutamine. Although glutamine has an amide group, similar to asparagine, the two are structurally different since asparagine has one methylene bridging amine and amide groups, while glutamine has an ethylene group bridging the amine and amides. The difference in their chlorine demands was considered to be due to the difference in the methylene and ethylene groups. It should be noted that the pKa of the methylene group is lower than that of the ethylene group, but the difference is small.

Six amino acids have chlorine demands ranging between 3 and 4, i.e., proline, serine, threonine, lysine, glycine, and aspartic acid (Table 1). These values were generally lower than those reported in previous studies conducted with longer contact times (Hureiki et al., 1994; Hong et al., 2009). This suggests that these amino acids may continue reacting even in distribution systems, because it is expected that treated water may still be in the distribution system after 24 h. While glutamic acid showed a chlorine demand of 2.68 mol-Cl$_2$/mol-Cp, both leucine and isoleucine, which isomers, showed the same chlorine demand of 2.54 mol-Cl$_2$/mol-Cp (Table 1). These results for the latter are similar to those of the 72 h study (Hureiki et al., 1994) but lower by about 3 units compared to the 96-h study (Hong et al., 2009).

Phenylalanine and alanine consumed 2.82 and 2.54 mol-Cl$_2$/mol-Cp, respectively, and the results are not significantly different from the previous 15 h and 72 h studies (de Laat et al., 1982; Hureiki et al., 1994). However, Hong et al. (2009) reported chlorine demands of 3.90 and 5.1 mol-Cl$_2$/mol-Cp for alanine and phenylalanine, respectively, which are higher than those in our experiments and in the other two studies mentioned above. The chlorine demands were increased by 1 and 2 units from those in the 15 h and 72 h studies, respectively. These observations indicated that alanine and phenylalanine would have reacted with chlorine if the contact time was increased to 96 h. Although
phenylalanine is an aromatic amino acid, it is not as reactive as tyrosine. The difference is that phenylalanine lacks strong ring-activating substituents, such as OH and NH₂. Therefore, electrophilic substitution may occur very slowly in the phenyl ring, which is why the chlorine demand of phenylalanine is expected to be slightly higher than that of alanine and less than that of tyrosine (Hong et al., 2009).

Amino sugars are nitrogenous organic compounds that have amine groups, and can react with chlorine similar to amino acids. In this study, glucosamine, galactosamine, and mannosamine were tested for chlorine consumption under the same conditions as those used for the amino acids. The results indicated that the chlorine demands were very close to each other and were also similar to those of some amino acids despite the differences in their structures (Table 1). These results also indicated that high levels of amino sugars in water sources may contribute significantly to the total chlorine demand of water. Finally, all nitrogenous compounds may contain amine, amide, or imine groups in their structures. In the present study, ammonia was used as the simplest nitrogenous compound in water. Chloramines are formed after the reaction of ammonia and chlorine. In the case of the nitrogenous organic compounds used in the present study, N-chloramines were considered to be formed by chlorination. This was supported by the frontier electron density (FED) calculation, which showed that the nitrogen in the amine group had the highest electron density of all atoms in alanine (Chu et al., 2009). Thus, it was considered that the amine group was an important site for chlorine substitution in nitrogenous compounds, as discussed previously (Bull et al., 2006). However, chlorination contact time at water treatment plants and distribution systems may not be sufficient to allow complete reaction with some compounds in water. Although there have been few reports of amino acids in tap water, an unpleasant smell in tap water indicates the formation of odorous compounds (e.g., chloroaldimines). Thus, the presence of such odorous compounds was considered to indicate the importance of determining the chlorine demand under the conditions used for chlorination in actual distribution systems.

Prediction of chlorine demand
Further analysis of the data reported in this work has shown that there is a weak linear relationship ($R^2 = 0.265, F = 7.93, P = 0.01, N = 24$) between the molecular weight (MW) and chlorine demand at 95% confidence level (Fig. 2). The weak relationship may be attributed to the fact that only certain sites of the molecule (functional groups) are involved in chemical reaction with chlorine and not the whole molecule.

Thus, the number of functional groups and relative position in the molecule are highly important in chemical reaction. That is why two molecules with different molecular weights may have the same chlorine demands as shown by the following pairs of molecules: lysine vs. glucosamine, alanine vs. leucine or serine vs. proline (Table 1). Although the MW of tyrosine differs from that of amino sugars by 2 units, its chlorine demand was 2 times higher than that of amino sugars (Table 1). Thus, models based on structure of molecules may be more useful for predicting chlorine demands than those models that were derived using bulk water parameters (e.g., pH, turbidity, UV-absorption, temperature) or molecular properties (e.g., MW, molecular volume, molecular surface area).
We used the QSPR model for predicting the chlorine demand in drinking water (Luilo and Cabaniss, 2010) to estimate the chlorine demand of the amino acids and amino sugars studied. The results showed that all compounds, with the exception of arginine and glutamine were predicted within ±2.48 standard deviations of prediction, for 2% error margins for both sides (Fig. 3).

The standard deviation of prediction is ±2SDE, and in this case the SDE for QSPR calibration at 95% confidence interval was 1.24 mol-Cl₂/mol-Cp. Any predicted data

![Graph showing weak correlation between chlorine demands and molecular weights of the 23 model compounds.](image)

**Fig. 2 -** Weak correlation between chlorine demands and molecular weights of the 23 model compounds.

![Graph showing relationship between predicted and experimental chlorine demands for the 23 amino acids.](image)

**Fig. 3 -** Relationship between predicted and experimental chlorine demands for the 23 amino acids. The solid line is for the ideal model and dotted lines are standard deviation of prediction margins (±2.48 mol-Cl₂/mol-Cp).
point outside ± 2.48 margins implies that its prediction by the QSPR has some uncertainty and therefore, the predicted value may not be reliable.

To determine the reliability of the prediction, we used the coefficient of determination for cross-validation, $q^2$, as described previously (Luilo and Cabaniss, 2010). The value of $q^2$ was 0.74 which is much higher than the cut off value of 0.5 (Golbraikh and Tropsha, 2002) and RMSE equal to 1.30 mol-Cl₂/mol-Cp is closer to the model's standard error of 1.24 mol-Cl₂/mol-Cp. However, the MBD of 11.9% indicates that the model predicted 15 out of 24 compounds, which is slightly higher than expected. Although chlorine demand prediction data for tryptophan and histidine were close to the observed data, the chlorine demand prediction for histidine may not be as reliable as that for tryptophan. This is because the QSPR calibration data did not include any molecules with endocyclic heteroatoms such as nitrogen (e.g., imidazole, indole, pyrrole). Thus, one limitation of QSPR is that it cannot give reliable prediction of chlorine demands of aromatic molecules without exocyclic ring activators such as histidine because RAI descriptor is zero. However, RAI may be calculated in tryptophan because the amine group, which is exocyclic to the benzene ring (pyrrolyl moiety) may induce electrophilic substitution reaction in the benzene ring through resonance. However, the model cannot estimate the contribution of chlorine demand from the chlorine substitution at the pyrrolyl ring which has the endocyclic amine group.

The predictive power of the QSPR on model compounds is tested by generating regression line in a plot of predicted chlorine demand against the observed chlorine demand with and without intercept (Fig. 4). A good prediction should not give significantly different slopes that are closer to 1 ($0.85 \leq k \leq 1.15$); y-intercept, b, should

![Fig. 4 - Normal regression of predicted chlorine demand against the experimental chlorine demand with y-intercept (solid line) and without y-intercept (dotted line) for 23 model compounds.](image-url)
be closer to zero ($b \approx 0$); and the ratio \((R_i^2 - R_o^2)/R_i^2 < 0.1\) (Golbraikh and Tropsha, 2002; Tropsha et al., 2003) where \(R_i^2\) is obtained using regression line with intercept and \(R_o^2\) is obtained from regression line through origin. In Fig. 4, \(k_i = 0.998; k_o = 1.092; R_i^2 = 0.82; R_o^2 = 0.81\) and the ratio \((R_i^2 - R_o^2) / R_i^2\) is 0.01, which imply that the predictive strength of the model is high.

It is interesting that arginine has three amines and one imine. Despite having the greatest number of reduced nitrogens (amine and imine) in the molecule, its chlorine demand (5.63 mol-Cl\(_2\)/mol-Cp) was comparable to those of asparagine (2 amine groups) and amino sugars (1 amine group). It is possible that if the contact time used in the present study was extended to 72 h or more, arginine would have consumed more chlorine. This is because chlorination of arginine for 72 h and 96 h indicated chlorine demands of 8.20 and 8.90 mol-Cl\(_2\)/mol-Cp, respectively (Hureiki et al., 1994; Hong et al., 2009), which are close to the predicted value of 9.34 mol-Cl\(_2\)/mol-Cp. These observations imply that arginine reacts slowly with chlorine and therefore, requires a longer contact time for depletion. Thus, arginine is likely to continue reacting with residual chlorine in the distribution system.

CONCLUSIONS AND RECOMMENDATIONS
Amino acids and amino sugars are included among the dissolved organic matter that must be removed from raw water. However, conventional water treatment processes cannot remove all fractions of nitrogenous compounds, and are therefore likely to contribute to the chlorine demand of water. This study showed that the aromatic amino acids consumed 10 - 13 mol-Cl\(_2\)/mol-Cp, whereas the chlorine consumption of S-amino acids was between 6 and 8 mol-Cl\(_2\)/mol-Cp. The remaining amino acids and the three amino sugars showed chlorine consumption ranging between 2 and 6 mol-Cl\(_2\)/mol-Cp. The findings of our 24-h study were compared to those of the previous studies with contact times of 72 h and 96 h, and the results indicated an increasing trend in chlorine demands with time for some of the nitrogenous compounds. The QSPR predictions showed that arginine was overpredicted, while the rest were within the model prediction error.

The results of the present study indicated that chlorine demand in 24 h showed a good correlation with prediction data. Furthermore, kinetic studies are required to determine how fast the precursors can react with chlorine at typical water treatment contact times and chlorine doses. The data will provide insight into the amounts of chlorine demand and production of disinfection byproducts, especially nitrogenous compounds related to odor, while the results of longer contact times, such as 72 h and 96 h, might be useful to drive the reaction to completion.

ACKNOWLEDGMENTS
The authors would like to express their sincere gratitude to Mr. Takafumi Ito (Department of Waterworks, Kushiro, Japan) and Ms. Kyoko Suzuki (Yokohama Waterworks Bureau, Yokohama, Japan) for their experimental support. This work was partly supported by a Grant-in-aid for Health, Labor, and Welfare Research.
REFERENCES


Hong H. C., Wong M. H. and Liang Y. (2009) Amino acids as precursors of


