Supplementary Materials

1. Analytical solution: the equations for calculating the propagated errors in molecular indicators.
2. Errors in primary measurements and propagated errors in molecular indicators
3. Table S1. Statistical parameters used for calculating the propagated error in DI in this study.
4. Table S2. Errors (standard deviation) and variance in the primary measurements in this study. For amino acids, n = 3, and for DOC, n = 8. Note that the standard deviation is regarded as an error, and variance is a statistical parameter which is used to calculate the propagated error. The total standard deviation of all amino acids in this table is 19.8 nM
5. Table S3. Propagated errors (as transferred standard deviation, in bold) in the molecular indicators in this study. The upper part is the absolute value, and the bottom part is the proportion of propagated error in the raw molecular indicator values.

1. Analytical solution: the equations (as a standard deviation) for the calculating the propagated errors.

For the D/L ratio, the propagated error (p.e.) can be calculated as:

$$p.e. = \sqrt{\frac{s_{11}^2}{\bar{x}_2^2} + \frac{\bar{x}_1^2 \times s_{22}^2}{\bar{x}_2^4}} \quad (1)$$

where $\bar{x}_1$, $\bar{x}_2$, s_{11}^2, and s_{22}^2 are the concentration of the L form amino acid, the concentration of the D form amino acid, the variance of the L form amino acid, and the variance of the D form amino acid, respectively.

For Gly%, GABA%, or any other amino acid mole percentage:

$$p.e. = \sqrt{\frac{\bar{x}_2^2 \times s_{11}^2 + \bar{x}_1^2 \times s_{22}^2}{(\bar{x}_1 + \bar{x}_2)^4}} \quad (2)$$

where $\bar{x}_1$, $\bar{x}_2$, s_{11}^2, and s_{22}^2 are the concentration of Gly (or GABA), the sum of amino acid concentrations excluding Gly (or GABA), the variance of Gly (or GABA), and the sum of the variance of amino acids excluding Gly (or GABA), respectively.

For the AA C yield:

$$p.e. = \sqrt{\frac{\bar{x}_m^2 \times \sum_{j=1}^{m-1} c_j^2 \times s_{jj}^2 + s_{mm}^2 \times (\sum_{j=1}^{m-1} c_j \times \bar{x}_j)^2}{(\sum_{j=1}^{m-1} c_j \times \bar{x}_j + \bar{x}_m)^4}} \quad (3)$$

where $\bar{x}_m$ is the amount of non-amino acid carbon in the bulk dissolved organic carbon (i.e., [DOC] – $[DOC]_{amino\ acids}$), c_j is the number of carbon atoms in amino acid j (i.e., for glycine, $C_2H_5NO_2$, c = 2), s_{jj}^2 is the variance of amino acid j, s_{mm}^2 is the variance of non-amino acid carbon in DOC (i.e., the variance of DOC subtracted by $\sum_{j=1}^{m-1} c_j^2 \times s_{jj}^2$), and $\sum_{j=1}^{m-1} c_j \times \bar{x}_j$ is the amino acid carbon concentration ($[DOC]_{amino\ acids}$). In this study, we detected 33 amino acid compounds (L and D form amino acids were taken as different compounds), and hence, here, m equals 33 plus 1 (34). Note that all concentrations and variance should have the same unit (e.g., nM for concentration and nM^2 for variance).

The DI was first proposed by Dauwe and Middelburg (1998) for use in particulate/sediment studies. Our focus was on the dissolved form, so we followed the DI calculation presented later by Kaiser and Benner (2009), but with slight modifications. The amino acid histidine is not routinely measured in our laboratory, while Lys is measured but with interference from other peaks. Therefore, these two amino acids (histidine and Lys) were not considered in the DI calculation. Hence, the slight modification from the Kaiser and Benner DI method is that we used 12 of the 14 amino acids originally used by Kaiser and Benner. Under this condition, the propagated error in DI can be calculated by (m = 12):

$$p.e. = \sqrt{\sum_{j=1}^{m} s_{jj}^2 \times (\frac{\gamma_j \times \sum_{k=1}^{m} \bar{x}_k - \sum_{k=1}^{m} \gamma_k \times \bar{x}_k}{(\sum_{k=1}^{m} \bar{x}_k)^2})^2} \quad (4)$$

where s_{jj}^2 is the variance of amino acid j among the 12 amino acids used for DI calculation (a combination of the variances of both L and D forms of amino acid j), $\sum_{k=1}^{m} \bar{x}_k$ is the sum of the 12 amino acid concentrations, and γ_j is a constant of amino acid j, which was calculated as the fac. divided by STD, using the terms in the original DI coefficient table provided by Kaiser and Benner (2009). The s_{jj}^2 and γ_j values for the 12 amino acids are listed in the supplemental materials (Table S1). Note that when Equation 4 is applied

to the DI using 14 amino acids (or any other number of amino acids), the values of m should be changed accordingly.

2. Precision in primary measurements and propagated errors in molecular indicators

The precision of measurements is strongly related to the concentration itself and to the parameters of laboratory processing (e.g., injection volume, instrument status). Precision is usually poorer when samples at lower concentrations are measured because interference increases from the noise or the baseline. The most appropriate approach to obtaining precision is to measure samples that are at similar concentrations as the study is proceeding. In this study, we used the deep seawater standard and SCS slope region surface seawater to determine the precision of the measurements of DOC and AAs, respectively. The selected DOC and AA concentrations were lower than (for DOC) or similar to (for AAs) those of the collected samples. Therefore, the precision used here was either sufficient (DOC) or appropriate (AAs).

Unlike precision, propagated errors cannot be determined by instrumental measurements. At a given precision, this is a mathematical (or statistical) question. There are usually two ways of calculating propagated errors: one is a numerical solution and the other is an analytical (theoretical) solution. For the analytical solution, the exact propagated error expression must first be derived (e.g., as an equation) via statistical inference, and the propagated error can then be calculated. The advantages of this method are that the propagated error generation mechanism is theoretically clear and it is sometimes more efficient than the numerical solution in determining the final propagated error. However, the disadvantages of this method are also very clear: it requires a good background in mathematics and/or statistics, and the equation can only be applied to a specific scenario. Another round of statistical inference is required when a new scenario arises. Numerical solutions (e.g., Monte Carlo simulation) are widely used in estimating propagated errors in various chemical oceanographic studies (e.g., Krause-Jensen and Duarte, 2016; Munro et al., 2013; Shankle et al., 2002), whereas the application of analytical solutions is very limited (Bronk et al., 2000). As far as we know, our work is the first study to use the analytical solution to calculate the propagated errors in AA-based molecular indicators (AA carbon yield, AA%, D/L ratio, and DI) (Kaufman, 2003; McCarthy et al., 2007).

Error in the primary measurements (precision).

For amino acids, the standard deviation of all detected amino acids ranged from 0.2 to 2.4 nM. Asparagine and glutamine were not detected due to acid-hydrolyzation, and hence no data was available. The mean TDAA concentration in the precision-test samples was 382 nM. For DOC, the error was 1.0 μM. All the errors in primary measurements are shown as standard deviations in Table S2. The mean DOC concentration in the precision-test samples was 42.1 μM.

Error in the molecular indicators (propagated error).

Equations 1–4 are the exact expressions used for the calculation of propagated errors for each molecular indicator. The errors in molecular indicators were all functions of their corresponding primary variables (e.g., for the error in the D/L ratio Ala, it was a function of both the D-Ala and L-Ala concentration). The propagated errors for all molecular indicators in this study were calculated, based on the observed amino acid and DOC data (Table S3). The average propagated error in the D/L ratio of Ala was 0.02 for autumn and 0.01 for spring. For Gly%, the propagated errors were 0.5% (autumn) and 0.4% (spring). The propagated errors in the AA C yield ranged from 0.03 to 0.08%, and the mean propagated errors were 0.04% (autumn) and 0.06% (spring), respectively. The DI values had a mean propagated error of 0.2 for both seasons, but the maximum was 0.6 (Table S3). When compared with the molecular indicator value itself, the proportional propagated error (%) was most significant for GABA. The proportional propagated error reached 126% in autumn, whereas in spring it was 66% (Table S3). For DI, the proportional propagated error relative to the DI value itself was also very large, reaching 92% (autumn), with a mean

proportion of 30% (autumn) and 13% (spring), respectively. With respect to Gly%, AA C yield, and the D/L ratio of Ala, the proportional propagated errors were all <10% (Table S3).

3. Table S1. Statistical parameters used for calculating propagated error in DI in this study.

	s_{jj}^2	γ
Ala	0.834	−0.0792
Arg	0.620	0.1293
Asp	1.060	−0.0192
Glu	0.396	0.0273
Gly	0.524	−0.0472
ILE	0.075	0.2033
leu	6.031	0.2184
Phe	0.588	0.1627
Ser	0.274	−0.0034
Thr	0.796	0.1409
Tyr	0.160	0.2262
Val	0.380	0.1607

4. Table S2. Errors (standard deviation) and variance in the primary measurements in this study. For amino acids, n = 3, and for DOC, n= 8. Note that the standard deviation is regarded as an error, and variance is a statistical parameter which is used to calculate the propagated error. The total standard deviation of all amino acids in this table is 19.8 nM.

		standard deviation	variance
L-Asp	nM	0.9	0.881
D-Asp	nM	0.4	0.179
L-Glu	nM	0.5	0.298
D-Glu	nM	0.3	0.097
L-Asn	nM	nd*	nd*
D-Asn	nM	nd*	nd*
L-Ser	nM	0.4	0.194
D-Ser	nM	0.3	0.080
L-Gln	nM	nd*	nd*
D-Gln	nM	nd*	nd*
L-Thr	nM	0.7	0.550
Gly	nM	0.7	0.524
D-Thr	nM	0.5	0.246
L-Arg	nM	0.5	0.260
D-Arg	nM	0.6	0.360
L-Ala	nM	0.8	0.689
GABA	nM	1	1.218
D-Ala	nM	0.4	0.145
L-Tyr	nM	0.4	0.160
L-Aba	nM	0.8	0.564
L-Val	nM	0.4	0.131
L-Met	nM	0.4	0.201
L-Trp	nM	0.8	0.649
D-Met	nM	0.5	0.249
D-Val	nM	0.5	0.249
L-Phe	nM	0.5	0.213
L-ILE	nM	0.2	0.040
D-Phe	nM	0.6	0.375
L-Leu	nM	0.5	0.261
D-ILe	nM	0.2	0.035
D-Leu	nM	2	5.770
L-Lys	nM	2	2.310
D-Lys	nM	2	3.487
bulk DOC	μM	1	1.038

*: no data.

5. Table S3. Propagated errors (as transferred standard deviation, in bold) in the molecular indicators in this study. The upper part is the absolute value, and the bottom part is the proportion of propagated error in the raw molecular indicator values.

		D/L alanine	Gly%	Thr%	Ser%	GABA%	AA% C yield	DI
absolute value								
autumn	mean	**0.02**	**0.50%**	**0.30%**	**0.26%**	**0.40%**	**0.04%**	**0.2**
	min	0.01	0.30%	0.20%	0.18%	0.30%	0.03%	0.1
	max	0.05	1.60%	0.81%	0.61%	1.00%	0.05%	0.6
spring	mean	**0.01**	**0.40%**	**0.21%**	**0.23%**	**0.30%**	**0.06%**	**0.2**
	min	0.01	0.20%	0.14%	0.14%	0.20%	0.04%	0.1
	max	0.02	0.70%	0.38%	0.40%	0.50%	0.08%	0.3
proportion								
autumn	mean	5%	1.60%	8.3%	2.0%	43%	3%	30%
	min	3%	1.10%	5.1%	1.3%	16%	2%	18%
	max	9%	4.30%	24%	5.9%	126%	8%	92%
spring	mean	4%	1.10%	5.4%	1.2%	33%	3%	13%
	min	2%	0.70%	2.5%	0.8%	17%	2%	8%
	max	6%	2.00%	12.4%	2.2%	66%	4%	23%