Correlation between Yield and Reduced Mass of Raw Materials in Intermolecular and Intramolecular Cyclization Reactions

Masatoshi Kawashima

Laboratory Section, Azuma Co., Ltd.
1-6-13 Tamasaki-Nishi, Ichihara, Chiba 290-0044

E-mail: kawashima@azuma-g.co.jp

Correlation between yield and reduced mass of raw materials in intramolecular C-N, C-C, C-O coupling reactions and intramolecular/intermolecular Diels-Alder reactions was revealed. The regression equation was found to be the same as that of intermolecular reactions; $yield = -0.1861 \frac{M_{AB}}{n_A n_B n_I} + 100.0$, where M_{AB} is the reduced mass per mole and n_A and n_B are the total number of each reaction site in molecular A and molecular B in the reaction system, and n_I is the number to distinguish whether it is a intramolecular reaction or intermolecular reaction.

Keywords: Yield |Reduced Mass |Correlation |Prediction

Previous report by the author has shown the correlation between yield and reduced mass of raw materials in coupling reaction.[1] The correlation is expressed by the following equation (1); $yield = -a \frac{M_{AB}}{n} + b$, where a and b are coefficients, M_{AB} is the reduced mass per mole and n is the number of reaction sites, and holds in intermolecular coupling reactions without consideration of leaving groups, catalysts, ligands, additives, solvents, reaction temperature, reaction time. Correlation in cyclization reactions has not been clarified. This paper reports on correlation between yield and reduced mass in intermolecular and intramolecular cyclization.

Although reduced mass is calculated from two raw materials in intermolecular reaction, raw material in intramolecular cyclization is only one. Moreover, there are many types of cyclization due to the number of reaction sites and the type of functional group of the reaction site. Therefore, to include intermolecular/intramolecular cyclization reactions in the scope of application of the regression equation, the n in the equation (1) was replaced by $n_A n_B n_I$, where n_A and n_B are the total number of each reaction site in molecular A and molecular B, respectively, and n_I is the number to distinguish whether it is a intramolecular reaction or intermolecular reaction. Provisionally, it was set to 2 in the case of an intramolecular reaction, and was set to 1 in the case of an intermolecular reaction. M_{AB} was tentatively calculated from the molecular weight (MW) of the same molecule in the case of an intramolecular reaction. The temporary correlation formula for intermolecular/intramolecular reactions is shown in the equation (2); $yield = -a \frac{M_{AB}}{n_A n_B n_I} + b$. All examples, that meet all of the following criteria, were selected for the regression analysis of the reaction in the selected literature; 1) reactions with cyclization 2) the structural formula can be identified, 3) the yield is specified, 4) the highest yield is adopted among the syntheses from the same raw materials, and 5) the highest yield is adopted among the H-NMR-, GC-, LC-yield, and isolated yield. It did not take into account

differences in leaving groups, catalysts, ligands, additives, solvents, reaction temperatures, or reaction times.

First, C-N coupling reaction was selected as a example. The synthesis of compound **8** using Pa-catalyzed intermolecular/intramolecular C-N coupling schematically consists of six steps, as depicted in Figure 1.[2]
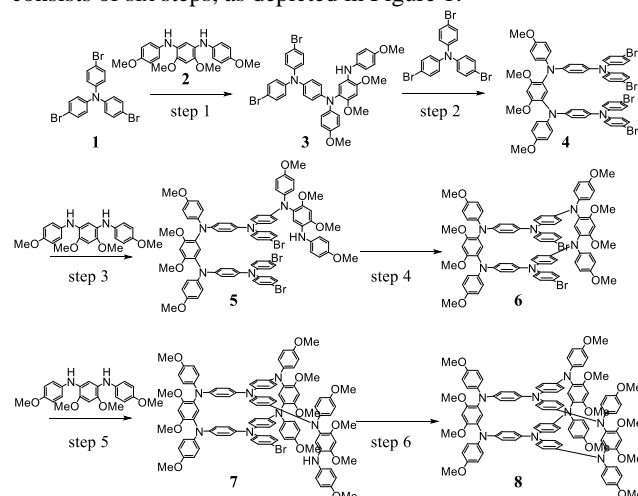
Figure 1. An example in the samples of Pa-catalyzed intermolecular and intramolecular C-N coupling.

Therewith, the analysis was conducted as presented in Table 1. The literature and predicted yield, calculated by the equation (2) using the coefficients of the value listed in the previous report, well match.[1] It turned out that there is not major problem with the provisional values and tentative calculation methods so far. The results of the statistical processing described later, along with other data, are presented in Table 8(a).

Table 1. Analysis of the example of Figure 1.

step	$M_A{}^a$	$M_B{}^b$	M_{AB}	M_{AB}/n	$n_A{}^c$	$n_B{}^d$	$n_I{}^e$	n^f	$Y_{lit.}$ /%^{g}	$Y_{int.}$ /%^{h}	$Y_{pred.}$ /%^{i}
1	482	380	213	35.4	3	2	1	6	97		
2	782	482	298	59.6	5	1	1	5	92		
3	1,183	380	288	36.0	4	2	1	8	97		
4	1,482	1,482	741	123.5	3	1	2	6	79		
5	1,401	380	299	74.8	2	2	1	4	89		
6	1,701	1,701	850	425.2	1	1	2	2	8	17	10

$^a M_A$: MW of each bromide. $^b M_B$: MW of each amine. ^{c}the total number of Br atoms in each raw material. ^{d}the total number of NH groups in each raw material. ^{e}intermolecular reaction=1, intramolecular reaction=2. $^f n = n_A n_B n_I$. ^{g}literature yield of **8** from **1**. ^{h}predicted yield in each step. ^{i}predicted yield of **8** from **1**.

Second, C-C coupling was selected to analyze. The synthesis of compound **11** using Pd-catalyzed direct arylation and cyclization schematically consists of two steps, as displayed in Figure 2, although the route is mechanistically denied in the literature.[3] The analysis was carried out as .
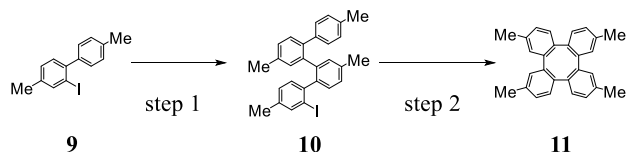
Figure 2. An example in the samples of Pd-catalyzed direct arylation and cyclization.
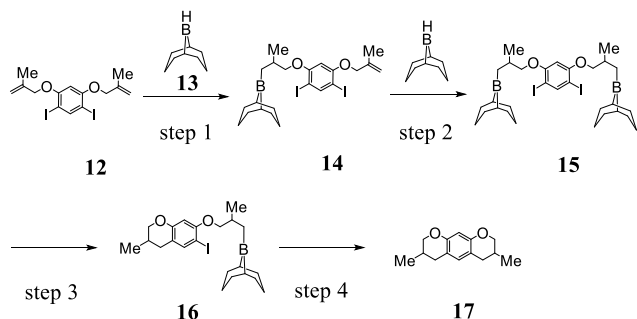
shown in Table 2. The literature and predicted yield well match also in this case. It is indicated that the exact reaction mechanism is not particularly important to this regression analysis. The results of the statistical processing described later, along with other data, are presented in Table 8(a).

Table 2. Analysis of the example of Figure 2.

step	$M_A{}^a$	$M_B{}^b$	M_{AB}	M_{AB}/n	$n_A{}^c$	$n_B{}^d$	$n_I{}^e$	n^f	$Y_{lit.}$ /%[g]	$Y_{int.}$ /%[h]	$Y_{pred.}$ /%[i]
1	308	308	154	38.5	2	2	1	4		96	
2	488	488	244	122.1	1	1	2	2	72	79	76

[a]M_A: MW of each iodide. [b]$M_B=M_A$. [c]the total number of I atoms in the formula of each raw material. [d]the total number of CH groups can participate in the reaction, in the formula of each raw material. [e]intermolecular reaction=1, intramolecular reaction=2. [f]$n=n_A n_B n_I$. [g]literature yield of **11** from **9**. [h]predicted yield in each step. [i]predicted yield of **11** from **9**.

Third, Suzuki-Miyaura coupling was selected as a example for analysis. The synthesis of compound **17** schematically consists of four steps, as depicted in Figure 3, which are two steps of hydroboration reaction and two steps of intramolecular Suzuki-Miyaura coupling.[4] The analysis

Figure 3. An example in the samples of intramolecular Suzuki-Miyaura coupling.
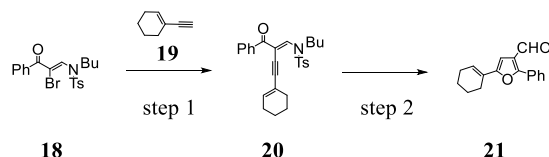
was conducted as shown in Table 3. The literature and predicted yield well match also in this case. The results of the statistical processing described later, along with other data, are presented in Table 8(a).

Fourth, the cascade reaction of Sonogashira coupling and the subsequent intramolecular cyclization was selected. The synthesis of compound **21** consists of two steps as

Table 3. Analysis of the example of Figure 3.

step	$M_A{}^a$	$M_B{}^b$	M_{AB}	M_{AB}/n	$n_A{}^c$	$n_B{}^d$	$n_I{}^e$	n^f	$Y_{lit.}$ /%[g]	$Y_{int.}$ /%[h]	$Y_{pred.}$ /%[i]
1	470	122	97	24.2	2	2	1	4		99	
2	592	122	101	101.2	1	1	1	1		83	
3	714	714	357	44.6	2	2	2	8		95	
4	466	466	233	116.6	1	1	2	2	68	80	63

[a]M_A: MW of each olefin or each iodide. [b]M_B: MW of each borane. [c]the total number of C=C double bonds or I atoms in the formula of each raw material. [d]the total number of B atoms in the formula of each raw material. [e]intermolecular reaction=1, intramolecular reaction=2. [f]$n=n_A n_B n_I$. [g]literature yield of **17** from **12**. [h]predicted yield in each step. [i]predicted yield of **17** from **12**.

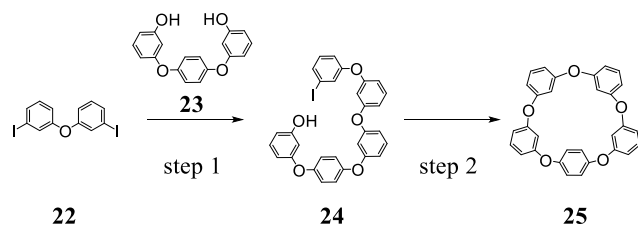**Figure 4.** An example in the samples of Pd/Cu-catalyzed Sonogashira coupling/cyclization.

displayed in Figure 4.[5] The analysis was conducted as indicated in Table 4. The literature and predicted yield very well match. The results of the statistical processing described later, along with other data, are presented in Table 8(a).

Table 4. Analysis of the example of Figure 4.

step	$M_A{}^a$	$M_B{}^b$	M_{AB}	M_{AB}/n	$n_A{}^c$	$n_B{}^d$	$n_I{}^e$	n^f	$Y_{lit.}$ /%[g]	$Y_{int.}$ /%[h]	$Y_{pred.}$ /%[i]
1	436	106	85	85.4	1	1	1	1		87	
2	462	462	231	115.4	1	1	2	2	69	81	70

[a]M_A: MW of bromide or acetylenic compound. [b]M_B: MW of acetylenic compound or N-tosyl enamine. [c]the total number of Br atoms or C≡C triple bonds in the formula of each raw material. [d]the total number of C≡C triple bonds or C=C-N bonds in the formula of each raw material. [e]intermolecular reaction=1, intramolecular reaction=2. [f]$n=n_A n_B n_I$. [g]literature yield of **21** from **18**. [h]predicted yield in each step. [i]predicted yield of **21** from **18**.

Fifth, C-O coupling reaction for a macrocyclic aryl ether was selected to analyze. The synthesis of compound 25 using Cu/Fe-catalyzed C-O coupling consists of two steps as depicted in Figure 5.[6] The analysis was conducted as

Figure 5. An example in the samples of Cu/Fe-catalyzed C-O coupling for a macrocyclic aryl ether.

indicated in Table 5. Although the literature and predicted yield relatively match, the correlation is lower than other

reactions described above. On the basis of the examination of reaction conditions in the literature, low solubility of metal phenolate is thought to be a cause. The results of the statistical processing described later, along with other data, are presented in Table 8(a).

Table 5. Analysis of the example of Figure 5.

step	$M_A{}^a$	$M_B{}^b$	M_{AB}	M_{AB}/n	$n_A{}^c$	$n_B{}^d$	$n_I{}^e$	n^f	$Y_{lit.}$ /%^{g}	$Y_{int.}$ /%^{h}	$Y_{pred.}$ /%^{i}
1	422	294	173	43.3	2	2	1	4		95	
2	588	588	294	147.1	1	1	2	2	54	74	71

$^a M_A$: *MW* of each iodide. $^b M_B$: *MW* of each phenolic compound. ^{c}the total number of I atoms in the formula of each raw material. ^{d}the total number of hydroxy groups in the formula of each raw material. ^{e}intermolecular reaction=1, intramolecular reaction=2. $^f n=n_A n_B n_I$. ^{g}literature yield of **25** from **22**. ^{h}predicted yield in each step. ^{i}predicted yield of **25** from **22**.

Sixth, Diels-Alder reaction was selected as a example for analyze. The synthesis of compound **28** using intermolecular Diels-Alder reaction consists of only one step as depicted in Figure 6(a).[7] The analysis was carried out as
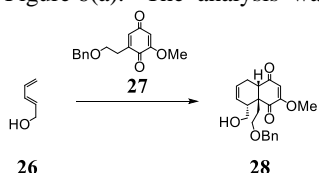
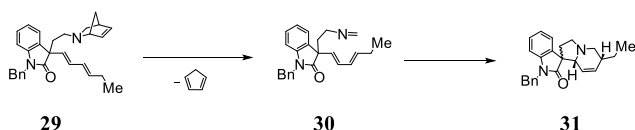**Figure 6(a).** An example in the samples of intermolecular Diels-Alder reaction.

reflected in Table 6(a). Since Diels-Alder reaction is energetically concerted reaction, both n_A and n_B were tentatively set to 1.[8] The literature and predicted yield very well match.

Table 6(a). Analysis of the example of Figure 6(a).

step	$M_A{}^a$	$M_B{}^b$	M_{AB}	M_{AB}/n	$n_A{}^c$	$n_B{}^d$	$n_I{}^e$	n^f	$Y_{lit.}$ /%^{g}	$Y_{pred.}$ /%^{h}
1	84.1	272	64.3	64.3	1	1	1	1	91	91

$^a M_A$: *MW* of enophile. $^b M_B$: *MW* of dienophile. ^{c}the total number of I atoms in the formula of each raw material. ^{d}the total number of hydroxy groups in the formula of each raw material. ^{e}intermolecular reaction=1. $^f n=n_A n_B n_I$. ^{g}literature yield. ^{h}predicted yield.

Intramolecular Diels-Alder reaction shown in Figure 6(b) was also analyzed. The synthesis of compound **31** consists of *retro*-Diels-Alder reaction and intramolecular Diels-Alder reaction.[7] The analysis was conducted as

Figure 6(b). An example in the samples of *retro*-Diels-Alder reaction and intramolecular aza-Diels-Alder reaction.
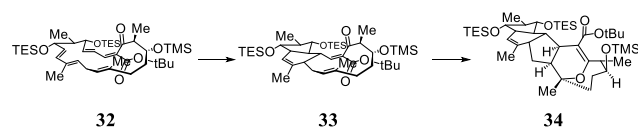
represented in Table 6(b). The literature and predicted yield relatively well match.

Table 6(b). Analysis of the example of Figure 6(b).

step	$M_A{}^a$	$M_B{}^b$	M_{AB}	M_{AB}/n	$n_A{}^c$	$n_B{}^d$	$n_I{}^e$	n^f	$Y_{lit.}$ /%^{g}	$Y_{int.}$ /%^{h}	$Y_{pred.}$ /%^{i}
1	425	425	212	106.1	1	1	2	2		82	
2	358	358	179	89.6	1	1	2	2	61	86	71

$^a M_A$: *MW* of each olefin. $^b M_B = M_A$. ^{c}concerted reaction=1. ^{d}concerted reaction=1. ^{e}intramolecular reaction=2. $^f n=n_A n_B n_I$. ^{g}literature yield of **31** from **29**. ^{h}predicted yield in each step. ^{i}predicted yield of **31** from **29**.

Transannular Diels-Alder reaction displayed in Figure 6(c) was also analyzed. The synthesis of **34** consists of transannular Diels-Alder reaction and transannular hetero Diels-Alder reaction.[7] The analysis was conducted as shown

Figure 6(c). An example in the samples of transannular Diels-Alder reaction and transannular hetero Diels-Alder reaction.

in Table 6(c). The literature and predicted yield very well match also in this case. The results of the statistical processing described later, along with other data, are presented in Table 8(a).

**Table 6(c).** Analysis of the example of Figure 6(c).

step	$M_A{}^a$	$M_B{}^b$	M_{AB}	M_{AB}/n	$n_A{}^c$	$n_B{}^d$	$n_I{}^e$	n^f	$Y_{lit.}$ /%^{g}	$Y_{int.}$ /%^{h}	$Y_{pred.}$ /%^{i}
1	775	775	388	193.8	1	1	2	2		65	
2	775	775	388	193.8	1	1	2	2	40	65	42

$^a M_A$: *MW* of each olefin. $^b M_B = M_A$. ^{c}concerted reaction=1. ^{d}concerted reaction=1. ^{e}intramolecular reaction=2. $^f n=n_A n_B n_I$. ^{g}literature yield of **34** from **32**. ^{h}predicted yield in each step. ^{i}predicted yield of **34** from **32**.

And Finally, outliers in all data of six types reactions described above were removed by Smirnov-Grubbs test as a significance level of 0.05. The coefficients a and b in the regression equation (2) were adjusted by using the least squares method to minimize the sample variance. The rewritten regression equation (3) is presented as follows; $yield = -0.1861 \frac{M_{AB}}{n_A n_B n_I} + 100.0$. The standard deviation (s) of the regression analysis of all data obtained by using the equation (3) became smaller from 0.13 to 0.12 and the regression coefficient was also improved from 1.05 to 1.00. Literature versus predicted yields plot of each type of reaction, obtained by linear regression through the origin, is depicted in Figure 7(a)-7(f) and Figure 8(a) and the regression coefficients and coefficients of determination are summarized in Table 8.

Data of intermolecular C-N, C-C, and C-O coupling reactions used in the previous report were reanalyzed with the regression equation (3) and the outliers were removed by Smirnov-Grubbs test as a significance level of 0.05.[1] The regression coefficient and coefficient of determination were improved. Literature versus predicted yields plot of the intermolecular C-N, C-C, and C-O coupling reactions is given in Figure 8(b). Namely, the regression equation (3) holds

regardless of whether it is an intermolecular reaction or an intramolecular reaction, and whether it is a cyclization reaction or not, as far as the author analyzed.
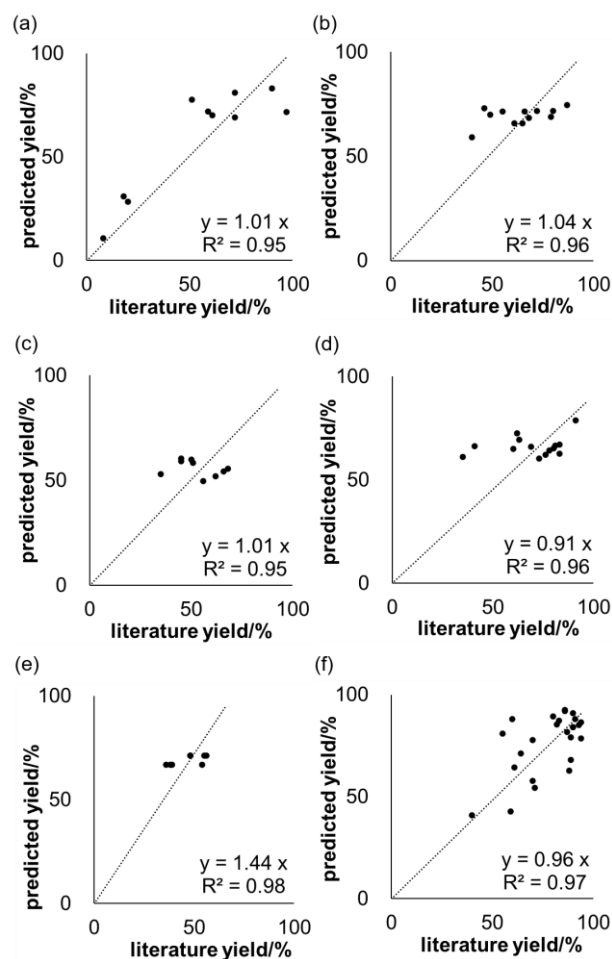
Figure 7. Literature versus predicted yields plot. (a) intramolecular C-N coupling. (b) direct arylation and cyclization. (c) intramolecular Suzuki-Miyaura coupling. (d) Sonogashira coupling/cyclization. (e) C-O coupling for macrocyclic aryl ether. (f) intramolecular and intramolecular Diels-Alder reaction.
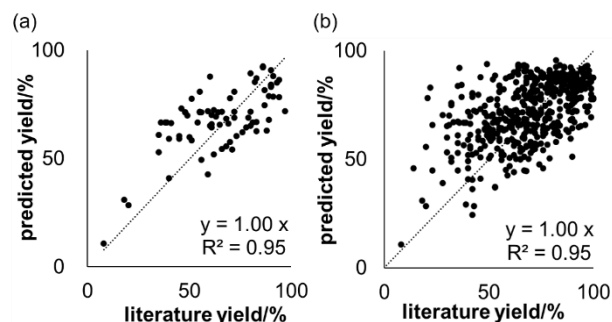
Figure 8. Literature versus predicted yields plot. (a) cyclization reactions included in this report. (b) intermolecular C-N, C-C, C-O coupling reactions in the previous report.[1]

Table 8. Linear regression through the origin.

type of reaction	N	β	R^2
Pa-catalyzed intramolecular C-N coupling	10	1.01	0.95
Pa-catalyzed direct arylation and cyclization,	12	1.04	0.96
Pa-catalyzed intramolecular Suzuki-Miyaura coupling	9	1.01	0.95
Pd/Cu-catalyzed Sonogashira coupling/cyclization	14	0.91	0.96
Cu/Fe-catalyzed C-O coupling for a macrocyclic aryl ether	8	1.44	0.98
intermolecular and intramolecular Diels-Alder reaction	25	0.96	0.97
all data	78	1.00	0.95

N: the sample size, β: regression coefficient, R^2: coefficient of determination

References

1   M. Kawashima, *ChemRxiv*. **2020**, Preprint. https://doi.org/ 10.26434/chemrxiv.12993218.v1

2   P. Ruiz-Castillo, S. L. Buchwald, *Chem. Rev.* **2016**, *116*, 12564; DOI: 10.1021/acs.chemrev.6b00512

3   C. Zhu, Y. Zhao, D. Wang, W.-Y. Sun, Z. Shi, *Sci. Rep.* **2016**, *6*, 33131; https://doi: 10.1038/srep33131

4   1) T. Willemse, W. Schepens, H. W. T. v. Vlijmen, B. U. W. Maes, S. Ballet, *Catalysts* **2017**, *7*, 74; https://doi:10.3390/catal7030074
      2) S. Ma, B. Ni, S. Lin, Z. Liang, *J. Organomet. Chem.* **2005**, *690*, 5389; https://doi.org/10.1016/j.jorganchem.2005.06.021

5   J. Yang, C. Wang, X. Xie, H. Li, E. Li, Y. Li, *Org. Biomol. Chem.* **2011**, *9*, 1342; https://doi.org/10.1039/C0OB00985G

6   Q. Zhou, L. Su, T. Jiang, B. Zhang, R. Chen, H. Jiang, Y. Ye, M. Zhu, D. Han, J. Shen, G. Dai, Z. Li, *Tetrahedron* **2014**, *70*, 1125; https://doi.org/10.1016/j.tet.2013.12.089

7   K. C. Nicolaou, S. A. Snyder, T. Montagnon, G. Vassilikogiannakis, *Angew. Chem. Int. Ed.* **2002**, *41*, 1668; https://doi.org/10.1002/1521-3773(20020517)41:10<1668::AID-ANIE1668>3.0.CO;2-Z

8   L. R. Domingo, E. Chamorro, P. Perez, *Org. Biomol. Chem.*, **2010**, *8*, 5495; https://doi.org/10.1039/C0OB00563K