Doctoral Dissertation (Censored)

博士論文 (要約)

Studies of a Cylindrical Host with Multiple Guests via Statistical Analyses of Stoichiometry

（統計手法を用いた化学量論解析による筒状ホスト分子による ゲスト分子包接の研究）

A Dissertation Submitted for the Degree of Doctor of Philosophy

December 2022

令和 4 年 12 月博士（理学）申請

Department of Chemistry, Graduate School of Science,

The University of Tokyo

東京大学大学院理学系研究科化学専攻

Kanato Takahashi

高橋 奏仁

Abstract

This thesis describes the studies of supramolecular complexes of cylindrical hosts with multiple guests via statistical analyses. Based on the statistical analyses, an information-theory approach of Akaike's information criterion (AIC) is introduced as a method to determine the stoichiometry of the supramolecular complexes.

Chapter 1 describes the general introduction about the stoichiometry determination of the supramolecular complexes. Conventionally, Job plot analysis has been known as a common method to determine stoichiometry. However, recent papers suggest caution on its applicability, and a statistical approach is suggested as an alternative method. This chapter also describes the emerging field of supramolecular complexes of cylindrical molecules, which allows for the encapsulation of multiple guests.

Chapter 2 describes the exploration of cooperativity in a cylindrical nanotube molecule, [4]cyclochrysenylene (**[4]CC**). The associations between **[4]CC** and halogenated corannulenes (X-**COR**s) are discussed. According to the precise titration experiments and statistical analysis using the *F*-test, the stoichiometry of the complexes was carefully determined to be 1:2. It was revealed that the cooperativity of the association changed with the size of the guest molecule.

Chapter 3 describes the introduction of AIC for the stoichiometry determination of host-guest complexes. It was found that the *F*-test also has a limited applicability in complicated supramolecular systems that encompass, for example, 1:2 and 2:1 models. Thus, AIC is first introduced to the stoichiometry determination of the supramolecular complexes in this study. Several test cases with cylindrical hosts are examined to demonstrate the usefulness of AIC.

Chapter 4 summarizes the achievements and outlook of this study.

Table of Contents

Abstract **2**

Chapter 1. General Introduction **7**

1.1 Determination of the stoichiometry 7

1.2 Host-guest chemistry of carbon nanotubes 11

1.3 This work 16

1.4 References 18

Chapter 2. Exploration of Cooperativity in Nanotubes **20**

2.1 Introduction 20

2.2 Confirmation of the formation of supramolecular complexes 23

2.3 Determination of stoichiometry by Job plot analysis 28

2.4 Determination of stoichiometry by *F*-test 29

2.5 Reanalysis of the H-COR complex 36

2.6 Crystal structures of the complexes 39

2.7 Thermodynamics and cooperativity of the systems 41

2.8 Conclusion 43

2.9 Experimental section 44

2.10 References 61

Chapter 3. Introduction of Akaike's Information Criterion after the Death of the Job Plot **64**

3.1 Introduction 64

3.2 Problem of stoichiometry determination 65

3.3 Dumbbell-and-cylinder systems 68

3.4 [4]CC+X-COR 73

3.5 [4]CC+C_{70} 75

3.6 [4]CC + pyrene 78

3.7 Conclusion 80

3.8 Experimental section 81

3.9 References 84

Chapter 4. Summary **88**

List of Publications **91**

Acknowledgments 92

List of abbreviations

[3]C^{db}C	[3]cyclodibenzochrysenylene
[4]CA	[4]cycloanthanthrenylene
[4]CC	[4]cyclochrysenylene
[4]CF	[4]cyclofulminenylene
χ^2	chi-squared value
α	interaction parameter
AIC	Akaike information criterion
CNT	carbon nanotube
COR	corannulene
cov_{fit}	the covariance of the fit
CPP	cycloparaphenylene
δ	chemical shift
*o*DCB	*ortho*-dichlorobenzene
df	degree of freedom
DWCNT	double-walled carbon nanotube
F	*F*-test value
FL	fluorescence spectroscopy
ΔG	molar free energy
GOF	goodness of fit
ΔH	molar enthalpy
ΔH_1	molar enthalpy of the first stage
ΔH_2	molar enthalpy of the second stage
ITC	isothermal titration calorimetry
k	number of parameters to be fitted
K	association constant
K_1	association constant of the first stage
K_2	association constant of the second stage
N	number of data
NMR	nuclear magnetic resonance
P	probability of *F* given the null hypothesis
R^2	coefficient of determination

ΔS	molar entropy
ΔS_1	molar entropy of the first stage
ΔS_2	molar entropy of the second stage
SS	sum of squared residuals
SWCNT	single-walled carbon nanotube
TEM	transmission electron microscopy
UV-vis	ultraviolet/visible light
w_i	Akaike weight of model *i*
x_{max}	apparent amount fraction maximum in Job plot
X-COR	halogenated corannulene

Chapter 1. General Introduction

1.1 Determination of the stoichiometry

Supramolecular chemistry has continued to develop since the 1960s with the pioneering studies of Pedersen, Lehn and Cram, who were awarded the Nobel prize in 1987.[1,2,3] Pedersen reported the first crown-ether supramolecular complex comprising dicyclohexyl-18-crown-6 and alkali metals in 1967 (Figure 1.1a).[4] The 1:1 complexation with LiI_3, $NaNO_2$, KI, KCNS, NH_4CNS, CN_3NH_3CN, RbCNS, CsCNS, CsCNS, $CaCl_2$, $Ba(OH)_2$, $CdCl_2$, $HgCl_2$ and $Pb(OAc)_2$ was confirmed in the crystalline solid state, but the exact stoichiometry of the solution state was elusive at the time because of the immaturity of the methodology. Subsequent work using calorimetric titration revealed that the association constant for one of the isomers of dicyclohexyl-18-crown-6 with K^+ was 100 M^{-1} in water.[5,6] Since then, a variety of host molecules have been synthesized for capturing specific guest molecules, which led to the discovery of unique host-guest interactions. This development allows for the investigation of complex systems, and one of the representative examples is squaring cooperative binding cycles by Rowan and coworkers with a porphyrin clip as a host, viologen as a guest and a pyridine ligand as an activator (Figure 1.1b).[7] In this system, bulky pyridine ligand A1 binds with Zn-porphyrin of the host externally, which influences the binding ability of the host with viologen guests. By fluorescence titration experiments, the full thermodynamic and kinetic circles concerning the activation of the porphyrin clip for the binding with viologen were measured and elucidated.

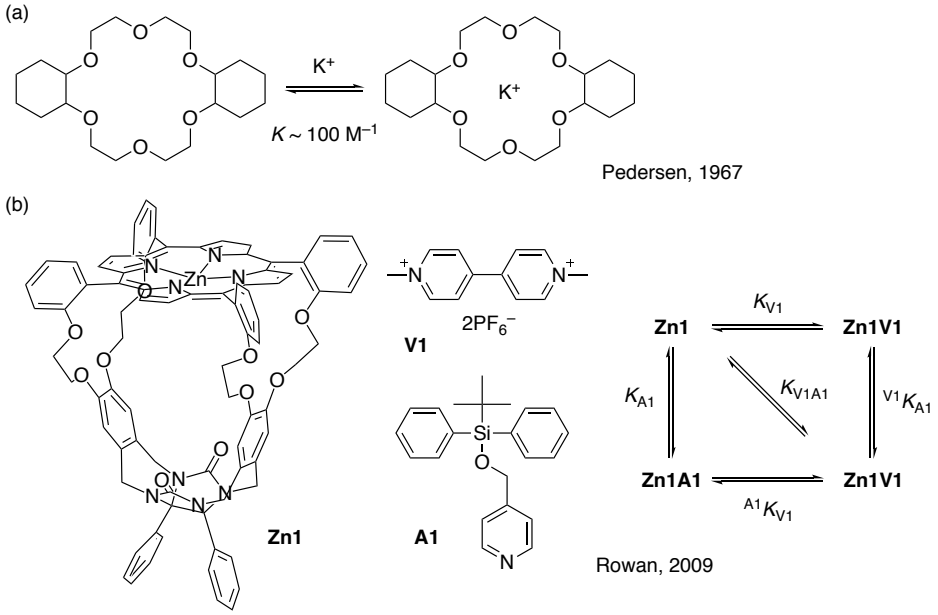


Figure 1.1. Pioneering works in supramolecular chemistry. (a) The first crown-ether supramolecular complex by Pedersen in ref. 4. (b) Representative examples of cooperative binding cycles by Rowan in ref. 7.

In supramolecular chemistry, one of the most important analyses is the quantitative measurement of the intermolecular interactions between host and guest molecules. The most common method for the quantitative analysis is the titration experiments. In titration experiments, one component (host/guest) is gradually added to the other component (guest/host) to monitor physical properties such as chemical shifts (in NMR spectroscopy) and absorption (in UV-vis absorption spectroscopy), which are sensitive to the supramolecular interactions. The change in the physical properties is then fitted to association models to obtain the parameters such as the association constants K, thermodynamic parameters (ΔG, ΔH and ΔS) and stoichiometry (1:1, 1:2, 2:1, etc.).

To obtain important thermodynamic parameters of interest, Determining the stoichiometry of complexes of host and guest is of particular importance (Figure 1.2). Depending upon the type of host and guest, a variety of complexation, including 1:1, 1:2 and 2:1 complexation, could be conceivable. To determine the stoichiometry, a Job plot is known as a common method employed in host-guest chemistry.[8,9] In the Job plot analysis, the change of the physical properties such as chemical shifts, UV-vis absorbance and fluorescence intensity are plotted versus molar ratio. The stoichiometry can be determined by the observed maximum in the Job plot. For example, the stoichiometry of

the association is indicated as 1:1 when the peak top of the plot x_{max} is at 0.5.

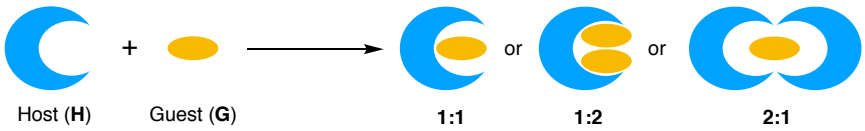
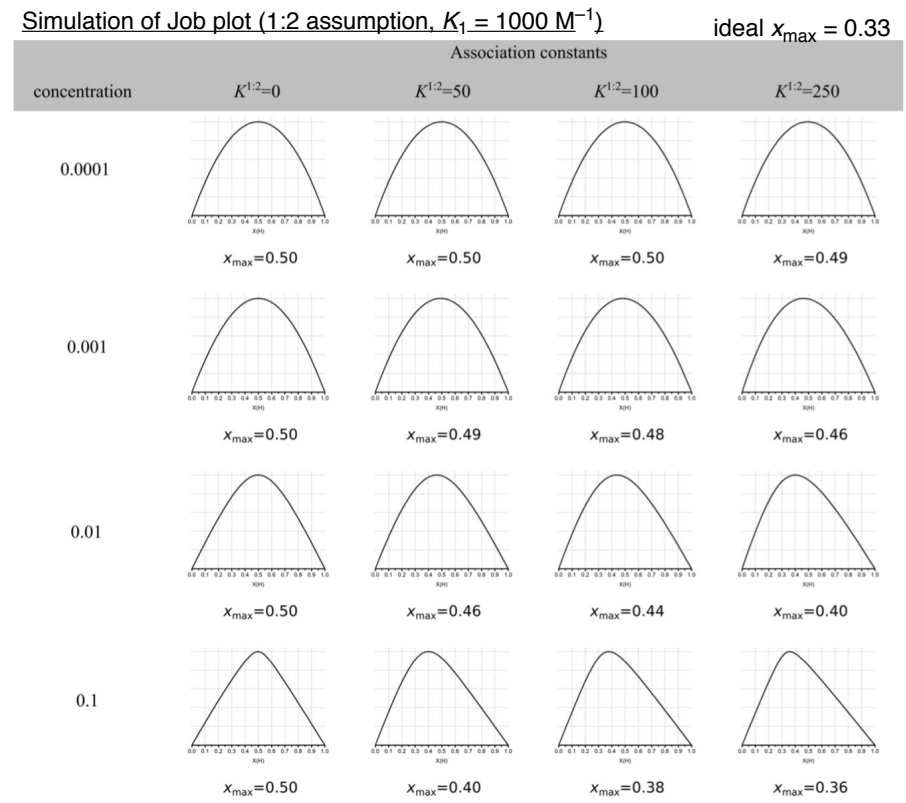


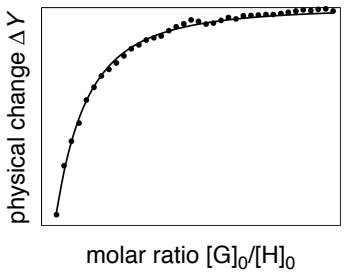
Figure 1.2. Example of stoichiometry of host-guest complexes. 1:1, 1:2 and 2:1 complexations are shown.

Although the Job plot analysis has been widely accepted as a common method for the determination of the stoichiometry of complexes, the limited applicability and "the death of Job plot" has been proclaimed by recent papers (Figure 1.3).[10,11] In these papers, it was proposed that Job plot does not necessarily give reliable results. Job plot is sensitive to host concentration and the ratio of K_1 to K_2. For example, when the 1:2 complex is unlikely to form due to the weak second association and the low concentration, the peak top of the Job plot x_{max} indicates 0.45~0.5 even with the 1:2 complexation. Misleading results are obtained because the expected x_{max} is 0.33 in the 1:2 association. Therefore, an alternative method to the Job plot was necessary to determine the stoichiometry, and statistical methods such as the *F*-test are proposed as an alternative in these papers.

(a)



(b)



"evaluation" : goodness of fit

χ^2 & cov_{fit}: smaller = better fit

"comparison" of two models

F-test & *P*-value:
F >> 1: "More complicated model" is suggested.
P < 0.05 : "More complicated model" is concluded.

Classic:
MCV → one fitting model → *K* determination

Modern (2016-):
multiple fitting models → evaluation → *K* determination
comparison

Figure 1.3. The death of Job plot. Adapted from ref. 10 (https://doi.org/10.1021/acs.joc.5b02909) according to the ACS guidance. This is an unofficial adaptation of an article that appeared in an ACS publication. ACS has not endorsed the content of this adaptation or the context of its use. Adapted from ref. 11 with permission from the Royal Society of Chemistry.

1.2 Host-guest chemistry of carbon nanotubes

Among the various artificial and natural host molecules examined, single-walled carbon nanotubes (SWCNTs) have also served as a unique host system in recent years (Figure 1.4a).[12] It was confirmed by transmission electron microscopy (TEM) that various guest molecules can be encapsulated inside SWCNTs (Figure 1.4b). In 1998, it was reported that multiple fullerene molecules were encapsulated and arranged in a one-dimensional fashion, which are called peapods.[13] Peapods have been widely studied because they have interesting features and can be easily prepared by placing CNTs and fullerenes together at 400-500 °C. Now, SWCNTs encapsulating higher fullerenes such as C_{70}, C_{76}, C_{78}, C_{80}, C_{82}, C_{84} and metallofullerenes $Sc_2@C_{84}$, $La@C_{82}$, $Sm@C_{82}$, $Gd@C_{82}$ can be synthesized in high yield. By encapsulating fullerenes, the electronic states of SWCNTs were significantly changed. For instance, the band gap of (11,9)-SWCNT was narrowed from 0.43 eV to 0.17 eV after the encapsulation of $Gd@C_{82}$, which was experimentally confirmed by scanning tunneling microscopy and spectroscopy (Figure 1.4c).[14] This phenomenon is caused by the strong interaction between the π-orbital of SWCNTs and fullerenes, thus showing the unique characteristics of SWCNTs to accommodate multiple fullerene molecules. Furthermore, in SWCNTs, novel reactions of guests, which do not usually proceed, were observed. The different reactivity of guests is attributed to an assumption that the π-electron cloud behaves as the catalyst. For example, double-walled carbon nanotubes (DWCNT) were obtained from a peapod comprised of SWCNTs and C_{60} upon heating at 1200 °C in vacuo.[15] Under such conditions, the C_{60} molecules inside the SWCNT coalesced to transform into a tubular structure.

Fullerenes are not the sole guest molecules to be accommodated inside SWCNTs. It was reported in 2011 that planar π-conjugated hydrocarbon, coronene molecules were encapsulated and stacked coaxially inside SWCNTs (Figure 1.4d).[16] Bowl-shaped compounds, corannulene and sumanene molecules can also be encapsulated (Figure 1.4e).[17,18] It was found that corannulenes were encapsulated with a random orientation, whereas sumanene molecules formed a stacked structure inside SWCNTs. As shown in these examples, SWCNTs can encapsulate various multiple molecules. Although SWCNTs serve as a unique system for host-guest chemistry, the studies remain phenomenological observations. This is largely because they are typically mixtures of various structures rather than molecular entities having a single definite structure. This

nature hampers the precise determination of thermodynamic parameters of the complexation, which is important for supramolecular chemistry.

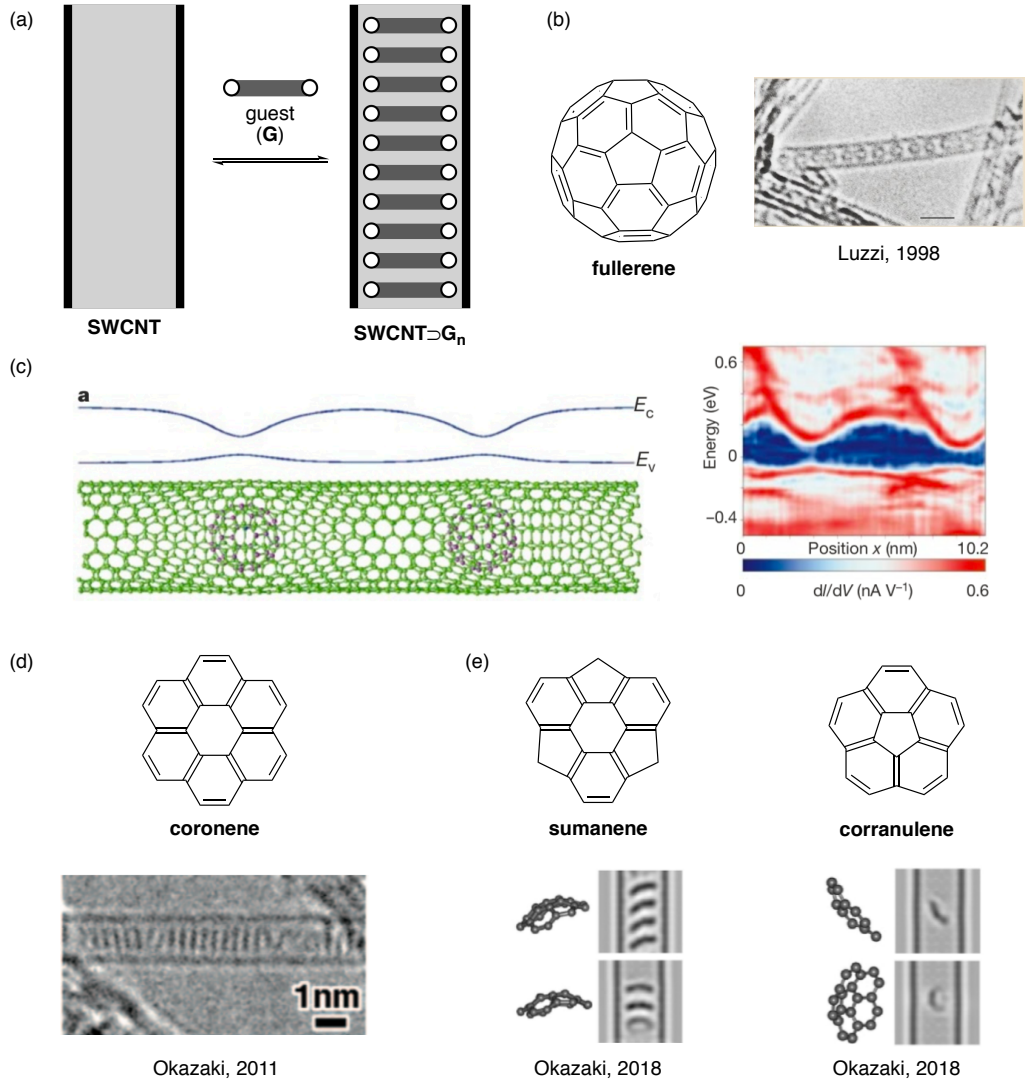


Figure 1.4. (a) Encapsulation of multiple molecules inside carbon nanotubes. (b) Encapsulation of fullerene in ref. 13. (c) A schematic illustration of the structure of (11,9)-SWCNT ⊃ Gd@C_{82} and its band structure showing conduction (E_c) and valence (E_v) edges in ref. 14. The dI/dV spectra taken along the 10.2-nm-long (11,9)-SWCNT ⊃ Gd@C_{82}. The x and y axes represent the position along the nanotube and the energy, respectively. The color shows the local density of states. (d) Encapsulation of coronene in ref. 16. (e) Encapsulations of sumanene and corannulene in ref. 17. Adapted from refs. 13, 14, 16 and 17 by permission from Springer Nature, Copyright 1998; Springer Nature, Copyright 2002; John Wiley and Sons, Copyright 2011; John Wiley and Sons, Copyright 2018.

Recently, the bottom-up approach to synthesize molecular entities to mimic SWCNT structures has flourished. Since the first report of hoop-shaped molecules, [*n*]cyclo-para-phenylene ([*n*]CPP) in 2008 by Jasti and Bertozzi,[19] various hoop-shaped and cylindrical-shaped molecules have been synthesized.[20,21,22,23,24] The concave cavity of hoop-shaped molecules provides the potential to act as a host for guest molecules. For instance, in 2011, it was found that [10]CPP can encapsulate C_{60} to form a supramolecular complex (Figure 1.5).[25] Job plot analysis using the UV-vis titration analysis suggested the 1:1 complexation, and the binding constant K_a in 1,2-dichlorobenzene was determined to be $(6.0 \pm 0.2) \times 10^3$ M^{-1}. Although the complex could be regarded as the "shortest fullerene-peapod", the single-linkage of phenylene panels allows for the free rotation of the panels, and the short length of the longitude axis precludes the accommodation of multiple guest molecules. Therefore, [*n*]CPPs do not perfectly mimic the belt-shaped character of SWCNTs, which should also affect the thermodynamic nature of the complexation.

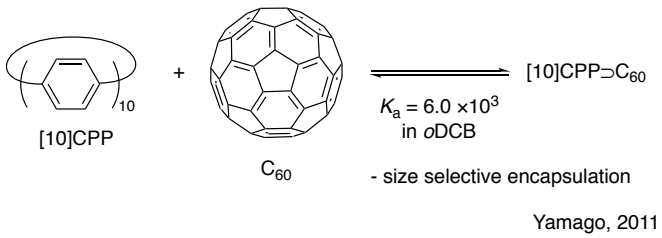


Figure 1.5. The supramolecular chemistry of [*n*]CPP in ref. 25.

In 2011, the first belt-persistent molecule, [4]cyclochrysenylene (**[4]CC**), was reported by the Isobe group (Figure 1.6). [26] Its π-lengthened version, [4]cycloanthanthrenylene (**[4]CA**), was also synthesized from abundant pigment as a starting material in 2013.[27] It was found that **[4]CC** and **[4]CA** serve as host molecules, in which various guest molecules can be encapsulated. It should be emphasized that precise determination of stoichiometry and thermodynamic parameters could be possible for these systems, which was in stark contrast to SWCNTs. First, it was revealed that C_{60} can be encapsulated in **[4]CC** in a 1:1 fashion by Job plot analysis using fluorescence titration, and inertial rotation of C_{60} inside **[4]CC** was observed.[28] Recently, a unique binding motif was discovered for **[4]CC** to form a supramolecular complex with a bowl-shaped molecule, corannulene (**COR**).[29] The complex was constructed by only CH-π hydrogen bonds. Interestingly, single-axis rotation of **COR** molecules in the solid-state was revealed by solid-state ^{2}H NMR analysis. This binding motif further allowed for the encapsulation of pyrene.[30] The pyrene molecule was static inside **[4]CC**, unlike **COR**. In the complex, **[4]CC** was deformed into an oval shape due to pyrene molecules. To examine an in-depth understanding of their structure-based properties, further diversification of the finite CNT segments was desired. In 2019, [3]cyclodibenzochrysenylene (**[3]C^{db}C**) was developed as a narrowed version of **[4]CC**. [31] It was found that **[3]C^{db}C** can encapsulate adamantane to form a supramolecular complex via CH-π hydrogen bonds.[32] Surprisingly, terahertz rotation of the adamantane molecule in **[3]C^{db}C** was observed. After the synthesis of a narrowed version of the finite CNT molecule, [4]cyclofulminenylene (**[4]CF**) was synthesized as a widened version in 2021.[33] It was also revealed that the association between **[4]CF** and **[3]C^{db}C** formed a complex whose structure was similar to that of DWCNTs.[15] Owing to the definite structure as molecular entities, a detailed thermodynamic discussion was possible for these systems. However, all the cases utilized the Job plot analysis for the determination of stoichiometry of the complexes, which may be reconsidered by using modern statistical analysis, such as the *F*-test.

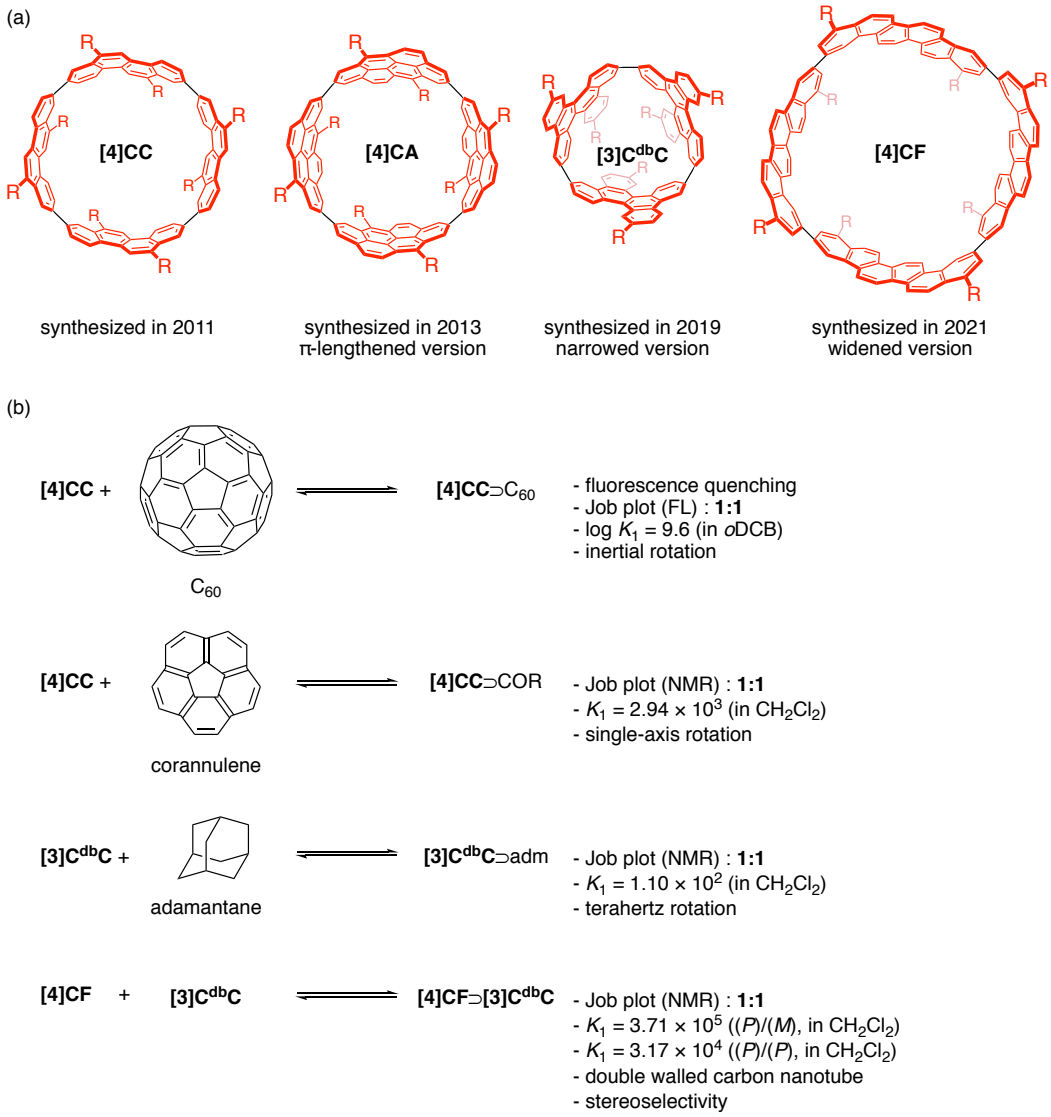


Figure 1.6. (a) Model molecules of the finite carbon nanotubes. (b) Supramolecular chemistry of the finite carbon nanotube molecule in ref. 28, 29, 32 and 33.

1.3 This work

Based on the abovementioned background information, this thesis describes the studies of supramolecular complexes of cylindrical hosts with multiple guests via statistical analyses as a unique showcase for highlighting the importance of the stoichiometry determination of complexes (Figure 1.7). As methods relying on statistical analyses, the *F*-test and an information-theory approach of Akaike's information criterion (AIC) are used for the determination of the stoichiometry of the supramolecular complexes.

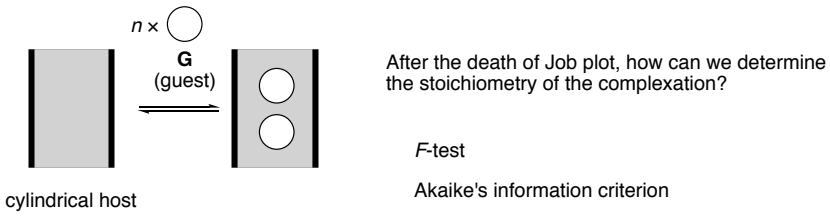


Figure 1.7. Survey of this work

In **Chapter 2**, supramolecular complexes comprising a cylindrical nanotube molecule, [4]cyclochrysenylene (**[4]CC**) and halogenated corannulenes (X-**COR**s) are surveyed (Figure 1.8). According to the precise titration experiments and statistical analysis using the *F*-test, the stoichiometry of the complexes was carefully determined to be 1:2. The precise determination of the stoichiometry allowed for the exploration of the cooperativity of the association. Interestingly, it was found that the cooperativity of the association changed with the size of the guest molecule.

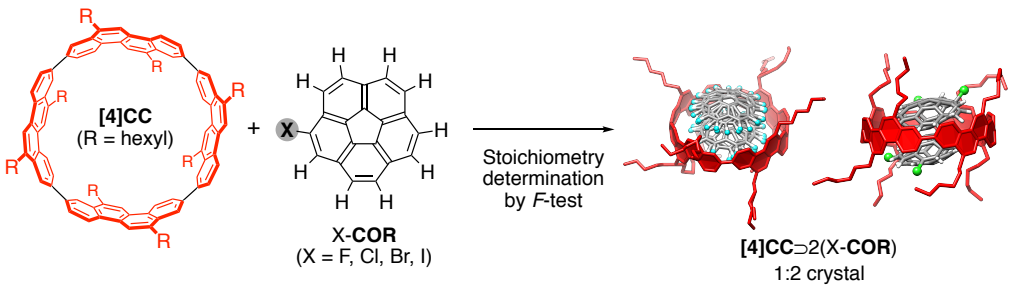


Figure 1.8. The association between **[4]CC** and X-**COR**s described in Chapter 2.

In **Chapter 3**, much more complicated supramolecular systems of cylindrical hosts are investigated. As a few test cases, supramolecular complexes of cylindrical hosts and C_{120}, which were revealed to show a 2:1 complexation by crystallographic analysis, are examined for the statistical analysis method to determine the stoichiometry (Figure 1.9). It was found that the *F*-test faced a limited applicability in these complicated supramolecular systems that encompass, for example, 1:2 and 2:1 models. Thus, AIC is introduced to the stoichiometry determination of the supramolecular complexes in this study. Several test cases with cylindrical hosts are also examined to demonstrate the usefulness of the AIC.

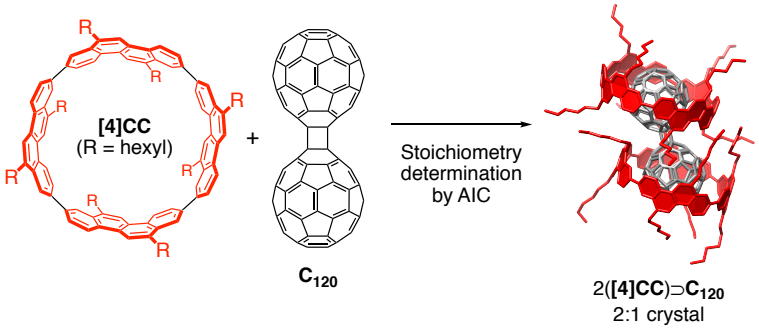


Figure 1.9. Application of AIC to the dumbbell-and-cylinder systems.

Chapter 4 summarizes the achievement and outlook of this study.

1.4 References

1. J.-M. Lehn, *Angew. Chem. Int. Ed.* **1988**, *27*, 90-112.
2. D. J. Cram, *Angew. Chem. Int. Ed.* **1988**, *27*, 1009-1020.
3. C. J. Pedersen, *Angew. Chem. Int. Ed.* **1988**, *27*, 1021-1027.
4. C. J. Pedersen, *J. Am. Chem. Soc.* **1967**, *89*, 2495-2496.
5. R. M. Izatt, J. H. Rytting, D. P. Nelson, B. L. Haymore, J. J. Christensen, *Science* **1969**, *164*, 443-444.
6. R. M. Izatt, J. S. Bradshaw, S. A. Nielsen, J. D. Lamb, J. J. Christensen, *Chem. Rev.* **1985**, *85*, 271-339.
7. A. B. C. Deutman, C. Monnereau, M. Moalin, R. G. E. Coumans, N. Veling, M. Coenen, J. M. M. Smits, R. de Gelder, J. A. A. W. Elemans, G. Ercolani, R. J. M. Nolte and A. E. Rowan, *Proc. Natl. Acad. Sci. U. S. A.*, **2009**, *106*, 10471-10476.
8. P. Job, *Ann. Chim.* **1928**, *9*, 113-203.
9. J. S. Renny, L. L. Tomasevich, E. H. Tallmadge, D. B. Collum, *Angew. Chem. Int. Ed.* **2013**, *52*, 11998-12013.
10. F. Ulatowski, K. Dabrowa, T. Balakier, J. Jurczak, *J. Org. Chem.* **2016**, *81*, 1746-1756.
11. D. B. Hibbert, P. Thordarson, *Chem. Commun.* **2016**, *52*, 12792-12805.
12. M. Monthioux, *Carbon* **2002**, *40*, 1809-1823.
13. B. W. Smith, M. Monthioux, D. E. Luzzi, *Nature* **1998**, *396*, 323-324.
14. J. Lee, H. Kim, S.-J. Kahng, G. Kim, Y.-W. Son, J. Ihm, H. Kato, Z. W. Wang, T. Okazaki, H. Shinohara, Y. Kuk, *Nature* **2002**, *415*, 1005-1008.
15. S. Bandow, M. Takizawa, K. Hirahara, M. Yudasaka, S. Iijima, *Chem. Phys. Lett.* **2001**, *337*, 48-54.
16. T. Okazaki, Y. Iizumi, S. Okubo, H. Kataura, Z. Liu, K. Suenaga, Y. Tahara, M. Yudasaka, S. Okada, S. Iijima, *Angew. Chem. Int. Ed.* **2011**, *50*, 4853-4857.
17. Y. Izumi, Z. Liu, K. Suenaga, S. Okada, S. Higashibayashi, H. Sakurai, T. Okazaki, *ChemNanoMat* **2018**, *4*, 557-561.
18. Y. Joko, R. Sasaki, K. Shintani, *Phys. Chem. Chem. Phys.* **2017**, *19*, 27704-27715.
19. R. Jasti, J. Bhattacharjee, J. B. Neaton, C. R. Bertozzi, *J. Am. Chem. Soc.* **2008**, *130*, 17646-17647.
20. S. Yamago, E. Kayahara, T. Iwamoto, *Chem. Rec.* **2014**, *14*, 84-100.

21. E. R. Darzi, R. Jasti, *Chem. Soc. Rev.* **2015**, *44*, 6401-6410.
22. S. E. Lewis, *Chem. Soc. Rev.* **2015**, *44*, 2221-2304.
23. Y. Segawa, H. Ito, K. Itami, *Nat. Rev. Mater.* **2016**, *1*, 15002.
24. E. J. Leonhardt, R. Jasti, *Nat. Rev. Chem.* **2019**, *3*, 672-686.
25. T. Iwamoto, Y. Watanabe, T. Sadahiro, T. Haino, S. Yamago, *Angew. Chem. Int. Ed.* **2011**, *50*, 8342-8344.
26. S. Hitosugi, W. Nakanishi, T. Yamasaki, H. Isobe, *Nat. Commun*. **2011**, *2*, 492.
27. T. Matsuno, S. Kamata, S. Hitosugi, H. Isobe, *Chem. Sci.* **2013**, *4*, 3179-3183.
28. H. Isobe, S. Hitosugi, T. Yamasaki, R. Iizuka, *Chem. Sci.* **2013**, *4*, 1293-1297.
29. T. Matsuno, M. Fujita, K. Fukunaga, S. Sato, H. Isobe, *Nat. Commun.* **2018**, *9*, 3779.
30. T. Matsuno, K. Fukunaga, S. Sato, H. Isobe, *Angew. Chem. Int. Ed.* **2019**, *58*, 12170-12174.
31. K. Kogashi, T. Matsuno, S. Sato, H. Isobe, *Angew. Chem. Int. Ed.* **2019**, *58*, 7385-7389.
32. T. Matsuno, S. Terasaki, K. Kogashi, R. Katsuno, H. Isobe, *Nat. Commun.* **2021**, *12*, 5062.
33. T. Matsuno, Y. Ohtomo, M. Someya, H. Isobe, *Nat. Commun.* **2021**, *12*, 1575.

Chapter 2. Exploration of Cooperativity in Nanotubes

2.1 Introduction

The inner space of carbon nanotubes (CNTs) is a unique place surrounded by a smoothly curved π-wall. Being nanometer-scale in diameter, the inner space captures various molecules as guests.[1,2,3,4,5] One of the unique features of CNTs as hosts is their longitudinal length, which allows encapsulation of multiple guest molecules. As guests, aromatic molecules are currently being explored, which allows for stacking the guest molecules in one-dimensional arrays (Figure 2.1a).[6,7,8] The recent advent of molecular versions of CNTs is further deepening our understanding of CNT chemistry with discrete molecular structures[9,10] including the host-guest chemistry of the aromatic guests in CNTs.[11,12] A segmental molecule of helical CNT, *i.e.*, (*P*)-(12,8)-[4]cyclochrysenylene (**[4]CC**), was found to encapsulate corannulene (H-**COR**) and allowed for dynamic, single-axis rotations of the guest in a solid. Although fundamental thermodynamic studies of the bowl-in-tube complex revealed its enthalpy-driven complexation *via* CH-π hydrogen bonds, there remained an interesting question related to the π-stack arrays in the cylindrical host. Thus, in solution, the formation of a 1:1 complex [**[4]CC**⊃1(H-**COR**)] was concluded, whereas a 1:2 π-stack complex [**[4]CC**⊃2(H-**COR**)] was found in the crystal. Realizing that π-stacks with multiple guests could be an interesting test case to examine cooperativity, I decided to perform structure-association relationship studies by adopting a series of halogenated corannulenes (X-**COR**) as guests (Figure 2.1c).

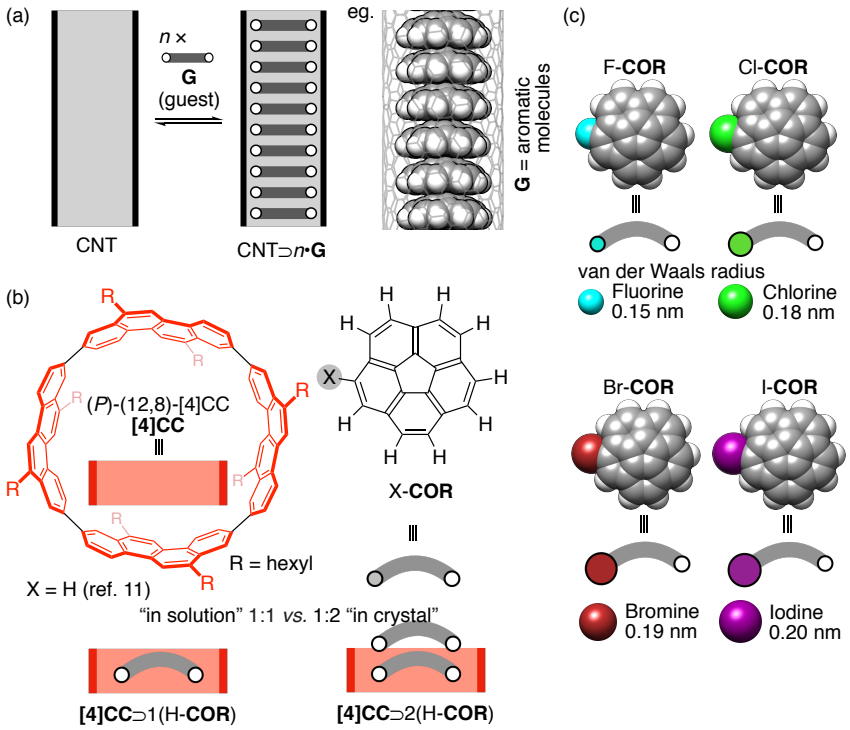


Figure 2.1. Host-guest complexes of cylindrical hosts. (a) Host-guest complexes of CNTs with guests. (b) Host-guest complexes of **[4]CC** and H-**COR**.[11] (c) Halogenated corannulenes (X-**COR**) used as guests in this study. The structures were obtained by DFT calculations [LC-BLYP/6-311G(d)]. The van der Waals radii are taken from the literature.[22]

Cooperativity is an important concept for understanding the fundamental association behaviors of multiple guests captured in a host. After the discovery of natural cooperative systems such as hemoglobin-oxygen complexes,[13] various artificial systems have been developed to deepen our understanding of the thermodynamics, and as a measure of cooperativity, the interaction parameter (α) has been introduced (Figure 2.2).[14,15,16] When $\alpha > 1$, the system is judged to have a positive cooperativity with multiple guests cooperating with each other to facilitate the association. When $\alpha < 1$, negative cooperativity is indicated to show the repelling relationships of multiple guests. With $\alpha = 1$, the system is judged to be non-cooperative with multiple guests being inert to each other. More specifically, for two-stage complexations of two guests at separate sites in a host (two-site host), the α value is obtained as "pre-factor × K_2/K_1", with a pre-factor of 4 being applied to the ratio of association constants.[15] Careful considerations are necessary, however, to apply this α measure to the host-guest systems of CNTs due to

structural features unique to the cylindrical host. Thus, unlike the two-site host, the cylindrical host possesses one association site for multiple guests (one-site host), and to derive the α value for two-stage association, the α measure needs to be modulated. As a relevant example of a one-site host possessing a cylindrical space to encapsulate two guest molecules, cucurbit[*n*]uril ([*n*]CB) has been extensively studied.[17, 18, 19, 20] Considering that two guests are encapsulated in one identical site, the pre-factor for the α measure has been reconsidered as 2 (Figure 2.2).[21] The revised analyses revealed that the two-stage associations of two guests in one-site host of [*n*]CB were often non-cooperative ($\alpha \sim 1$). Although the second association gains favorable enthalpic contributions, counterbalancing entropic contributions cancel the cooperativity. In this study, structure-association relationship studies with a one-site cylindrical host composed solely of hydrogen and carbon atoms showed that a size mismatch in the host-guest complex can favorably result in the cooperative associations of two guests.[22]

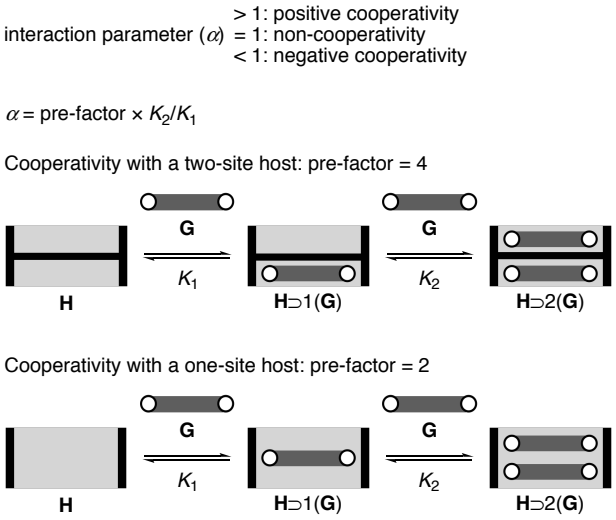


Figure 2.2. Cooperativity in two-stage associations of two guests in two-site and one-site hosts.

2.2 Confirmation of the formation of supramolecular complexes

Encapsulations of X-**COR** in **[4]CC** were first investigated by ^{1}H NMR spectroscopy to reveal basic association behaviors. As was the case with H-**COR**,[11] the ^{1}H resonances of the guest served as excellent references for the association analyses. Thus, the aromatic resonances of X-**COR** commonly exhibited large upfield shifts upon mixing with the **[4]CC** host in dichloromethane-d_2 (Figures 2.3-2.6). The upfield shifts indicated the formation of bowl-in-tube structures, which resulted in shielding effects from the surrounding π-wall toward inner guests. When the mixing ratios were varied for continuous variation analyses, gradual changes in the upfield shifts were observed (Figures 2.3-2.6). The observation indicated the presence of fast in-and-out exchange, which should result in the time-averaged resonances of unbound and bound guests.

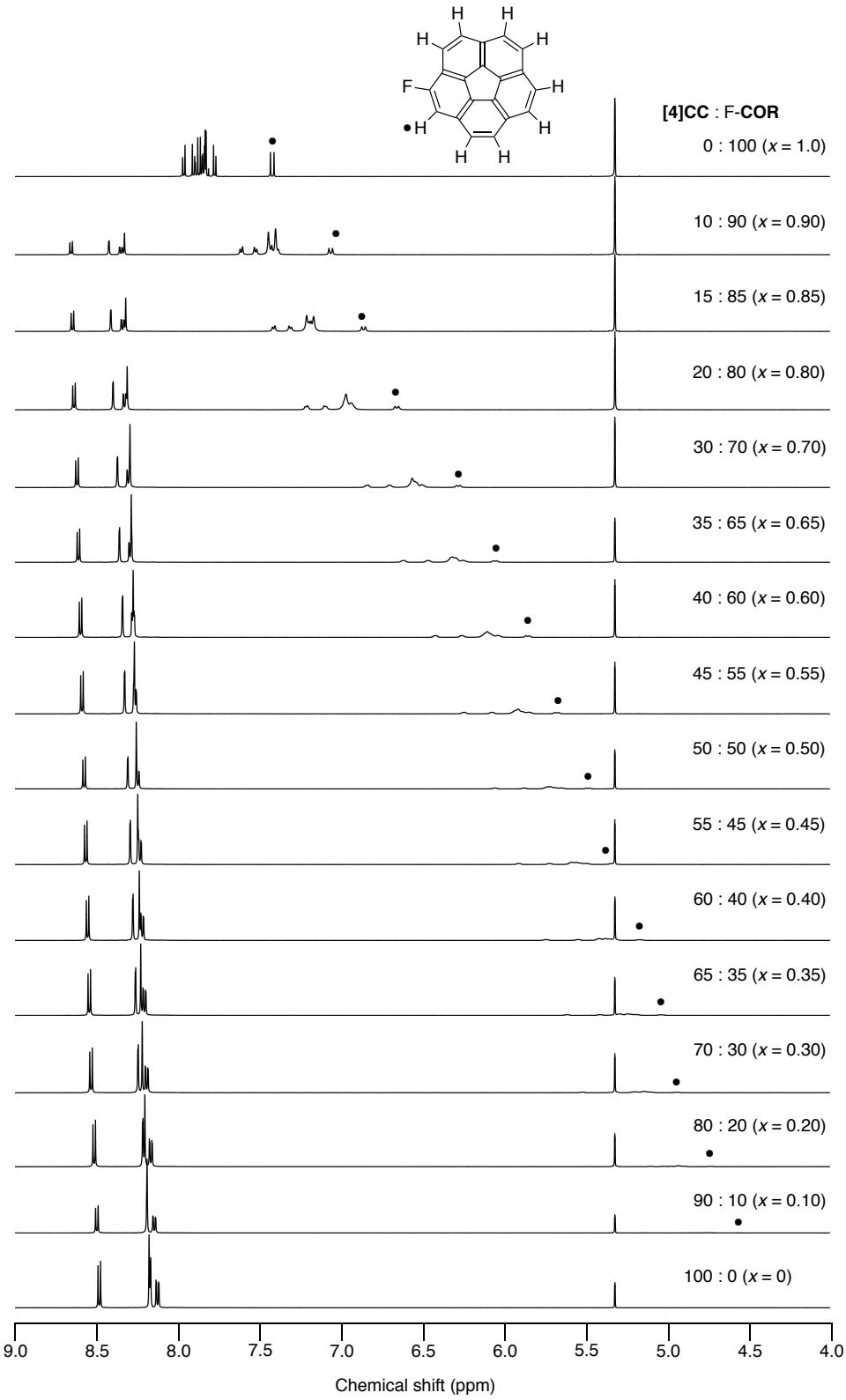


Figure 2.3. ^{1}H NMR spectra of the mixture of **[4]CC** and F-**COR** for Job plot analysis. Total concentration = 1.0×10^{-2} M (CD_2Cl_2, 298 K). Resonances of F-**COR** used for Job plot analysis are labeled with dots.

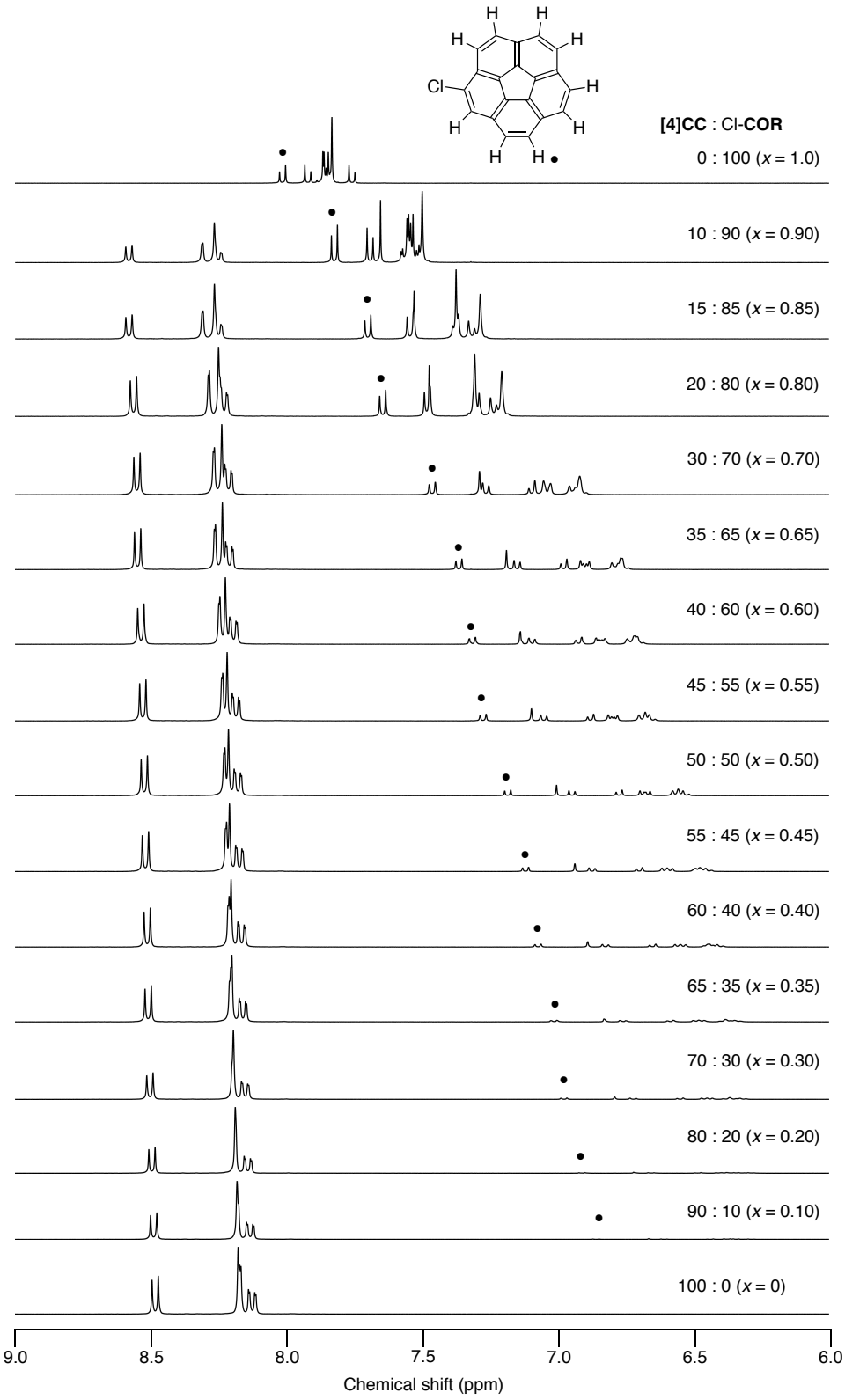


Figure 2.4. ^{1}H NMR spectra of the mixture of **[4]CC** and Cl-**COR** for Job plot analysis. Total concentration = 1.0×10^{-2} M (CD_2Cl_2, 298 K). Resonances of Cl-**COR** used for Job plot analysis are labeled with dots.

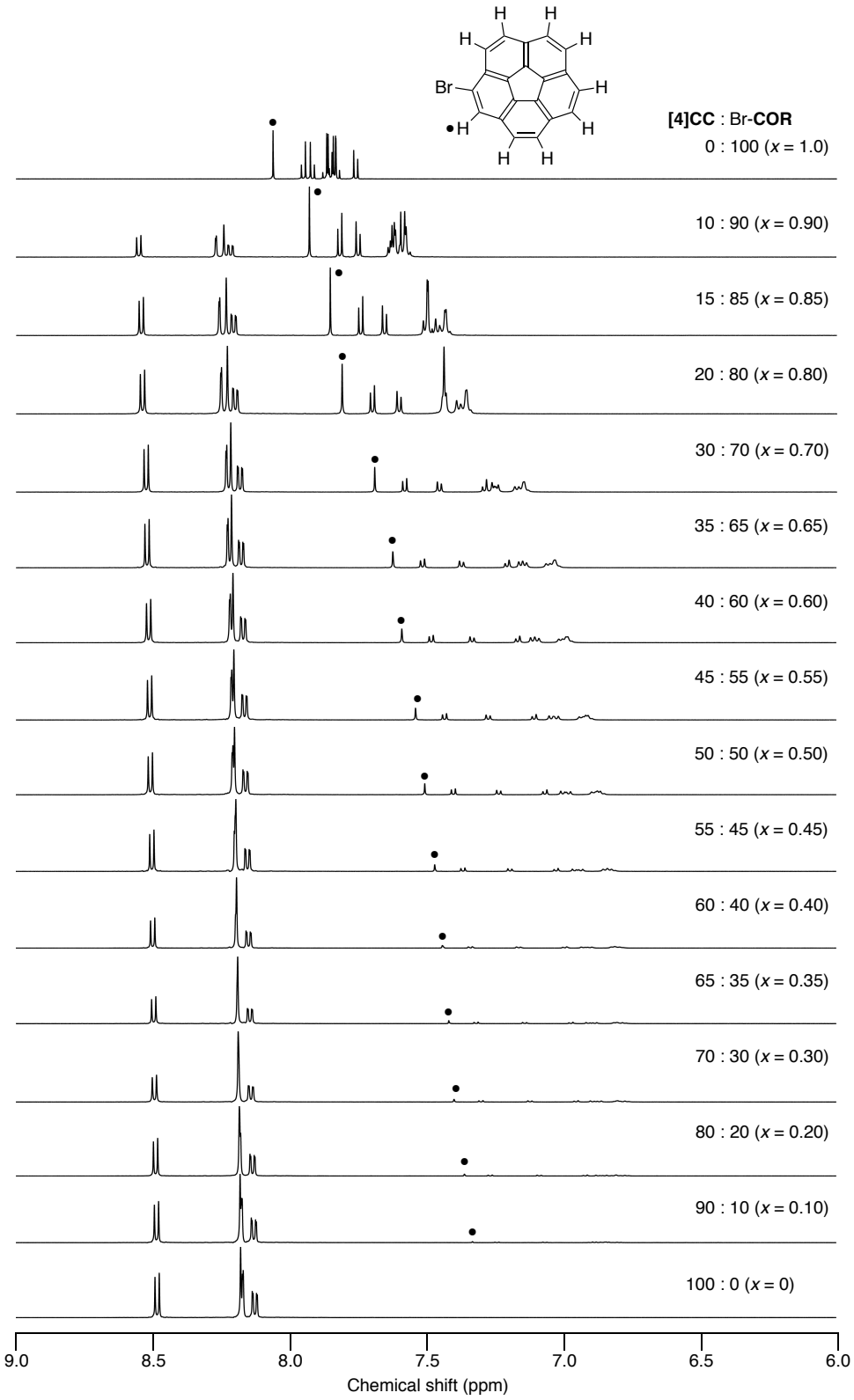


Figure 2.5. ^{1}H NMR spectra of the mixture of **[4]CC** and Br-**COR** for Job plot analysis. Total concentration = 1.0×10^{-2} M (CD_2Cl_2, 298 K). Resonances of Br-**COR** used for Job plot analysis are labeled with dots.

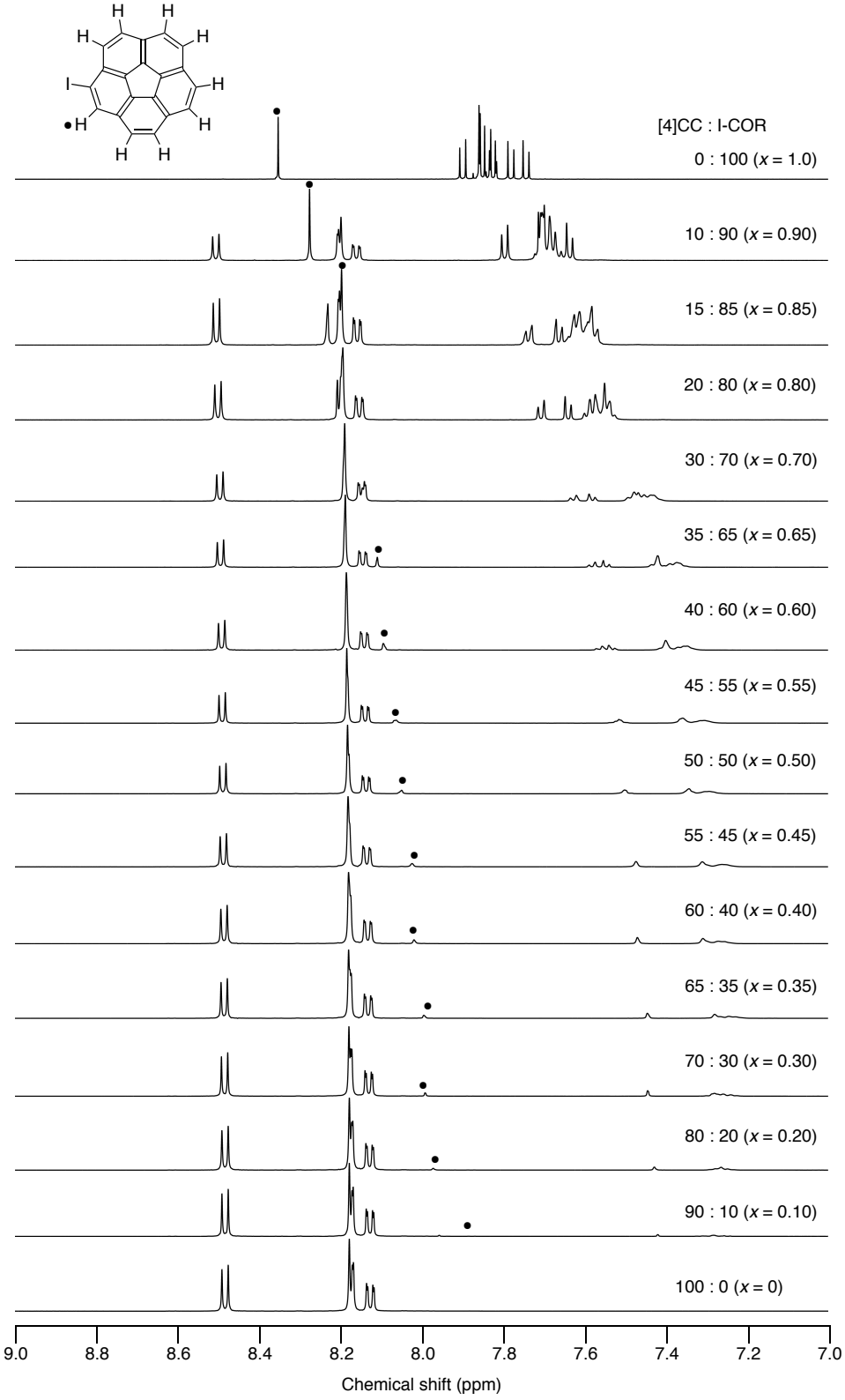


Figure 2.6. ^{1}H NMR spectra of the mixture of **[4]CC** and I-**COR** for Job plot analysis. Total concentration = 1.0×10^{-2} M (CD_2Cl_2, 298 K). Resonances of I-**COR** used for Job plot analysis are labeled with dots.

2.3 Determination of stoichiometry by Job plot analysis

As was the case with H-**COR**,[11] the stoichiometry of the association was first investigated by a continuous variation method (Job plot) using ^{1}H NMR spectra (Figure 2.7). Although the variation curve indicated a 1:1 stoichiometry with F-**COR** with a peak at x_{max} = 0.50, the peak shifted toward a 1:2 stoichiometry (x_{max} > 0.5) with other congeners (X=Cl, Br and I). This discrepancy observed for F-**COR** was in question because of the similarity of the structures of halogenated corannulenes. While investigating the origins of this discrepancy, a few papers in the literature caught my attention, which mentioned the limited applicability of the results/interpretations of the Job plot data. Specifically, when there is no and/or negative cooperativity in associations of multiple guests, the Job plot may well fail to indicate the correct stoichiometry, and alternative statistical analysis is advocated for determining the stoichiometry of host-guest complexes.[23]

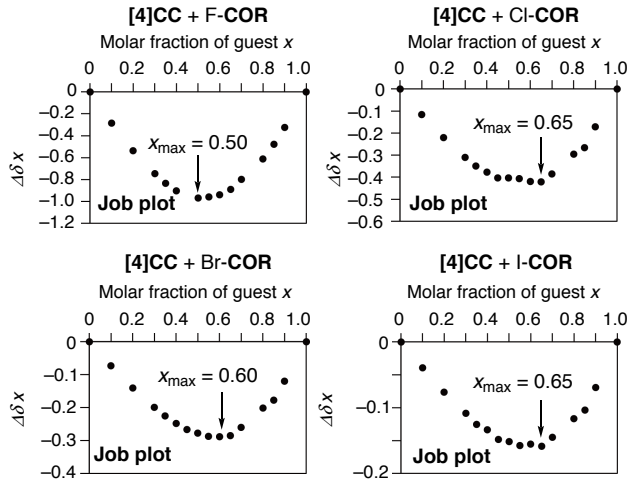


Figure 2.7. Job plot analysis of **[4]CC** and X-**COR**

2.4 Determination of stoichiometry by *F*-test

Statistical analysis was used to determine the stoichiometry of the host-guest complexes of **[4]CC** and X-**COR**. Because titration experiments with careful analyses are necessary to clarify the stoichiometry for such cases, isothermal titration calorimetry (ITC) analyses were performed. The titration experiments were performed in triplicate for each X-**COR** congener, and the thermograms were analyzed by fitting both the 1:1 and 1:2 models (Figures 2.8–2.11). The fit of the model was first evaluated by two indicators, *i.e.*, χ^2 and cov_{fit} (Table 2.1).[23] These two indicators allow us to evaluate the fit of each model by providing goodness-of-fit measures. For all the X-**COR** congeners, smaller values of χ^2 and cov_{fit} indicators were found with the 1:2 models, showing the superiority of the 1:2 association model.

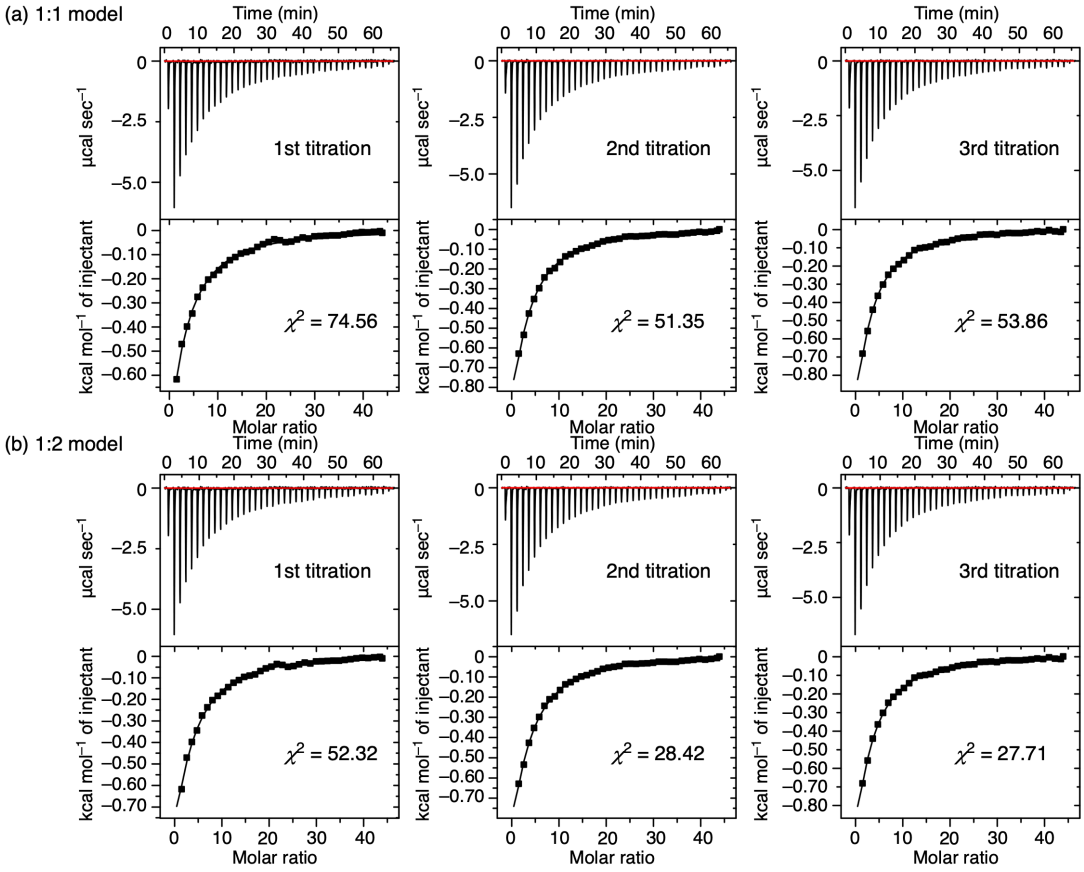


Figure 2.8. ITC analyses of **[4]CC** and F-**COR** in CH_2Cl_2 at 298 K. Concentrations: **[4]CC** = 4.86×10^{-4} M, F-**COR** = 9.89×10^{-2} M. (a) Fitting with the 1:1 model. (b) Fitting with the 1:2 model.

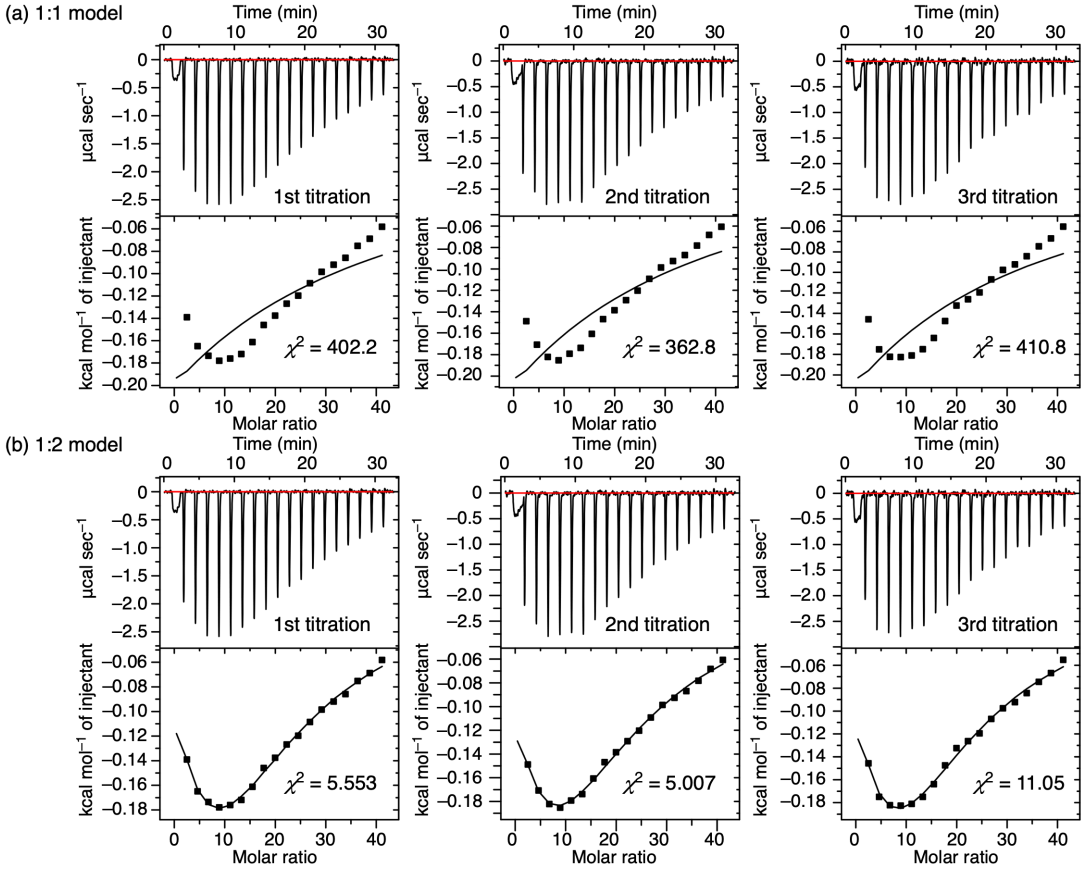


Figure 2.9. ITC analyses of **[4]CC** and Cl-**COR** in CH_2Cl_2 at 298 K. Concentrations: **[4]CC** = 4.89×10^{-4} M, Cl-**COR** = 1.06×10^{-1} M. (a) Fitting with the 1:1 model. (b) Fitting with the 1:2 model.

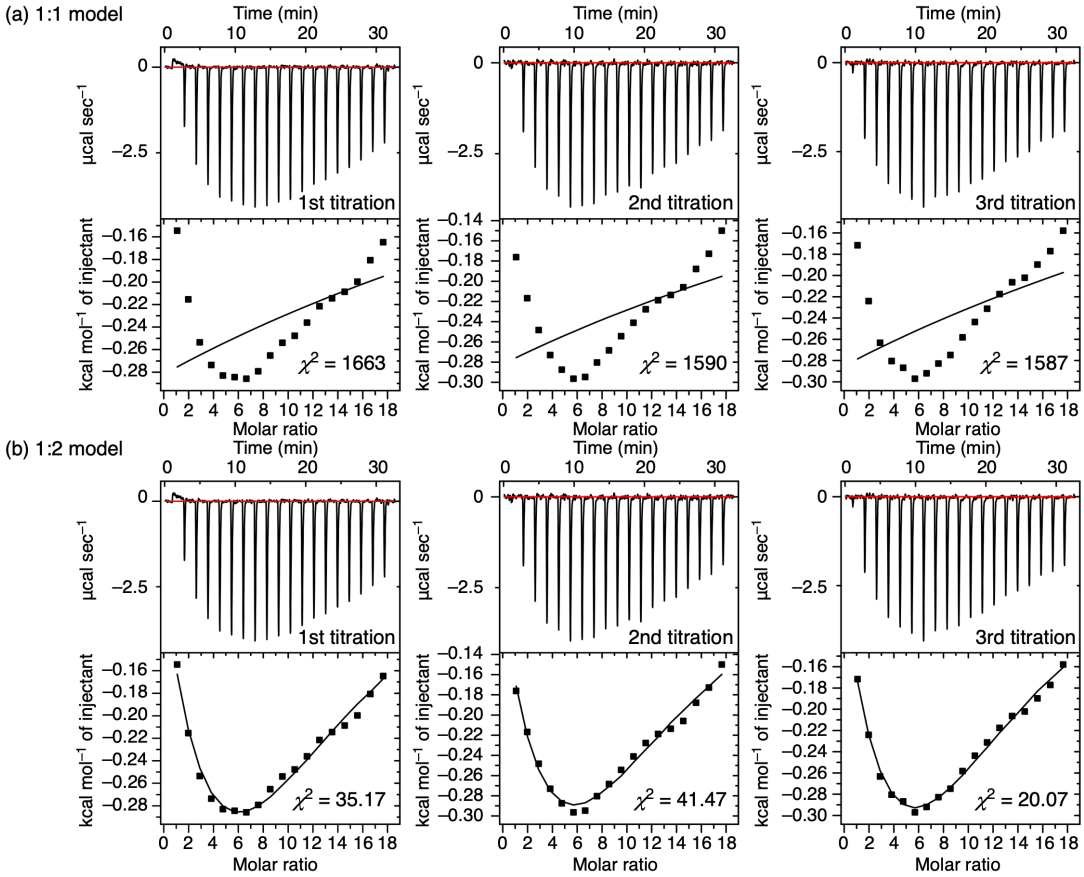


Figure 2.10. ITC analyses of **[4]CC** and Br-**COR** in CH_2Cl_2 at 298 K. Concentrations: **[4]CC** = 1.04×10^{-3} M, Br-**COR** = 9.67×10^{-2} M. (a) Fitting with the 1:1 model. (b) Fitting with the 1:2 model.

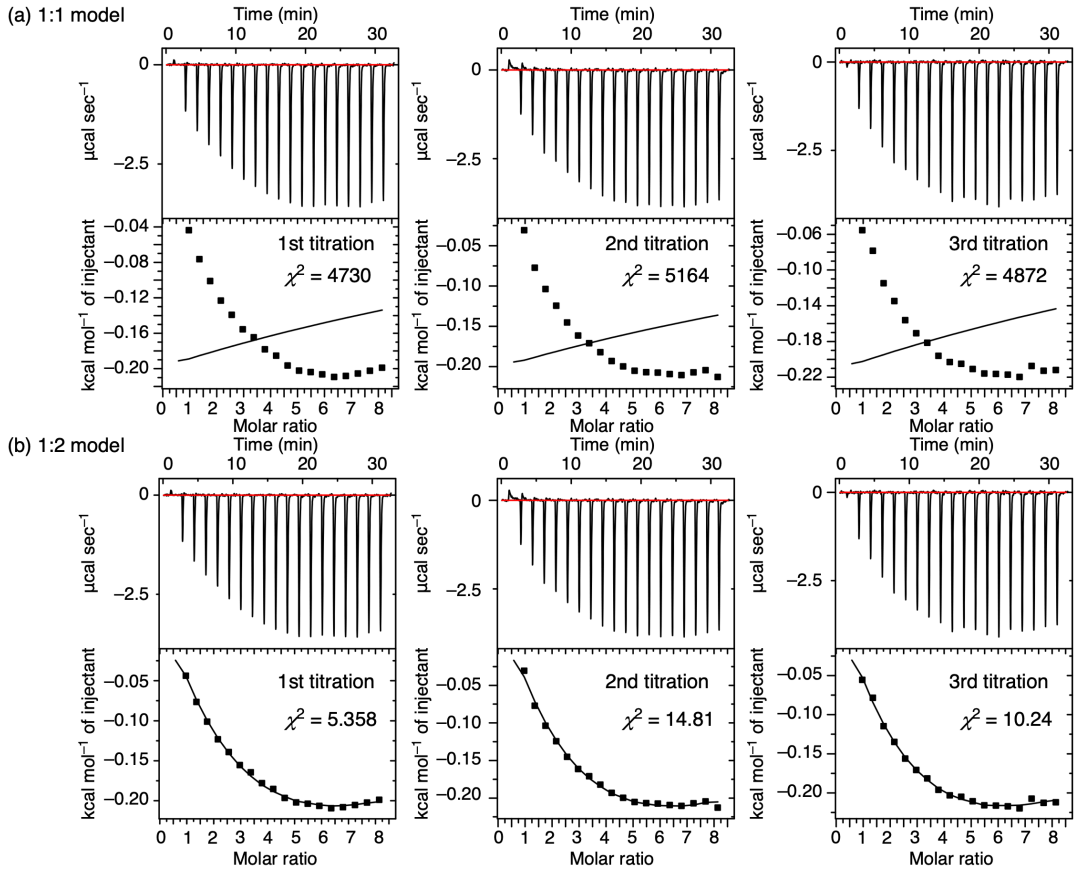


Figure 2.11. ITC analyses of **[4]CC** and I-**COR** in CH_2Cl_2 at 298 K. Concentrations: **[4]CC** = 2.20×10^{-3} M, I-**COR** = 8.85×10^{-2} M. (a) Fitting with the 1:1 model. (b) Fitting with the 1:2 model.

The quality of the fitting was further analyzed by a residual plot (Figures 2.12-2.15).[18,23] In residual plots, y_{fit} is plotted versus the molar ratio $[G]_{tot}/[H]_{tot}$, where y_{fit} is fitted residuals defined as the difference between the experimental values y_{data} and the calculated values y_{calc}. When data are fitted with a good model, the residuals can be distributed randomly because of the randomness of errors. On the other hand, when the data are fitted with a bad model, a sigmoidal distribution of the residuals is observed because the residuals are affected not only by random errors but also arbitrary and biased fitting. When the ITC data of **[4]CC** and X-**COR** were subjected to residual plot analysis, regular distributions were found for the residuals, specifically for the 1:1 models for Cl-, Br- and I-**COR**, which could arise from improper model selection. Therefore, the residual plot also corroborated the superior fitting of the 1:2 models rather than the 1:1 models.

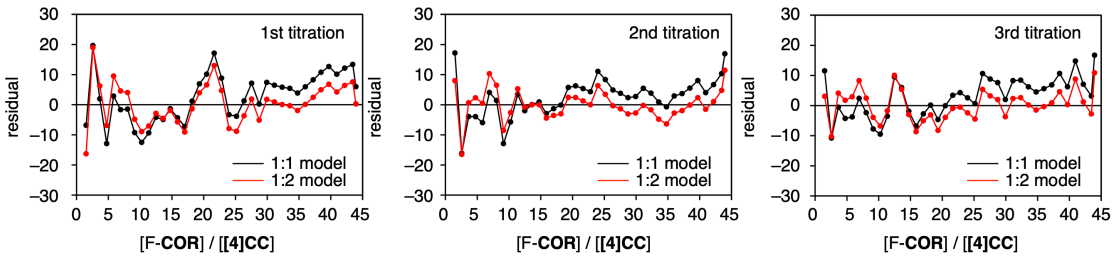


Figure 2.12. Residual plot for ITC analyses of **[4]CC** and F-**COR**.

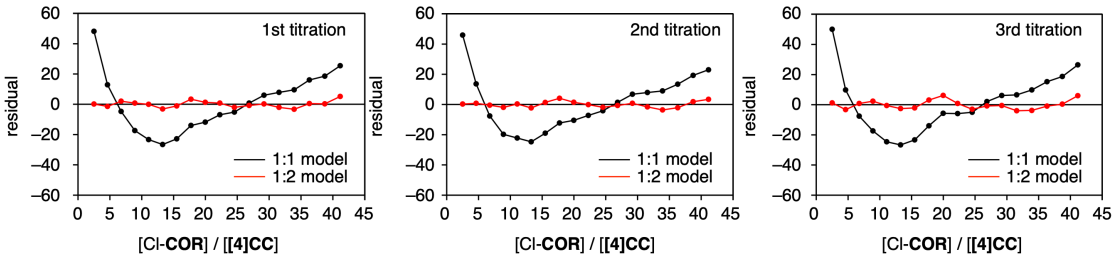


Figure 2.13. Residual plot for ITC analyses of **[4]CC** and Cl-**COR**.

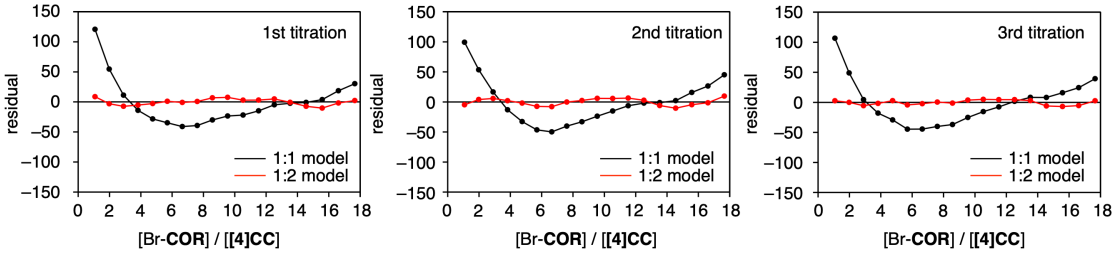


Figure 2.14. Residual plot for ITC analyses of **[4]CC** and Br-**COR**.

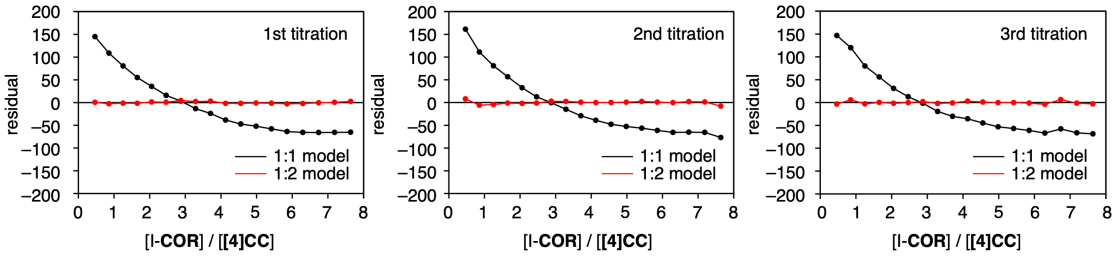


Figure 2.15. Residual plot for ITC analyses of **[4]CC** and I-**COR**.

The two models were further compared by a statistical comparison test (*F*-test) with the probability indicators (*P*-values). Thus, with all the X-**COR** congeners as guests, the *F*-values were larger than 1, which suggested two possible scenarios: 1. The more complicated 1:2 model is correct. 2. The simpler 1:1 model is correct, but random scatter leads the more complicated model to fit better.[23] The latter scenario was disproved by the small *P*-values, *i.e.*, the small probability of this coincidence happening. Therefore, it was concluded that the 1:2 model gave a better fit for the titration data for all the X-**COR** guests. The data are summarized in Table 2.1 and Figure 2.16. These results allowed for further detailed examinations of thermodynamic parameters (see below).[24]

Table 2.1. Titration experiments and analyses. After the analysis was performed independently in triplicate to obtain association constants and statistical values, the values were averaged with standard deviations. With comparison analyses, more likely data are underlined.[a]

X-**COR** model	K_1 (M^{-1})	K_2 (M^{-1})	χ^2	cov_{fit}	F	P
F-**COR**	305	-	59.9	0.00228	14.1	3×10^{-12}
1:1	±18		±12.7	±0.00080	±4.8	
F-**COR**	289	20.0	<u>36.2</u>	<u>0.00150</u>	-	-
1:2	±17	±4.5	<u>±14.0</u>	<u>±0.00074</u>		
Cl-**COR**	18.4	-	392	0.213	479	2×10^{-16}
1:1	±1.5		±26	±0.020	±163	
Cl-**COR**	81.9	98.5	<u>7.20</u>	<u>0.00337</u>	-	-
1:2	±3.9	±4.4	<u>±3.34</u>	<u>±0.00133</u>		
Br-**COR**	0.132	-	0.0161	0.789	432	5×10^{-16}
1:1	±0.014		±0.0043	±0.088	±171	
Br-**COR**	65.3	89.4	<u>32.2</u>	<u>0.0139</u>	-	-
1:2	±1.0	±4.8	<u>±11.0</u>	<u>±0.0050</u>		
I-**COR**	0.00378	-	0.00492	1.73	4750	3×10^{-23}
1:1	±0.00010		±0.00221	±0.04	±2230	
I-**COR**	32.4	42.6	<u>10.1</u>	<u>0.00311</u>	-	-
1:2	±5.0	±5.3	<u>±4.7</u>	<u>±0.00129</u>		

[a] Details of the parameters can be found in ref. 23.

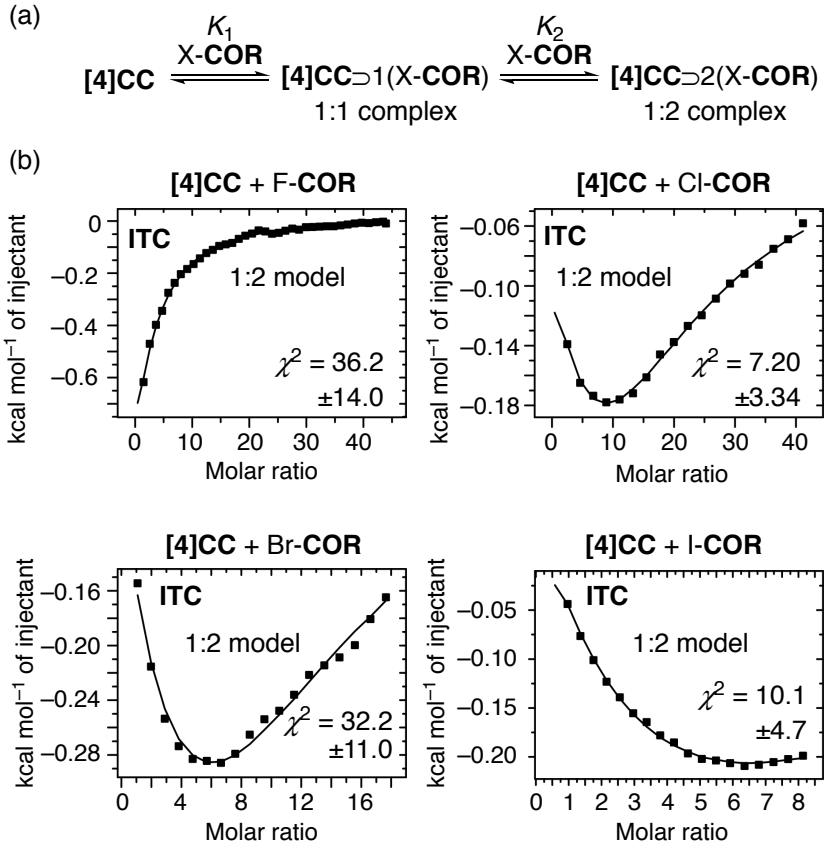


Figure 2.16. Bowl-in-tube association of **[4]CC**⊃*n*(X-**COR**). (a) Equilibrium to form possible 1:1 and 1:2 complexes. (b) Stoichiometry and association of **[4]CC**⊃*n*(X-**COR**). With careful titration analyses, $n = 2$ was concluded for each X-**COR** congener. For the ITC data, a representative thermogram from the triplicate titration experiments is shown.

2.5 Reanalysis of the H-COR complex

Based on the 1:2 complexation of X-**COR** with **[4]CC**, it was decided that the titration data of H-**COR**, which was reported previously to show the 1:1 complexation in the Job plot analysis,[11] was reanalyzed in this study using statistical analysis. The titration data, including careful fitting/statistics analyses, are summarized in Figures 2.17-2.18 and Table 2.2, and the analysis gave a stoichiometry of 1:2 even for H-**COR**.

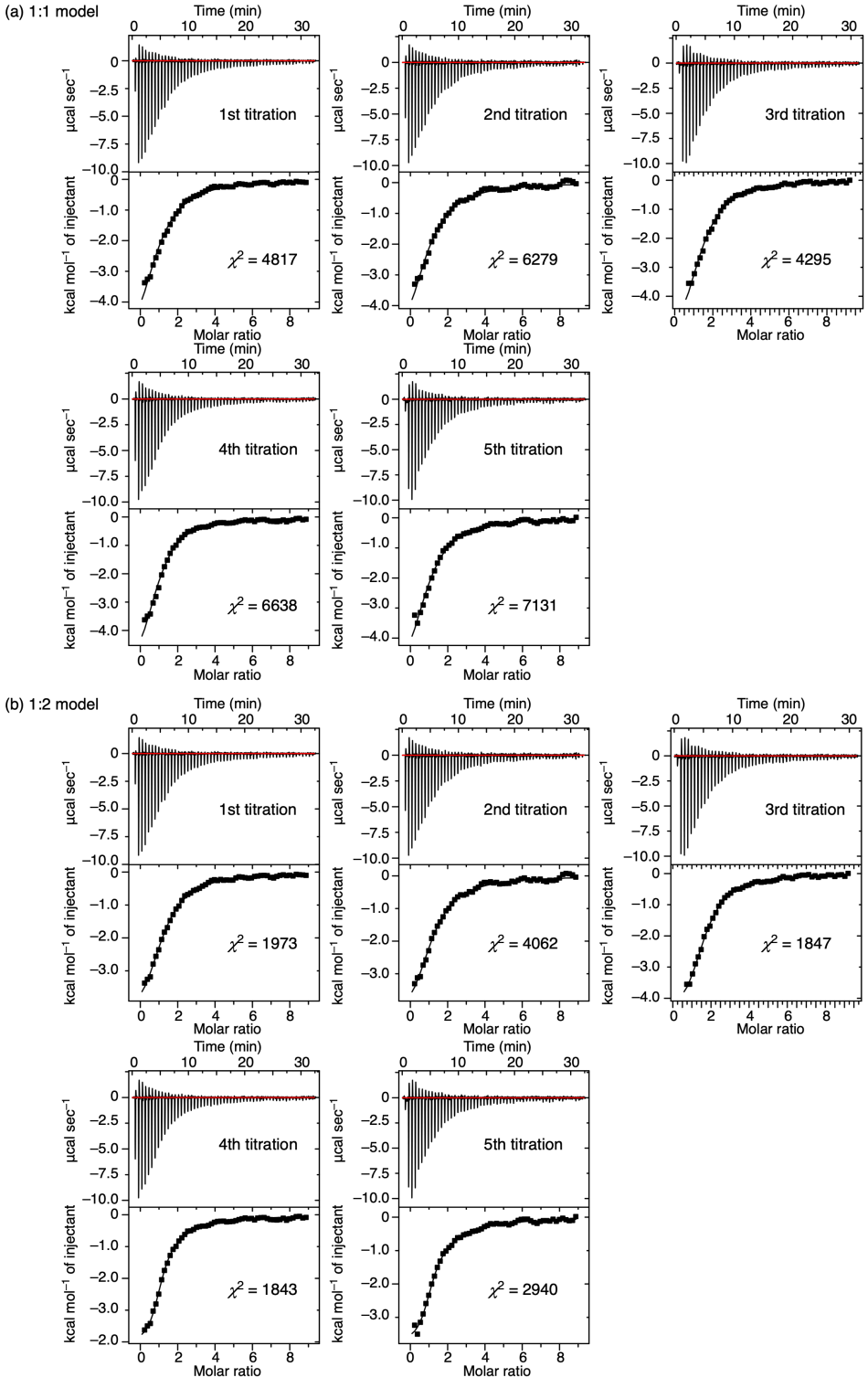


Figure 2.17. ITC analyses of **[4]CC** and H-**COR** in CH_2Cl_2 at 298 K. Concentrations: **[4]CC** = 5.23 × 10^{-4} M, H-**COR** = 2.32 × 10^{-2} M. (a) Fitting with the 1:1 model. (b) Fitting with the 1:2 model.

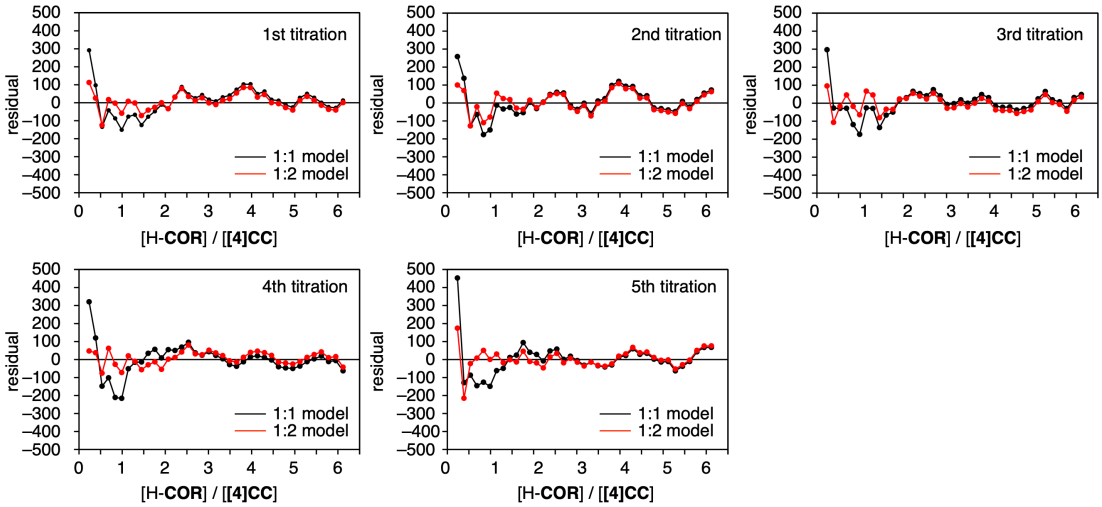


Figure 2.18. Residual plot for ITC analyses of **[4]CC** and H-**COR**

Table 2.2. Titration experiments and analyses. After the analysis was performed independently in triplicate to obtain association constants and statistical values, the values were averaged with standard deviations. With comparison analyses, more likely data are underlined.

X-**COR** model	K_1 (M^{-1})	K_2 (M^{-1})	χ^2	cov_{fit}	F	P
H-**COR**	2940	-	5830	0.00636	39.2	2×10^{-27}
1:1	±240		±1220	±0.00149	±19.1	
H-**COR**	16400	1140	2530	0.00272	-	-
1:2	±8200	±110	±970	±0.00122		

2.6 Crystal structures of the complexes

Molecular structures of the 1:2 complexes were revealed by X-ray crystallographic analyses of single crystals for F-**COR** and Cl-**COR**. The single crystals were grown in dichloromethane by mixing an excess amount of X-**COR** (X=F or Cl) with **[4]CC** in the presence of the vapor of a poor solvent (2-propanol for F and acetonitrile for Cl) at 25 °C. Both F-**COR** and Cl-**COR** resulted in 1:2 complexes [**[4]CC**⊃2(X-**COR**)], which confirmed the validity of the conclusions of the titration experiments (Figure 2.19). As was the case with H-**COR**,[11] the X-**COR** guests were found in the **[4]CC** host at different rotational orientations. By scrutinizing the structures, an interesting difference between **[4]CC**⊃2(F-**COR**) and **[4]CC**⊃2(Cl-**COR**) was found. Whereas the F-**COR** guests were axially aligned along the cylinder, the Cl-**COR** guests were tilted in the cylinder. For the quantitative measure of tilting, the axes of the X-**COR** bowl and the **[4]CC** cylinder are defined by using Mercury CSD (2020.3.1) (Figure 2.20)[25]. For **[4]CC**, carbon atoms at single-bond linkages were first adopted to define a mean plane, and the cylinder axis was defined as an axis perpendicular to this plane. For X-**COR**, carbon atoms at the central pentagon were adopted to define a mean plane, and the bowl axis was defined as an axis perpendicular to this plane. With the crystal structures, the tilt angles were obtained by using these axes, which revealed the large inclination for Cl-**COR** guests with tilt angles of 18.6°-25.9° (Figure 2.19).

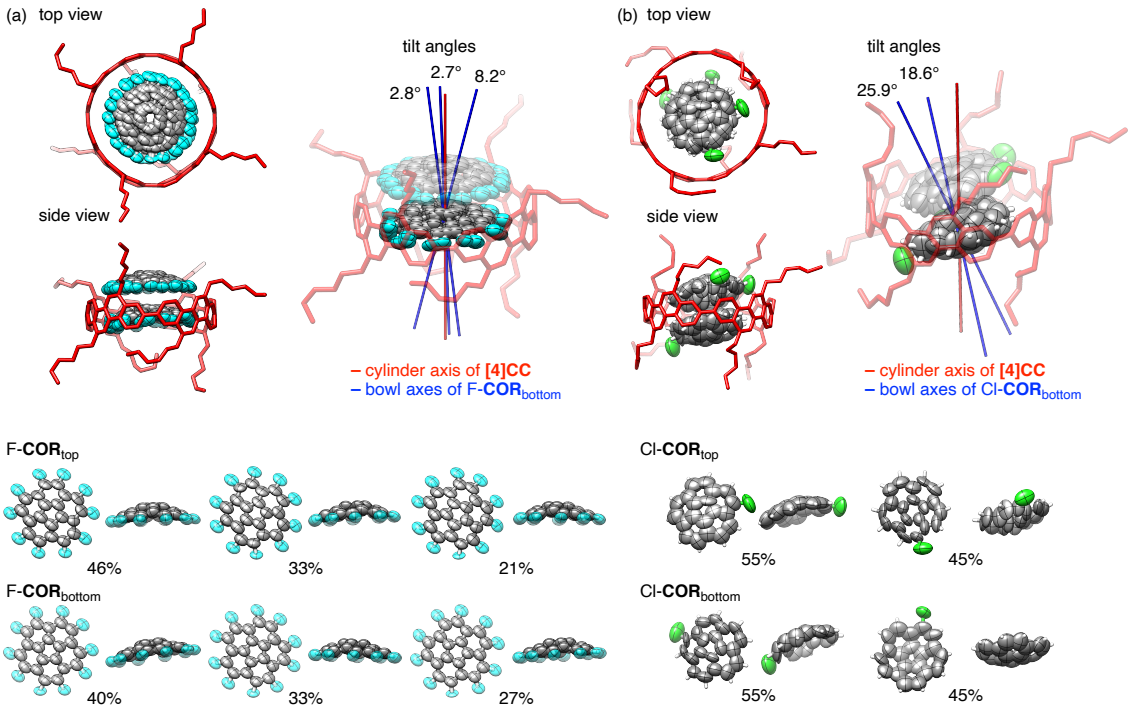


Figure 2.19. Crystal structures of 1:2 complexes. Thermal ellipsoids of X-**COR** are shown at 50% probability. The hydrogen atoms of **[4]CC** are omitted for clarity. (a) Structures of **[4]CC**⊃2(F-**COR**). (b) Structures of **[4]CC**⊃2(Cl-**COR**).

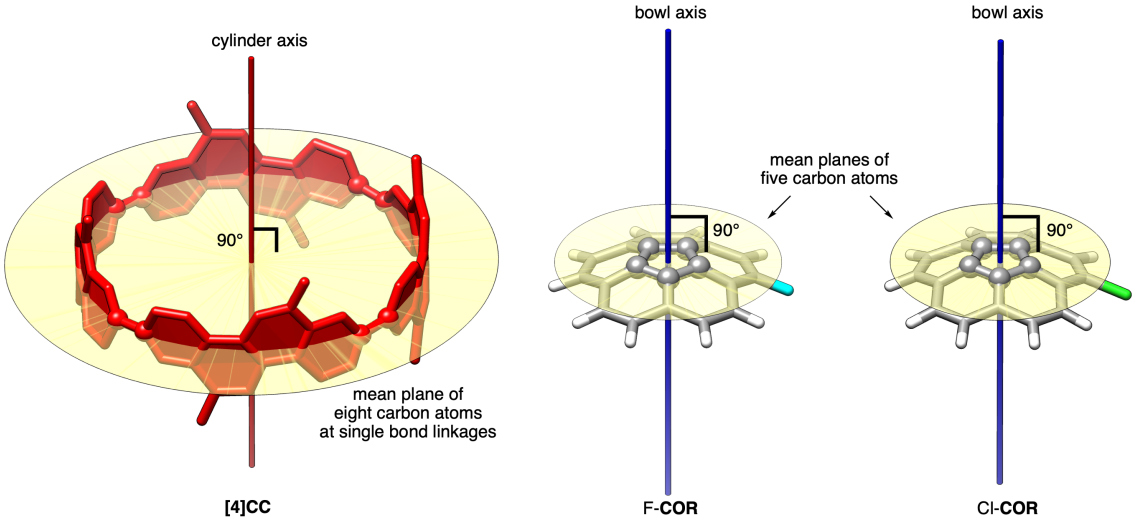


Figure 2.20. Axes of **[4]CC** cylinder and X-**COR** bowl for measurements of tilt angles.

2.7 Thermodynamics and cooperativity of the systems

Having confirmed the 1:2 complex as an appropriate model, the ITC data were further analyzed to elucidate the thermodynamics and cooperativity. Association constants for the two-stage association (K_1 and K_2) were first obtained from the ITC data. The first-stage association constants qualitatively depended on the size of the halogen substituents, varying from 289 ± 17 M^{-1} for F-**COR** to 32.4 ± 5.0 M^{-1} of I-**COR** (Figure 2.21). The second-stage association constant for F-**COR** was much smaller than the first one, whereas for other X-**COR** congeners (X=Cl, Br and I), the second-stage constants became larger. When these association constants were converted to the interaction parameters ($\alpha = 2\ K_2/K_1$), *i.e.*, the measure of cooperativity,[15,21] the α values became 0.14 for F-**COR**, 2.41 for Cl-**COR**, 2.74 for Br-**COR** and 2.63 for I-**COR**. Thus, negative cooperativity was indicated for F-**COR**, and positive cooperativity was demonstrated for the other X-**COR** congeners (X=Cl, Br and I). Thermodynamic parameters for the two-stage associations provided insight into the cooperativity (Figure 2.21). The first-stage associations for all the X-**COR** congeners were enthalpically favored, although the ΔH_1 values depended on the halogen species and decreased from 6.35 ± 0.16 kcal mol^{-1} for F-**COR** to 0.31 ± 0.11 kcal mol^{-1} for I-**COR**. This enthalpy dependence on the halogen species may be understood by taking into account the crystal structures of F-**COR** and Cl-**COR** (Figure 2.19). Thus, when the halogen atom is small enough for the **[4]CC** host to wrap around the rim, hydrogen atoms can make good contact with the cylinder wall for CH-π interactions.[11] However, when the halogen atom becomes larger, the bowl guest cannot be fully wrapped by the **[4]CC** host and is tilted to lose some of the CH-π hydrogen bonds. Based on this picture of the first-stage associations, the positive cooperativity with large X-**COR** guests can also be understood. Thus, when the first X-**COR** guest is tilted, a space in the cylinder remains to accommodate the second X-**COR** guest via CH-π contacts. The enthalpy dependence of the ΔH_2 values on the halogen size fits well with this picture. The positive cooperativity is activated by the size mismatch between X-**COR** and **[4]CC**, which is accompanied by the inclination of the guest in the cylinder. Further investigations are necessary to obtain a complete picture of energetics, and other factors, such as π-stack interactions,[26] dispersion forces,[27] halogen-π repulsions[28] and entropy contributions,[29] should be carefully considered.

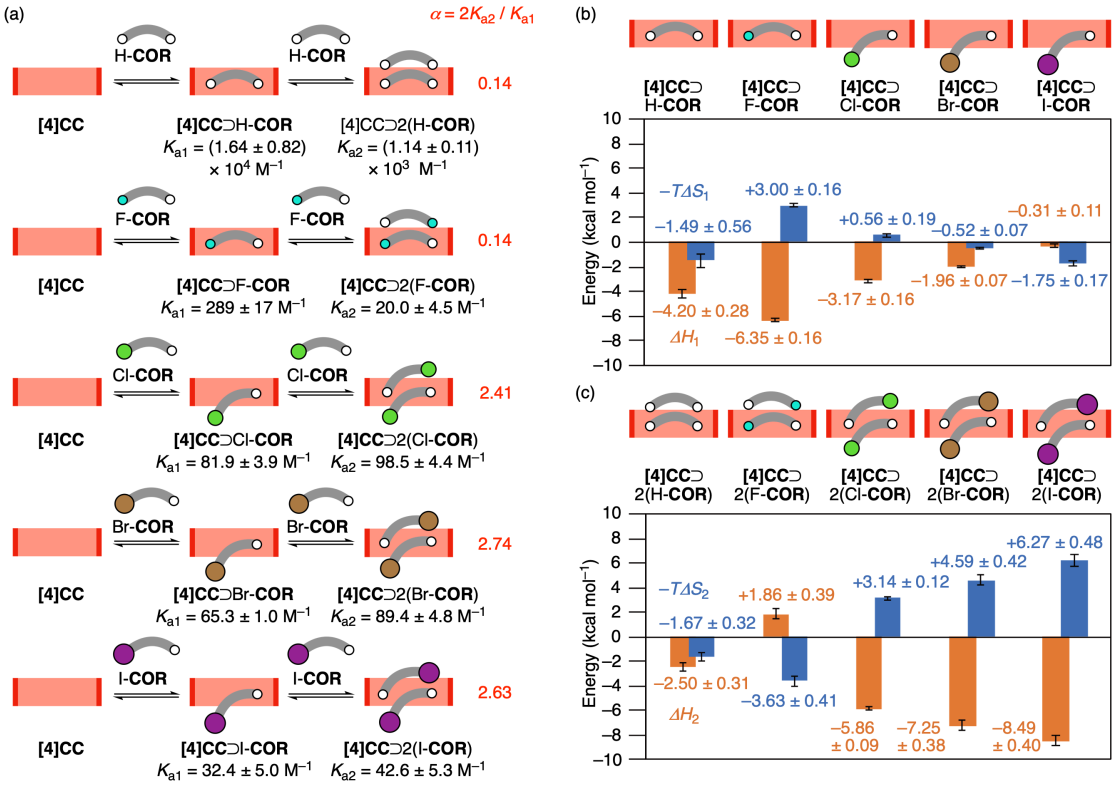


Figure 2.21. Summary of the association of **[4]CC** and X-**COR** including X = H. (a) Association behaviors and cooperativity of **[4]CC** and X-**COR**. (b) Association thermodynamics of the first-stage association. (c) Association thermodynamics of the second-stage association.

2.8 Conclusion

A series of bowl-in-tube complexes were assembled by a cylindrical host with four different halogenated bowl guests. The association stoichiometry was carefully determined to reveal the presence of 1:2 complexes for all the X-**COR** guests. The crystal structures with F-**COR** and Cl-**COR** were revealed to confirm the 1:2 complexation, which also showed the inclination of the guest in the size-mismatch assembly with large halogen substituents. The thermodynamic studies afforded fundamental parameters for the two-stage associations in the one-site host that revealed the positive cooperativity with large X-**COR** guests. As cylindrical shapes are commonly shared, such positive cooperativity may also be present for the encapsulation of guest entities in infinite CNTs. The size-mismatch activation of the positive cooperativity is unique and interesting for the design of various complexes with one-site hosts, and the expansion of structural variations such as heteromeric combinations of guests should be exploited in the future.

2.9 Experimental section

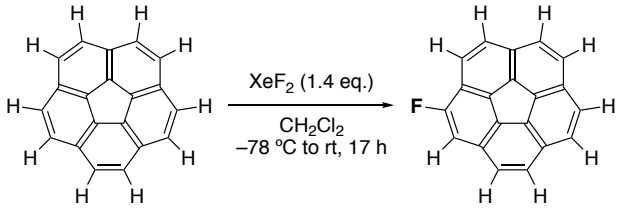
General

Flash silica gel column chromatography was performed on silica gel 60N (spherical and neutral gel, 40-50 μm, Kanto). ^{1}H NMR measurements were conducted on a JEOL RESONANCE JNM-ECAII 600 spectrometer (600 MHz) equipped with an UltraCOOL probe. The chemical shift values (δ) were given in ppm relative to internal CDHCl2 (δ 5.32). High pressure liquid chromatography (HPLC) was performed with cholesterylated columns (COSMOSIL® Cholester, Nacalai Tesque). The analytical conditions of HPLC (JASCO CO2060PLUS) were as follows: COSMOSIL® Cholester = 4.6φ × 250 mm, column temperature = 40 °C, UV-vis detector = 420 nm, CD detector = 300 nm and flow rate = 1.0 mL/min. The preparative HPLC conditions were as follows: COSMOSIL® Cholester = 20φ× 250 mm, column temperature = ambient temperature, UV-vis detector = 420 nm and flow rate = 70 mL/min. The combustion elemental analysis of the specimens was performed on a vario MICRO cube for CHN and Thermo Scientific Dionex ICS-1600 for Cl. UV-vis absorption spectra were recorded on a JASCO V-670 spectrophotometer. ITC analysis was performed using a Malvern MicroCal iTC200 microcalorimeter.

Materials

[4]CC was synthesized and isolated by a method reported in the literature.[30] Corannulene was purchased from Kanto Chemical Co., Inc. F-**COR**, Cl-**COR**, Br-**COR** and I-**COR** were synthesized by a procedure reported in the literatures[31,32,33], but purification procedures were modified to obtain pure materials. CD_2Cl_2, xenon difluoride, gold(III) chloride, iron(III) chloride were purchased from Sigma Aldrich. Dichloromethane, 1,2-dichloroethane, *N*-chlorosuccinimide and *N*-iodosuccinimide were purchased from Wako. Bromine and boron trifluoride diethyl ether complex were from TCI.

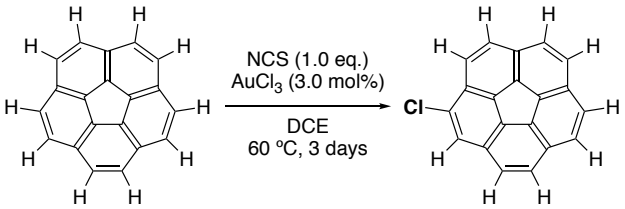
Fluorocorannulene (F-COR)



COR **F-COR** 26%

To a solution of **COR** (401 mg, 1.60 mmol) in dichloromethane (20 mL) was added xenon difluoride (379 mg, 2.24 mmol, 1.4 equiv) at –78 °C. The mixture was allowed to warm to room temperature and stirred for 17 hours. The brown slurry was diluted with dichloromethane and silica gel was added. The solvent was evaporated under reduced pressure and the crude product was purified by flash column chromatography (eluent: 0-5% ethyl acetate/hexane) to obtain a mixture containing F-**COR**, recovered **COR** and polyfluorinated by-products. The mixture was further purified by preparative HPLC (Eluent: 5% chloroform/methanol) to give F-**COR** (113 mg, 26%) as white solid. Elemental analysis revealed that the compound contains solvent molecules: Anal. Calcd for $C_{20}H_9F\cdot(CH_4O)_{0.12}$: C, 88.80; H, 3.51; F, 6.98. Found: C, 88.83; H, 3.81; F, 6.97. ε = 7.05×10^4 M^{-1} cm^{-1} (255 nm). Spectra data were consistent with the reported data.

Chlorocorannulene (Cl-COR)

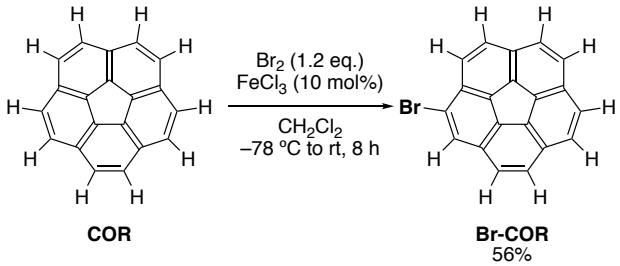


COR **Cl-COR** 51%

The mixture of **COR** (300 mg, 1.20 mmol), *N*-chlorosuccinimide (161 mg, 1.20 mmol) and gold(III) chloride (12.0 mg, 3.0 mol%) were dissolved in 1,2-dichloroethane (17 mL). After stirring for overweekend, the mixture was diluted with dichloromethane and silica gel was added. The solvent was evaporated under reduced pressure and the crude product was purified by flash column chromatography (eluent: 0-5% ethyl acetate/hexane) to obtain a mixture containing Cl-**COR**, recovered **COR** and polychlorinated by-products. The mixture was further purified by preparative HPLC (Eluent: 35% chloroform/methanol) to give Cl-**COR** (175 mg, 51%) as white solid. Anal. Calcd for $C_{20}H_9Cl\cdot(CH_4O)0.2$: C, 83.33; H, 3.39; Cl, 12.18. Found: C, 83.33; H, 3.75; Cl,

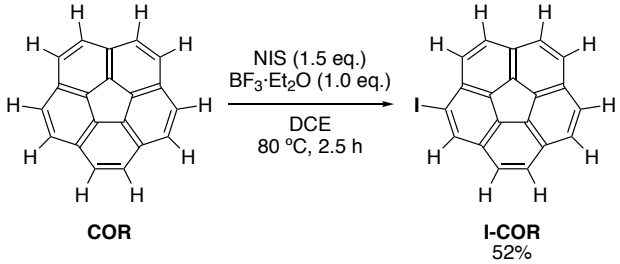
12.32. $\varepsilon = 7.04 \times 10^4$ M^{-1} cm^{-1} (257 nm). Spectra data were consistent with the reported data.

Bromocorannulene (Br-COR)



To a solution of **COR** (300 mg, 1.20 mmol) and iron(III) chloride (21 mg, 0.129 mmol) in dichloromethane (30 mL), was slowly added bromine (0.75 mL, 1.46 mmol) in dichloromethane (9 mL) at –78 °C. The mixture was allowed to warm to room temperature and stirred for 5 hours. The reaction was quenched with a saturated aqueous solution of sodium bisulfate and extracted with dichloromethane (*ca.* 20 mL × 3). The organic layer was washed with water (ca. 30 mL), dried over sodium sulfate and the solvent was evaporated under reduced pressure. The crude product was purified by flash column chromatography (eluent: hexane) to obtain a mixture containing Br-**COR**, recovered **COR** and polybrominated by-products. The mixture was further purified by preparative HPLC (Eluent: 35% chloroform/methanol) to give Br-**COR** (221 mg, 56%) as white solid. Anal. Calcd for $C_{20}H_9Br\cdot(CH_4O)_{0.1}$: C, 72.63; H, 2.85; Br, 24.04. Found: C, 72.80; H, 3.10; Cl, 23.78. $\varepsilon = 6.59 \times 10^4$ M^{-1} cm^{-1} (258 nm). Spectra data were consistent with the reported data.

Iodocorannulene (I-COR)



To a solution of **COR** (300 mg, 1.20 mmol) and *N*-iodosuccinimide (406 mg, 1.80 mmol) in 1,2-dichloroethane (6 mL) were added boron trifluoride diethyl ether complex (0.15 mL, 1.19 mmol) at ambient temperature. After stirring at 80 °C for 2.5 hours, the reaction mixture was diluted with dichloromethane (*ca.* 10 mL). The combined

organic layer was washed with water (*ca.* 20 mL), dried over sodium sulfate and the solvent was evaporated under reduced pressure. The crude product was purified by flash column chromatography (eluent: hexane) to obtain a mixture of I-**COR**, recovered **COR** and polyiodinated by-products. The mixture was further purified by preparative HPLC (Eluent: 35% chloroform/methanol) to give I-**COR** (236 mg, 52%) as white solid. Anal. Calcd for $C_{20}H_9I\cdot(CHCl_3)_{0.02}$: C, 63.32; H, 2.40; I, 33.52; Cl, 0.56. Found: C, 63.59; H, 2.76; Cl, 0.23. $\varepsilon = 4.11 \times 10^4\ M^{-1}\ cm^{-1}$ (254 nm). Spectra data were consistent with the reported data.

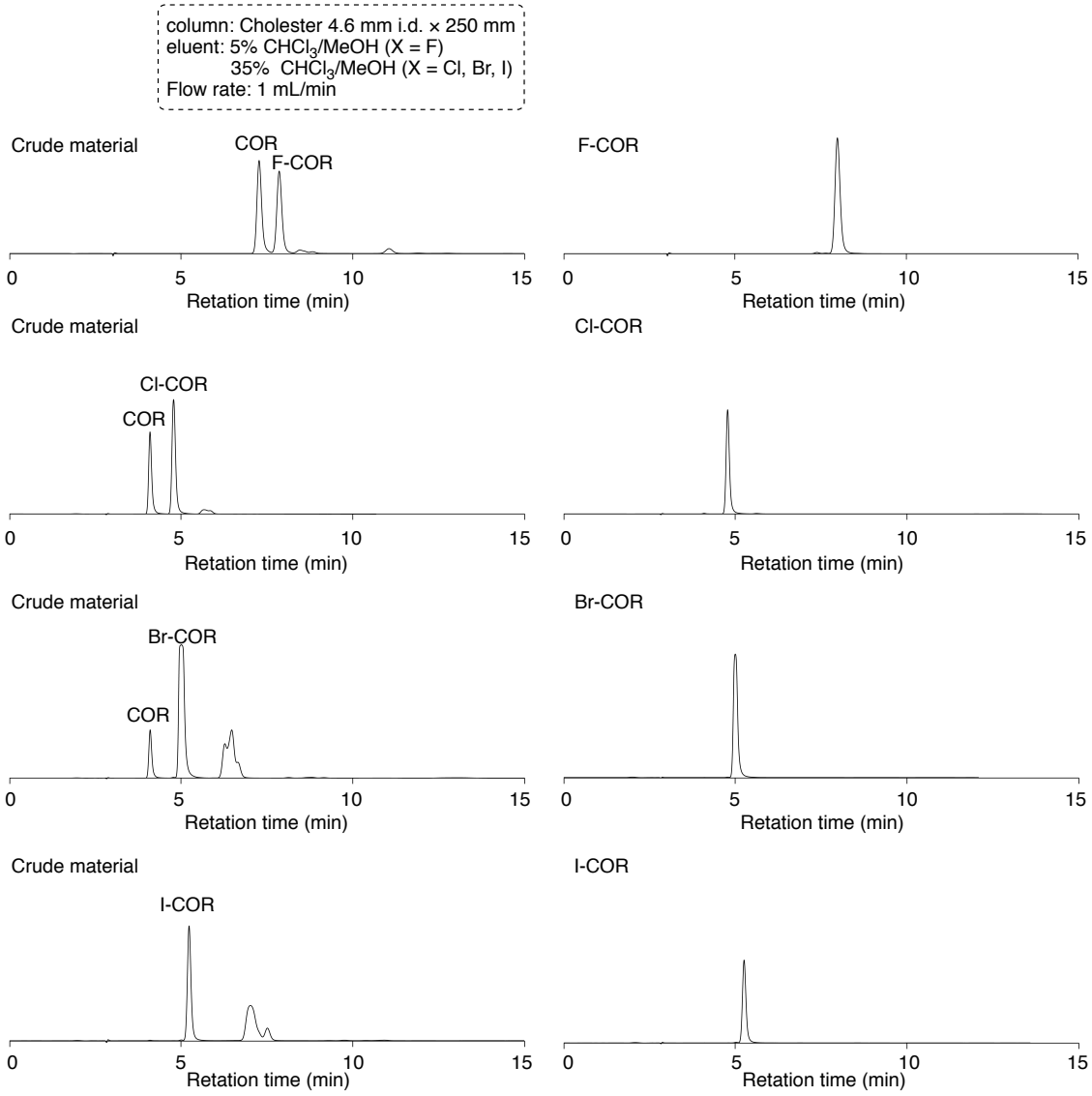


Figure 2.22. Analytical HPLC charts of X-**COR** before and after purification by preparative HPLC.

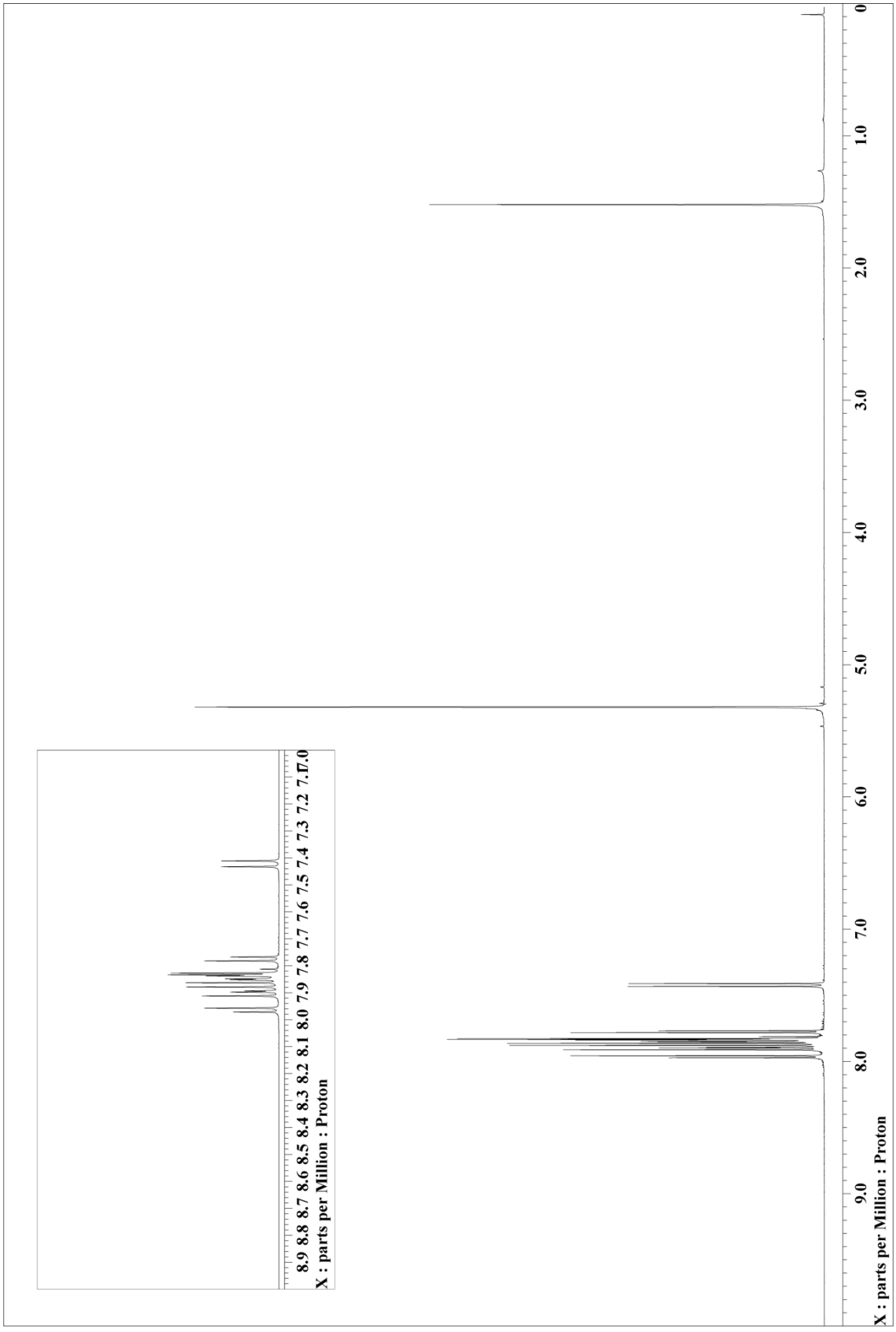


Figure 2.23. ^{1}H NMR spectra of F-**COR**. (600 MHz, CD_2Cl_2, 298 K).

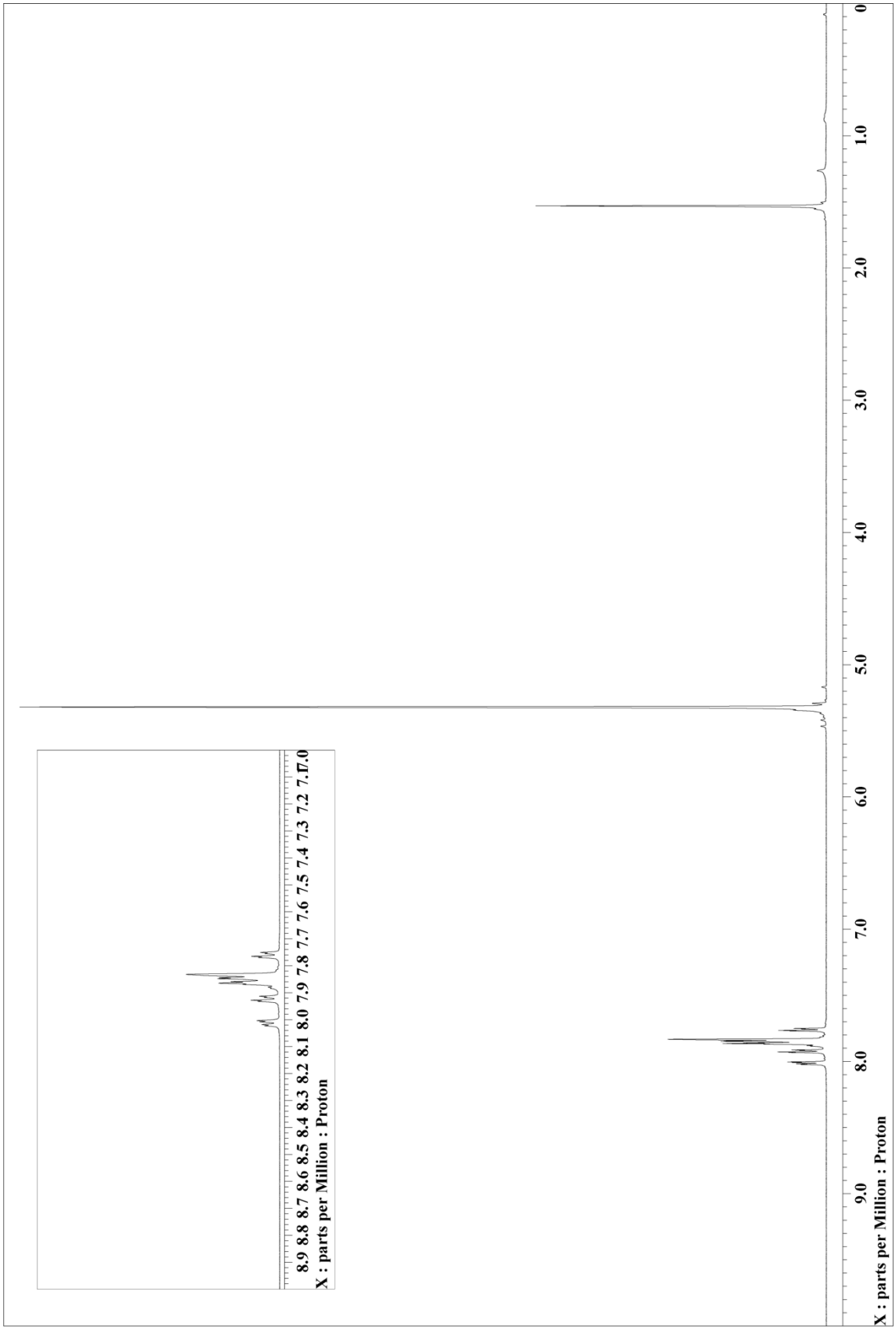


Figure 2.24. ^{1}H NMR spectra of Cl-**COR**. (600 MHz, CD_2Cl_2, 298 K).

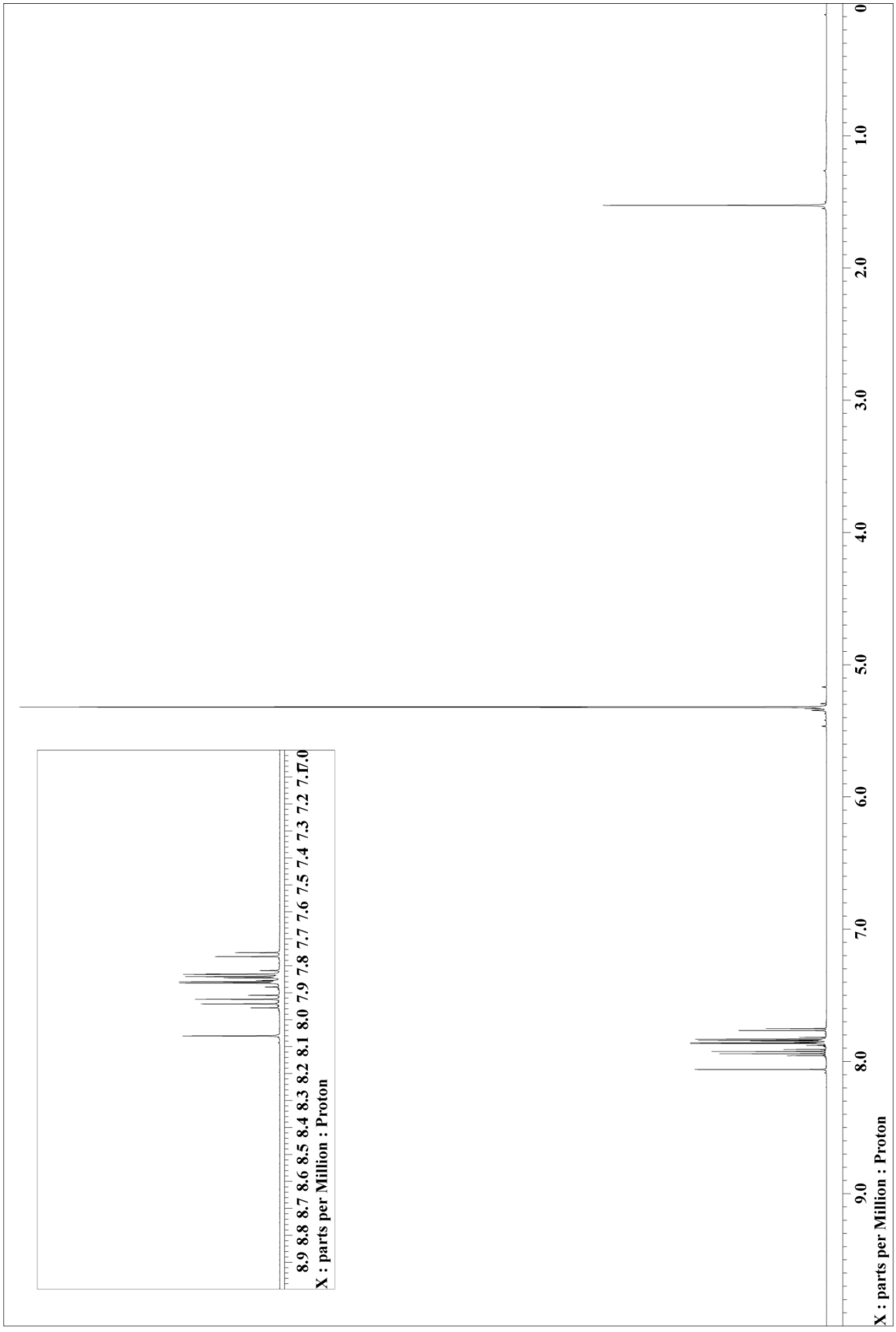


Figure 2.25. ^{1}H NMR spectra of Br-**COR**. (600 MHz, CD_2Cl_2, 298 K).

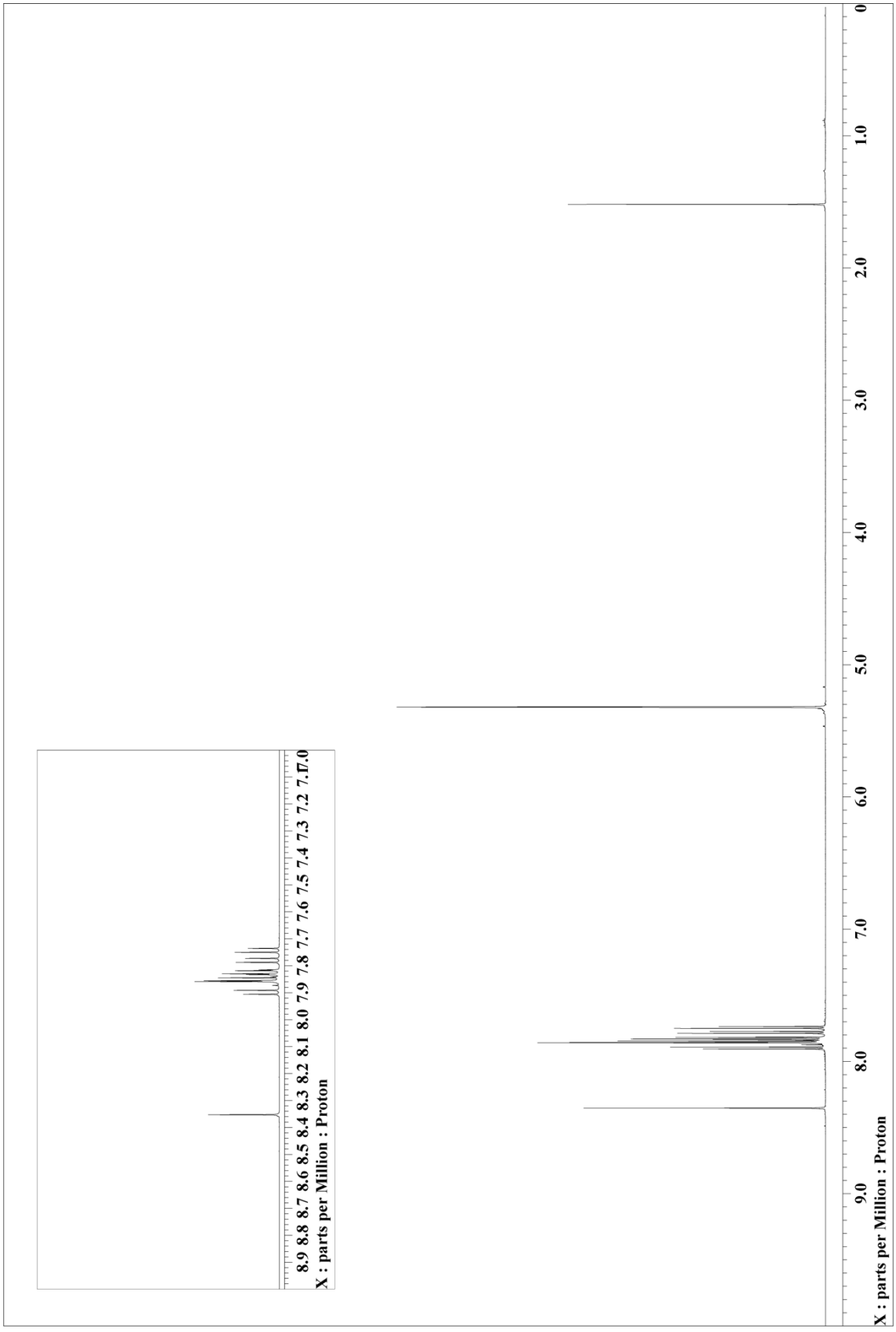


Figure 2.26. ^{1}H NMR spectra of I-**COR**. (600 MHz, CD_2Cl_2, 298 K).

Experimental details

Job plot analysis

Job plot analysis was performed with ^{1}H NMR spectra in CD_2Cl_2 at 298 K. Typical procedures: A series of solutions of **[4]CC** and X-**COR** (total concentration: 10 mM, in CD_2Cl_2) by varying a mole fraction of X-**COR** (χ) to record ^{1}H NMR spectra were collected the data at 298 K. Chemical shift difference $\Delta\delta$ was obtained as the differences in chemical shifts of the four aromatic resonances of **[4]CC** and an aromatic resonance of X-**COR** relative to those of free **[4]CC** and X-**COR**. The product of $\Delta\delta$ and χ was plotted against χ.

ITC titration

ITC in CH_2Cl_2 was performed using a Malvern MicroCal iTC200 microcalorimeter. To a solution of **[4]CC** in a microcalorimeter cell was added a solution of X-**COR** by using the automated titration mode via a syringe. The concentrations of solutions were set as follows: **[4]CC** = 4.86×10^{-4} M and F-**COR** = 9.89×10^{-2} M for the titration of **[4]CC** and F-**COR**, **[4]CC** = 4.89×10^{-4} M and Cl-**COR** = 1.06×10^{-1} M for the titration of **[4]CC** and Cl-**COR**, **[4]CC** = 1.04×10^{-3} M and Br-**COR** = 9.67×10^{-2} M for the titration of **[4]CC** and Br-**COR**, **[4]CC** = 2.20×10^{-3} M and I-**COR** = 8.85×10^{-2} M for the titration of **[4]CC** and I-**COR**. The titration experiments were performed in triplicate for each guest. By using the ORIGIN program provided with the instrument,[34] the titration thermograms were analyzed with 1:1 and 1:2 models. The χ^2 goodness-of-fit indicators were calculated by the software.

Here, procedures of fitting with equations are briefly described for 1:1 and 1:2 models.[15]

1:1. The binding constant K_1 can be expressed by using concentration of a host ([H]), a guest ([G]) and a 1:1 complex ([HG]) as follows:

$$K_1 = \frac{[\mathrm{HG}]}{[\mathrm{H}][\mathrm{G}]} \qquad \text{(Eq. 2.1).}$$

The physical property (Y) that is related to the absolute concentration of each species can be represented as

$$Y = Y_{\mathrm{H}}[\mathrm{H}] + Y_{\mathrm{G}}[\mathrm{G}] + Y_{\mathrm{HG}}[\mathrm{HG}] \qquad \text{(Eq. 2.2),}$$

where Y_{H}, Y_{G} and Y_{HG} are mole-dependent physical properties of pure H, pure G and pure HG, respectively.

When pure G does not contribute to the physical property, one can simplify the equation as follows:

$$Y = Y_{\mathrm{H}}[\mathrm{H}] + Y_{\mathrm{HG}}[\mathrm{HG}] \qquad \text{(Eq. 2.3).}$$

Here, the change of the physical property (ΔY) is considered. ΔY represents the difference of physical properties such as absorbance ΔA in UV, chemical shift $\Delta\delta$ in NMR and heat production Q in ITC.

ΔY is expressed as a subtraction of the initial value (Y_0) from Y and thus is derived as

$$\Delta Y = Y - Y_0 = \left(Y_{\mathrm{H}}[\mathrm{H}] + Y_{\mathrm{HG}}[\mathrm{HG}]\right) - Y_{\mathrm{H}}\left([\mathrm{H}] + [\mathrm{HG}]\right)$$

$$= \left(Y_{\mathrm{HG}} - Y_{\mathrm{H}}\right)[\mathrm{HG}] = Y_{\Delta\mathrm{HG}}[\mathrm{HG}] \qquad \text{(Eq. 2.4).}$$

In this equation, the concentration of HG during a titration experiment is expressed with the total concentration of host ($[\mathrm{H}]_0$) and the total concentration of guest ($[\mathrm{G}]_0$) as follows:

$$[\mathrm{HG}] = \frac{1}{2}\left\{\left([\mathrm{G}]_0 + [\mathrm{H}]_0 + \frac{1}{K_1}\right) + \sqrt{\left([\mathrm{G}]_0 + [\mathrm{H}]_0 + \frac{1}{K_1}\right)^2 - 4[\mathrm{H}]_0[\mathrm{G}]_0}\right\} \qquad \text{(Eq. 2.5).}$$

In the calorimetry, quantity of heat (Q) becomes the corresponding physical property. Thus, the ITC titration data were fitted to an equation,

$$Q = \Delta H_1 V[\mathrm{HG}] = \frac{\Delta H_1 V}{2}\left\{\left([\mathrm{G}]_0 + [\mathrm{H}]_0 + \frac{1}{K_1}\right) + \sqrt{\left([\mathrm{G}]_0 + [\mathrm{H}]_0 + \frac{1}{K_1}\right)^2 - 4[\mathrm{H}]_0[\mathrm{G}]_0}\right\}$$

(Eq. 2.6),

where V is the volume of solution and ΔH_1 is the enthalpy term in the association. In this non-linear fitting analysis, we used the Solver program in Microsoft Excel to obtain the fitting parameters, ΔH_1 and K_1 with R^2 and χ^2 for evaluations of GOF.

1:2. The binding constant K_1 and K_2 can be expressed by using concentration of a host ([H]), a guest ([G]), a 1:1 complex ([HG]) and a 1:2 complex ([HG_2]) as follows:

$$K_1 = \frac{[\mathrm{HG}]}{[\mathrm{H}][\mathrm{G}]} \qquad \text{(Eq. 2.7)}$$

$$K_2 = \frac{[\mathrm{HG_2}]}{[\mathrm{HG}][\mathrm{G}]} \qquad \text{(Eq. 2.8).}$$

In the 1:2 system, the physical property (Y) that is related to the absolute concentration of each species can be represented as

$$Y = Y_{\mathrm{H}}[\mathrm{H}] + Y_{\mathrm{G}}[\mathrm{G}] + Y_{\mathrm{HG}}[\mathrm{HG}] + Y_{\mathrm{HG_2}}[\mathrm{HG_2}] \qquad \text{(Eq. 2.9),}$$

where Y_{H}, Y_{G}, Y_{HG} and Y_{HG2} are mole-dependent physical properties of pure H, pure G, pure HG and pure HG_2 respectively. When pure G does not contribute to the physical property, one can simplify the equation as follows:

$$Y = Y_{\mathrm{H}}[\mathrm{H}] + Y_{\mathrm{HG}}[\mathrm{HG}] + Y_{\mathrm{HG_2}}[\mathrm{HG_2}] \qquad \text{(Eq. 2.10).}$$

The change of the physical property (ΔY) is expressed as a subtraction of the initial value (Y_0) from Y and thus is derived as

$$\Delta Y = Y - Y_0 = \left(Y_{\mathrm{H}}\left[\mathrm{H}\right] + Y_{\mathrm{HG}}\left[\mathrm{HG}\right] + Y_{\mathrm{HG}_2}\left[\mathrm{HG}_2\right]\right) - Y_{\mathrm{H}}\left(\left[\mathrm{H}\right] + \left[\mathrm{HG}\right] + \left[\mathrm{HG}_2\right]\right)$$

$$= \left(Y_{\mathrm{HG}} - Y_{\mathrm{H}}\right)\left[\mathrm{HG}\right] + \left(Y_{\mathrm{HG}_2} - Y_{\mathrm{H}}\right)\left[\mathrm{HG}_2\right] = Y_{\Delta\mathrm{HG}}\left[\mathrm{HG}\right] + Y_{\Delta\mathrm{HG}_2}\left[\mathrm{HG}_2\right] \quad \text{(Eq. 2.11).}$$

[HG] and [HG_2] are represented as the following equations:

$$\left[\mathrm{HG}\right] = \frac{\left[\mathrm{H}\right]_0 K_1\left[\mathrm{G}\right]}{1 + K_1\left[\mathrm{G}\right] + K_1K_2\left[\mathrm{G}\right]^2} \quad \text{(Eq. 2.12).}$$

$$\left[\mathrm{HG}_2\right] = K_2\left[\mathrm{G}\right]\left[\mathrm{HG}\right] = \frac{\left[\mathrm{H}\right]_0 K_1K_2\left[\mathrm{G}\right]^2}{1 + K_1\left[\mathrm{G}\right] + K_1K_2\left[\mathrm{G}\right]^2} \quad \text{(Eq. 2.13).}$$

Using (eq. 2.11), (eq. 2.12) and (eq. 2.13), one can obtain

$$\Delta Y = \frac{Y_{\Delta\mathrm{HG}}\left[\mathrm{H}\right]_0 K_1\left[\mathrm{G}\right] + Y_{\Delta\mathrm{HG}_2}\left[\mathrm{H}\right]_0 K_1K_2\left[\mathrm{G}\right]^2}{1 + K_1\left[\mathrm{G}\right] + K_1K_2\left[\mathrm{G}\right]^2} \quad \text{(Eq. 2.14).}$$

In this equation, the concentration of G during a titration experiment is obtained by solving the following cubic equation,

$$K_1K_2\left[\mathrm{G}\right]^3 + \left(2K_1K_2\left[\mathrm{H}\right]_0 - K_1K_2\left[\mathrm{G}\right]_0 + K_1\right)\left[\mathrm{G}\right]^2 + \left(K_1\left[\mathrm{H}\right]_0 - K_1\left[\mathrm{G}\right]_0 + 1\right)\left[\mathrm{G}\right] - \left[\mathrm{G}\right]_0 = 0$$

(Eq. 2.15).

In the calorimetry, quantity of heat (Q) becomes the corresponding physical property. Thus, the ITC titration data were fitted to an equation,

$$Q = \frac{\Delta H_1 V\left[\mathrm{H}\right]_0 K_1\left[\mathrm{G}\right] + \left(\Delta H_1 + \Delta H_2\right)V\left[\mathrm{H}\right]_0 K_1K_2\left[\mathrm{G}\right]^2}{1 + K_1\left[\mathrm{G}\right] + K_1K_2\left[\mathrm{G}\right]^2} \quad \text{(Eq. 2.16),}$$

where V is the volume of solution, and ΔH_1 and ΔH_2 are the enthalpy term in the first and second association, respectively. In this non-linear fitting analysis, we used the Solver program in Microsoft Excel to obtain the fitting parameters, ΔH_1, ΔH_2, K_1 and K_2 with R^2 and χ^2 for evaluations of GOF.

Fitting analyses of ITC data with cov_{fit} indicators

The cov_{fit} goodness-of-fit indicators were manually obtained as follows. The ITC data were obtained from the ORIGIN program in csv format. The values of experimental heat (y_{data}), calculated heat (y_{calc}) and the residual ($y_{fit} = y_{data} - y_{calc}$) at each titration data point were extracted from the csv file. The variance of experimental values, cov(y_{data}), and the variance of the residual, cov(y_{fit}), were calculated by using Microsoft Excel.[35] The cov_{fit} values were then obtained as the ratio of two variances as shown below.[15]

$$\mathrm{cov}_{\mathrm{fit}} = \frac{\mathrm{cov}\left(y_{\mathrm{fit}}\right)}{\mathrm{cov}\left(y_{\mathrm{data}}\right)} \qquad \text{(Eq. 2.17)}$$

Comparisons of models for ITC fittings by F-test

The statistical comparison test (*F*-test) was performed for the comparison of 1:1 and 1:2 models for the fitting analyses. With the data in csv format, *F* values were calculated by[23]

$$F = \frac{\left(SS_1 - SS_2\right)/SS_2}{\left(df_1 - df_2\right)/df_2} \qquad \text{(Eq. 2.18),}$$

where SS_1 = sum of squared residuals for 1:1 fitting, SS_2 = sum of squared residuals for 1:2 fitting, df_1 = degrees of freedom for 1:1 fitting, df_2 = for degrees of freedom for 1:2 models. The SS_1 and SS_2 were obtained as the summation of squared residuals y_{fit}^2. The df_1 and df_2 were obtained by subtracting the numbers of parameters (2 for 1:1 model or 4 for 1:2 models) from the number of titration data points (54, 38, 18, 18 and 18 for H-**COR**, F-**COR**, Cl-**COR**, Br-**COR** and I-**COR**, respectively). Therefore, the values of df_1 were 52, 36, 16, 16 and 16 for H-**COR**, F-**COR**, Cl-**COR**, Br-**COR** and I-**COR**, respectively. The values of df_2 were 50, 34, 14, 14 and 14 for H-**COR**, F-**COR**, Cl-**COR**, Br-**COR** and I-**COR**, respectively. The *P* values for calculated *F* values are calculated by using FDIST function of Microsoft Excel.[35]

Comparisons of models for ITC fittings by residual plot analyses

The residual plot analyses[23] were also performed for the comparison of 1:1 and 1:2 models for the fitting analyses. With the data in csv format, the residual y_{fit} values

were plotted against the equivalence of added guest [X-**COR**]/[**[4]CC**].

Crystallographic analyses

Single crystals of 1:2 complexes suitable for X-ray analyses were obtained as follows. Crystals of **[4]CC**⊃2(F-**COR**) was obtained by slow evaporation from a 1:2 mixture of **[4]CC** and F-**COR** in dichloromethane/2-propanol (1:1 vol/vol) at 25 °C. Crystals of **[4]CC**⊃2(Cl-**COR**) was obtained by vapor diffusion method from a 1:10 mixture of **[4]CC** and Cl-**COR** in dichloromethane with acetonitrile as a poor solvent at 25 °C. A single crystal was mounted on a thin polymer tip with cryoprotectant oil and frozen *via* flash cooling (**[4]CC**⊃2(F-**COR**)) or sealed in a glass capillary (**[4]CC**⊃2(Cl-**COR**)). The diffraction analyses of a single crystal with synchrotron X-ray sources were conducted at 95 K (**[4]CC**⊃2(F-**COR**)) or 298 K (**[4]CC**⊃2(Cl-**COR**)) by using a diffractometer equipped with a Dectris EIGER X 16M PAD detector at KEK BL17A. The diffraction data were processed with the XDS software program.[36] The structures were solved by a direct method with SIR2014 (**[4]CC**⊃2(F-**COR**)) [37] or SHELXD (**[4]CC**⊃2(Cl-**COR**)) program.[38] and refined by full-matrix least-squares on F^2 using the SHELXL-2014/7 program suite [39] running with the Yadokari-XG 2009 software program.[40] In the refinements, alkyl groups and solvent molecules were restrained by SIMU, DFIX and DANG. F-**COR** and Cl-**COR** molecules were treated as rigid bodies with AFIX6 instructions. Non-hydrogen atoms were analyzed anisotropically and hydrogen atoms were input at the calculated positions and refined with a riding model. In the refinements, the electron density attributed to solvent molecules was not properly modelled due to the severe disorders, and the structures were refined without these solvents by the PLATON/SQUEREZE technique.[41] Crystal of **[4]CC**⊃2(F-**COR**) and **[4]CC**⊃2(Cl-**COR**) contained two inequivalent complexes. One representative complex is shown in Figure 2.19, and the other complex is shown in Figure 2.27. Crystal data are summarized in Tables 2.3 and 2.4.

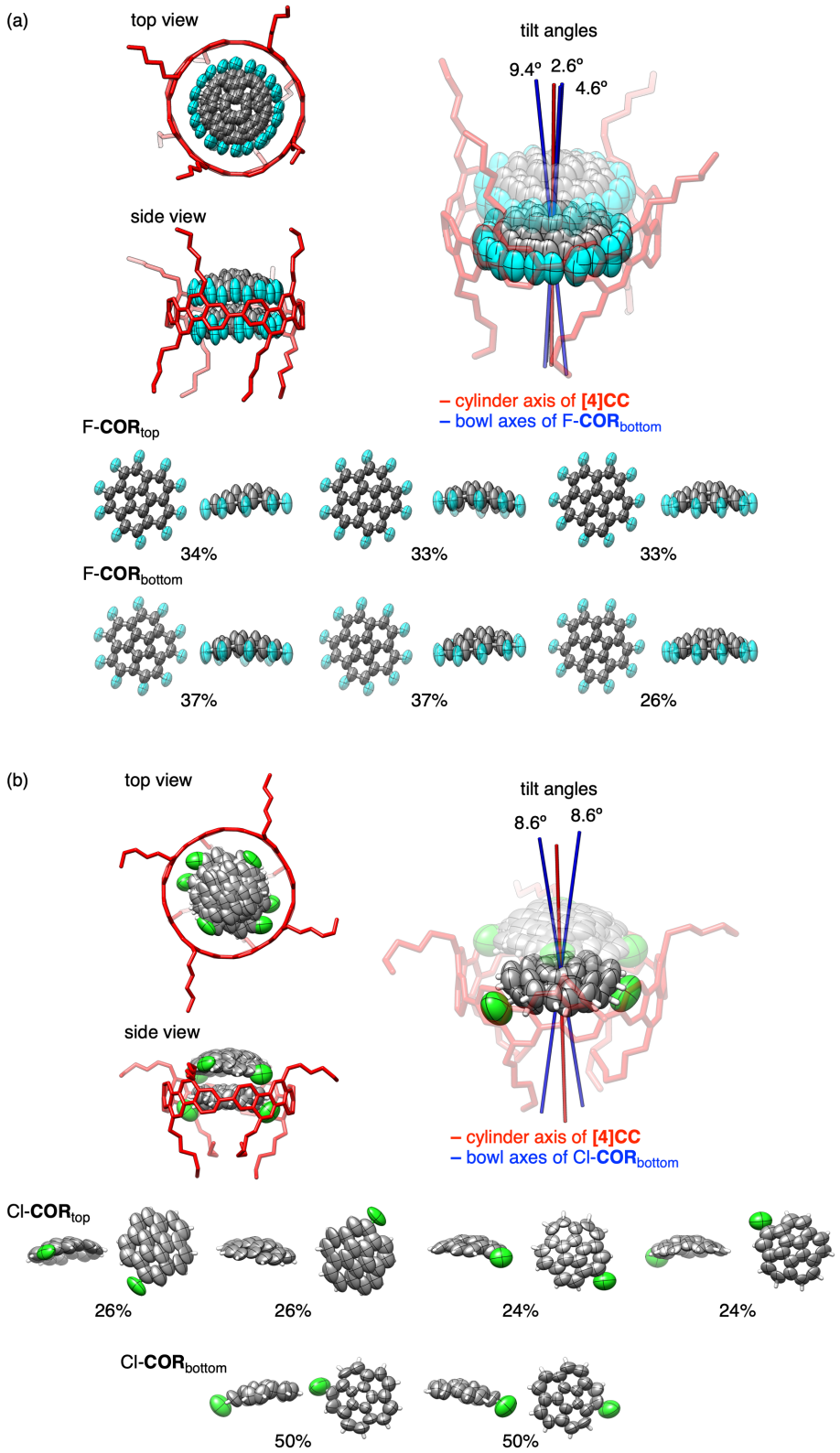


Figure 2.27. Crystal structures of 1:2 complexes. Thermal ellipsoids of X-**COR** are shown in 50% probability. Hydrogen atoms of **[4]CC** are omitted for clarity. (a) Structures of **[4]CC**⊃2(F-**COR**). (b) Structures of **[4]CC**⊃2(Cl-**COR**).

Crystallographic Analysis

Table 2.3. Crystal data and structure refinement for **[4]CC⊃2(F-COR)**.

CCDC No.	2142996
Empirical formula	$C_{160}\ H_{154}\ F_2$
Formula weight	2114.82
Temperature	95(2) K
Wavelength	0.9000 Å
Crystal system	Trigonal
Space group	$P\ 3_2\ 2\ 1$
Unit cell dimensions	a = 28.390(4) Å a = 90°.
	b = 28.390(4) Å b = 90°.
	c = 64.060(13) Å g = 120°.
Volume	44714(15) Å^3
Z	12
Density (calculated)	0.942 Mg/m^3
Absorption coefficient	0.091 mm^{-1}
F(000)	13584
Crystal size	0.100 × 0.100 × 0.050 mm^3
Theta range for data collection	1.123 to 28.299°.
Index ranges	–29<=h<=29, –29<=k<=29, –67<=l<=67
Reflections collected	359555
Independent reflections	36395 [R(int) = 0.0333]
Completeness to theta = 28.299°	99.7 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	1.000 and 0.693
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	36395 / 34764 / 4819
Goodness-of-fit on F^2	1.480
Final R indices [I>2sigma(I)]	R_1 = 0.1197, wR_2 = 0.3234
R indices (all data)	R_1 = 0.1378, wR_2 = 0.3527
Flack parameter	0.9(2)
Extinction coefficient	n/a
Largest diff. peak and hole	0.516 and –0.359 e.Å^{-3}

Table 2.4. Crystal data and structure refinement for **[4]CC⊃**2(Cl-**COR**).

CCDC No.	2142997
Empirical formula	$C_{160}\ H_{154}\ Cl_2$
Formula weight	2147.72
Temperature	298(2) K
Wavelength	0.9000 Å
Crystal system	Orthorhombic
Space group	$P\ 2_1\ 2_1\ 2$
Unit cell dimensions	a = 29.960(6) Å a = 90°.
	b = 30.410(6) Å b = 90°.
	c = 21.140(4) Å g = 90°.
Volume	19260(7) Å^3
Z	6
Density (calculated)	1.111 Mg/m^3
Absorption coefficient	0.182 mm^{-1}
F(000)	6888
Crystal size	0.050 × 0.050 × 0.020 mm^3
Theta range for data collection	1.208 to 28.330°.
Index ranges	–31<=h<=31, –32<=k<=32, –22<=l<=22
Reflections collected	152331
Independent reflections	23610 [R(int) = 0.0950]
Completeness to theta = 28.330°	99.8 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	1.000 and 0.512
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	23610 / 2454 / 2567
Goodness-of-fit on F^2	1.257
Final R indices [I>2sigma(I)]	R_1 = 0.1542, wR_2 = 0.3657
R indices (all data)	R_1 = 0.2046, wR_2 = 0.4083
Flack parameter	0.14(9)
Extinction coefficient	n/a
Largest diff. peak and hole	0.337 and –0.485 e.Å^{-3}

2.10 References

1. B. W. Smith, M. Monthioux, D. E. Luzzi, *Nature* **1998**, *396*, 323-324.
2. M. Monthioux, *Carbon* **2002**, *40*, 1809-1823.
3. I. V. Krive, R. I. Shekhter, M. Jonson, *Low Temp. Phys.* **2006**, *32*, 887-905.
4. S. Iijima, *Physica B* **2003**, *323*, 1-5.
5. A. de Juan, E. M. Perez, *Nanoscale* **2013**, *5*, 7141-7148.
6. T. Okazaki, Y. Iizumi, S. Okubo, H. Kataura, Z. Liu, K. Suenaga, Y. Tahara, M. Yudasaka, S. Okada, S. Iijima, *Angew. Chem. Int. Ed.* **2011**, *50*, 4853-4857.
7. Y. Izumi, Z. Liu, K. Suenaga, S. Okada, S. Higashibayashi, H. Sakurai, T. Okazaki, *ChemNanoMat* **2018**, *4*, 557-561.
8. Y. Joko, R. Sasaki, K. Shintani, *Phys. Chem. Chem. Phys.* **2017**, *19*, 27704-27715.
9. T. Iwamoto, Y. Watanabe, T. Sadahiro, T. Haino, S. Yamago, *Angew. Chem. Int. Ed.* **2011**, *50*, 8342-8344.
10. Z. Sun, T. Matsuno, H. Isobe, *Bull. Chem. Soc. Jpn.* **2018**, *91*, 907-921.
11. T. Matsuno, M. Fujita, K. Fukunaga, S. Sato, H. Isobe, *Nat. Commun.* **2018**, *9*, 3779.
12. T. Matsuno, K. Fukunaga, S. Sato, H. Isobe, *Angew. Chem. Int. Ed.* **2019**, *58*, 12170-12174.
13. M. F. Perutz, *Q. Rev. Biophys.* **1989**, *22*, 139-236.
14. C. A. Hunter, H. L. Anderson, *Angew. Chem. Int. Ed.* **2009**, *48*, 7488-7499.
15. P. Thordarson, *Chem. Soc. Rev.* **2011**, *40*, 1305-1323.
16. L. K. S. von Krbek, C. A. Schalley, P. Thordarson, *Chem. Soc. Rev.* **2017**, *46*, 2622-2637.
17. Z. Huang, X. Chen, G. Wu, P. Metrangolo, D. Whitaker, J. A. McCune, O. A. Scherman, *J. Am. Chem. Soc.* **2020**, *142*, 7356-7361.
18. Z. Huang, K. Qin, G. Deng, G. Wu, Y. Bai, J.-F. Xu, Z. Wang, Z. Yu, O. A. Scherman, X. Zhang, *Langmuir* **2016**, *32*, 12352-12360.
19. X. Shi, X. Zhang, X.-L. Ni, H. Zhang, P. Wei, J. Liu, H. Xing, H.-Q. Peng, J. W. Y. Lam, P. Zhang, Z. Wang, H. Hao, *Macromolecules* **2019**, *52*, 8814-8825.
20. H. Barbero, E. Masson, *Chem. Sci.* **2021**, *12*, 9962-9968.
21. (a) L. M. Heitmann, A. B. Taylor, P. J. Hart, A. R. Urbach, *J. Am. Chem. Soc.* **2006**, *128*, 12574-12581. (b) E. Cavatorta, P. Jonkheijm, J. Huskens, *Chem. Eur. J.* **2017**, *23*,

4046-4050.

22. A. Bondi, *J. Phys. Chem.* **1964**, *68*, 441-451.

23. (a) F. Ulatowski, K. Dabrowa, T. Balakier, J. Jurczak, *J. Org. Chem.* **2016**, *81*, 1746-1756. (b) D. B. Hibbert, P. Thordarson, *Chem. Commun.* **2016**, *52*, 12792-12805.

24. It was also found that the stoichiometry for the H-**COR** guest should be revised from 1:1 to 1:2 through the same careful analyses.

25. C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. A. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler, P. A. Wood, *J. Appl. Cryst.* **2020**, *53*, 226-235.

26. (a) S. L. Cockroft, C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch, *J. Am. Chem. Soc.* **2005**, *127*, 8594-8595. (b) H. Isobe, S. Hitosugi, T. Matsuno, T. Iwamoto, J. Ichikawa, *Org. Lett.* **2009**, *11*, 4026-4028.

27. (a) C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch, *J. Chem. Soc.-Perkin Trans.* **2001**, *2*, 651-669. (b) H. Isobe, K. Nakamura, S. Hitosugi, S. Sato, H. Tokoyama, H. Yamakado, K. Ohno, H. Kono, *Chem. Sci.* **2015**, *6*, 2746-2753. (c) H. Isobe, K. Nakamura, S. Hitosugi, S. Sato, H. Tokoyama, H. Yamakado, K. Ohno, H. Kono, *Chem. Sci.* **2016**, *7*, 2929-2932.

28. H. Adams, S. L. Cockroft, C. Guardigli, C. A. Hunter, K. R. Lawson, J. Perkins, S. E. Spey, C. J. Urch, R. Ford, *ChemBioChem* **2004**, *5*, 657-665.

29. Entropy contributions might well be complicated by possible involvements of host-guest/guest-guest relationships, and, moreover, desolvation at contorted π-systems should also be considered. For example, see (a) K. Ikemoto, R. Kobayashi, S. Sato, H. Isobe, *Angew. Chem. Int. Ed.* **2017**, *56*, 6511-6514. (b) K. Ikemoto, J. Lin, R. Kobayashi, S. Sato, H. Isobe, *Angew. Chem. Int. Ed.* **2018**, *57*, 8555-8559.

30. S. Hitosugi, W. Nakanishi, T. Yamasaki, H. Isobe, *Nat. Commun.* **2011**, *2*, 492.

31. B. M. Schmidt, B. Topolinski, M. Yamada, S. Higashibayashi, M. Shionoya, H. Sakurai, D. Lentz, *Chem. Eur. J.* **2013**, *19*, 13872-13880.

32. D. Wu, T. Shao, J. Men, X. Chen, G. Gao, *Dalton Trans.* **2014**, *43*, 1753-1761.

33. Y. Wang, O. Allemann, T. S. Balaban, N. Vanthuyne, A. Linden, K. K. Baldridge, J. S. Siegel, *Angew. Chem. Int. Ed.* **2018**, *57*, 6470-6474.

34. Origin 7, OriginLab Corporation, Northampton (USA), **2002**.

35. Microsoft Excel for Mac (v. 16.57), Microsoft Corporation, Redmond (USA), **2021**.
36. W. Kabsch, *J. Appl. Cryst.* **1993**, *26*, 795-800.
37. M. C. Burla, R. Caliandro, M. Camalli, B. Carrozzini, G. L. Cascarano, L. De Caro, C. Giacovazzo, G. Polidori, D. Siliqi, R. Spagna, *J. Appl. Cryst.* **2007**, *40*, 609-613.
38. G. M. Sheldrick, *Acta Crystallogr. A* **2008**, *64*, 112-122.
39. G. M. Sheldrick, T. R. Schneider, *Methods in Enzymology, Part B, Macromolecular Crystallorgaphy*, Academic Press, Cambridge, MA, **1997**, Vol. 277, pp. 319-343.
40. C. Kabuto, S. Akine, T. Nemoto, K. Kwon, *J. Crystallogr. Soc. Jpn.* **2009**, *51*, 218-224.
41. A. L. Spek, *J. Appl. Cryst.* **2003**, *36*, 7-13; P. van der Sluis, A. L. Spek, A. L. *Acta Crystallogr. A* **1990**, *46*, 194-201.

Chapter 3. Introduction of Akaike's Information Criterion after the Death of the Job Plot

3.1 Introduction

Defining and measuring equilibria are two of the most important steps in supramolecular chemistry.[1,2] Studies on equilibria expand the scope of supramolecular chemistry to interdisciplinary fields such as chemical biology and sensing technology and deepen our understanding of the underlying chemistry of noncovalent interactions. Equilibria leading to host-guest complexes are commonly analyzed by titration experiments, which track changes in physical properties (ΔY), and essential physical parameters such as association constants (K_n) are determined to define the equilibria with physical measures (Figure 3.1).[3] Although the analytical methods related to supramolecular host-guest chemistry may seem mature, one of the most popular methods has recently been questioned. Thus, the Job plot,[4] which is a method to determine the stoichiometry of the host-guest complex, was proclaimed dead upon being revisited in 2016.[5] In the "antemortem" era, the Job plot was adopted as a method of continuous variables to define the stoichiometry of host-guest complexes. The Job plot analyses allowed us to decide on one equilibrium model with a specific equilibrium equation for fitting analyses (Figure 3.1).[6] Consequently, fitting the equation with titration data afforded the K_n values and relevant thermodynamic parameters with a goodness-of-fit (GOF) measure for fitting evaluations. However, in the classical approach, stoichiometry is not defined at the initial stage, and multiple possible models must be examined. The titration data are fitted with multiple equilibrium equations for all possible models, such as the 1:1, 1:2 or 2:1 model. After the GOF evaluation of the fitting, the fitting results for each model are compared, which eventually affords the K_n values from the equilibrium equation of the best model. This transition from the classical approach to the modern approach is not unique to supramolecular chemistry. The critical question of "Which model is the best to use?" has been ubiquitously asked and examined in various scientific fields, such as statistics and biology.[7] A null hypothesis testing approach using the *F*-test with the *P*-value from statistics was adopted in various scientific fields and introduced to the field of host-guest chemistry,[5,8,9] and an alternative information theoretic approach, called Akaike's information criterion (AIC), was invented.[10] Herein, this AIC method is

applied to host-guest chemistry to identify the best model of equilibria. This method can also rank the models by using the weight of evidence as quantitative measures. To the best of my knowledge, this study is the first example to apply AIC measures to the supramolecular chemistry of host-guest complexes and should be informative for relevant works performed in the future.

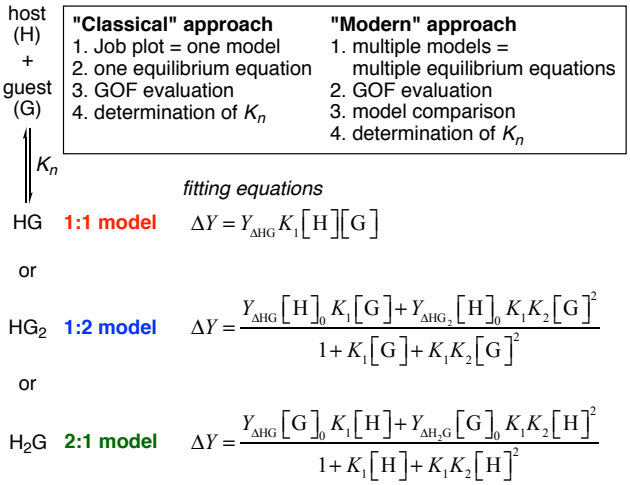


Figure 3.1. Determination of the association constants (K_n). Procedures altered before and after the death of the Job plot.[5] GOF: goodness of fit; Y: physical property; ΔY: change in Y; K_n: association constant.

3.2 Problem of stoichiometry determination

Due to their segmental molecular design,[11,12] the study of the supramolecular chemistry of carbon nanotubes is currently flourishing.[13] The cylindrical structures of the segments have unique environments surrounded by curved π-systems,[14] which also provide interesting cases to investigate the stoichiometry of host-guest complexes using the modern approach. A host molecule is normally defined as a molecular entity with convergent binding sites,[15] but the cylindrical structures of nanotubes and their segments can have one-dimensional structural divergency along the cylindrical axis. Consequently, multiple guest molecules can be encapsulated in the cylindrical host,[13] which complicates the model selections for the equilibria. In previous studies with (*P*)-(12,8)-[4]cyclochrysenylene (**[4]CC**) as a host and corannulene as a guest, a 1:1 model was initially concluded via classical Job plot analyses (see Figure 2.1 in Chapter 2),[16] but

modern analyses via statistical comparisons of models with the *F*-test/*P*-value revealed the superiority of a 1:2 model that was also observed as a crystal structure.[17] In this work, the relevant host-guest systems of previous studies were scrutinized, and the survey found interesting systems to investigate that may further develop modern analyses. The first set of examples adopted dumbbell-shaped guests of fullerene dimers.[18,19,20,21] Previously, C_{120} was examined as the dumbbell-shaped guest for cylindrical hosts of **[4]CC** and its lengthened congener, [4]cycloanthanthrenylene (**[4]CA**). A series of 2:1 host-guest complexes was concluded via classical analyses with the Job plot in previous studies,[18,19] and a similar dumbbell-shaped guest of $C_{118}N_2$ was likewise concluded by classical methods to be a 2:1 complex with **[4]CC** host.[20] However, considering the unique structural divergency of the host and guest,[15] it was noticed that the 1:1, 1:2 and 2:1 models should be examined using modern analyses of dumbbell-and-cylinder complexes. Based on three fitting equations, these systems should provide an interesting choice for the best model.[7] The fitting equations for the three models have two different relationships (Figure 3.2). First, the 1:1 model and 1:2 model are so-called nested, *i.e.*, one model (*e.g.,* 1:1) is a simpler case of the other (*e.g.,* 1:2).[6,7] This relationship is also found in the fitting equations of the 1:1 model and 2:1 model, and these models are also nested. However, the 1:2 model and 2:1 model are not nested, as the fitting equations in Figures 3.1 and 3.2 show. Unnested models cannot be compared by null hypothesis testing with the *F*-test/*P*-value,[7] and an alternative method is needed.

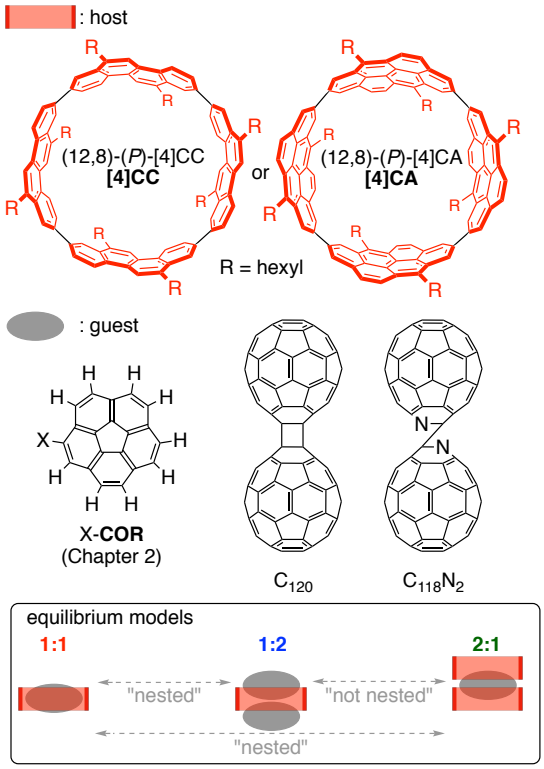


Figure 3.2. Host-guest chemistry of cylindrical hosts.

In this study, a measure from the information-theory approach was adopted, and the three models, 1:1, 1:2 and 2:1, were compared by AIC and Akaike weights (w_i).[7,10] AIC and Akaike weights can be calculated by the following equations. (Eq. 3.1, 3.2)

$$\mathrm{AIC} = N \ln\left(\frac{SS}{N}\right) + 2k \qquad \text{(Eq. 3.1)},$$

where N is the number of data points, SS is the sum of the squared residuals, and k is the number of parameters. The first term with SS and N tends to decrease as more parameters are added to the model, whereas the second term with k increases as more parameters are added to the model. Accordingly, the trade-off between underfitting and overfitting is estimated by AIC, with a smaller AIC inferring a better model than a model with a larger AIC. The relative likelihood of a model is also interpreted by using the Akaike weight. The Akaike weight is considered the weight of evidence for model i and is scaled to sum to 1 in total for all the models. The Akaike weight (w_i) is defined as

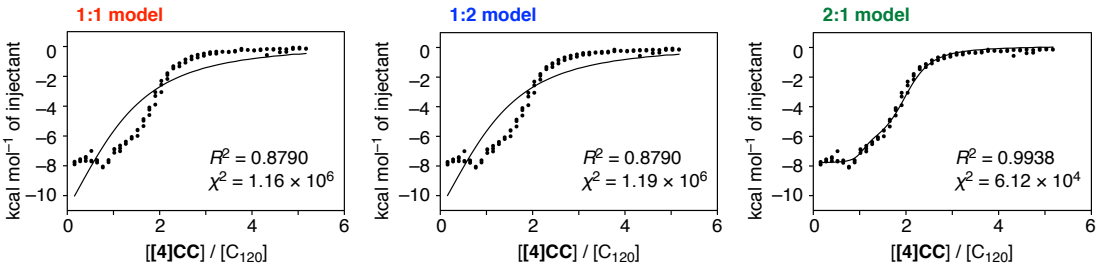
$$w_i = \frac{\exp\left(\left(\mathrm{AIC}_{\min} - \mathrm{AIC}_i\right)/2\right)}{\sum_{i=1}^{M} \exp\left(\left(\mathrm{AIC}_{\min} - \mathrm{AIC}_i\right)/2\right)} \qquad \text{(Eq. 3.2)},$$

where M is the number of models in the comparison, AIC_{min} is the smallest value of AIC recorded in the best model, and AIC_i is the AIC value of the model i of interest. Based on these equations, analyses were conducted for the supramolecular complexes of **[4]CC** and **[4]CA**

3.3 Dumbbell-and-cylinder systems

Previous classical analyses of the dumbbell-and-cylinder complexes utilized a combination of fluorescence quenching titration and isothermal titration calorimetry (ITC) experiments.[18,19,20] For the modern analyses in the present study, I noticed that ITC data were superior for global fitting analyses of complexation processes that potentially involved two-stage associations,[3,22,23] and fitting analyses were performed with the ITC data of three dumbbell-and-cylinder systems (**[4]CC**+C_{120}, **[4]CA**+C_{120} and **[4]CC**+$C_{118}N_2$) that were originally investigated for the classical analyses.[18,19,20] The results are summarized in Figures 3.3-3.5. After fitting with three models (**1:1**, **1:2** and **2:1**), GOF was first evaluated using R^2 and χ^2 values (Table 3.1). It was found that the R^2 values were intuitively superior to determine the GOF level, and most models for the three dumbbell-and-cylinder systems showed a sufficient GOF (>0.9). Although the GOF levels were not recommended for the model comparisons,[5,7] the nonscaled values of χ^2 could falsely cause such misuses. To compare three models (**1:1**, **1:2** and **2:1**), AIC was introduced. The theoretical basis of AIC is complicated, since it is composed of a combination of the maximum likelihood theory, information theory and the concept of entropy of information;[10] however, computations of AIC can be easily performed using the following equation: $AIC = N \log (SS/N) + 2k$, where N is the number of data points, SS is the sum of the squared residuals and k is the number of parameters.[7] The AIC values should be compared within an identical system, *e.g.*, among three models of **[4]CC**+C_{120}. The AIC values rank these models in the order of **2:1**, **1:1** and **1:2**, and the smallest value of 1294 for **2:1** quantitatively shows that it is the best model. The other two dumbbell-and-cylinder systems were likewise analyzed by AIC, which shows that **2:1** serves as the best model for these systems. Another measure of AIC, *i.e.*, Akaike weight (w_i), was also found to be useful for model comparison. Similar to the R^2 value, the w_i values were scaled to sum to 1, which can be interpreted as the weight of evidence for each model.[7,10] The w_i values of 1 for **2:1** and negligible w_i values for **1:1**/**1:2** of all dumbbell-and-

cylinder complexes confirmed that **2:1** was the best model for these systems (Table 3.1).



model	*N*	*k*	K_1 (M^{-1})	K_2 (M^{-1})	ΔH_1 (kcal mol^{-1})	ΔH_2 (kcal mol^{-1})	*SS*	AIC	w_i
1:1	117	2	1.44×10^4	–	–18.8	–	1.34×10^8	1636	3.5×10^{-75}
1:2	117	4	1.41×10^4	2.31×10^{-2}	–18.8	–2.6	1.34×10^8	1640	4.7×10^{-76}
2:1	117	4	1.16×10^7	1.30×10^5	–7.8	–6.8	6.92×10^6	1294	1

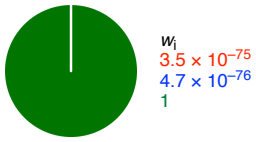
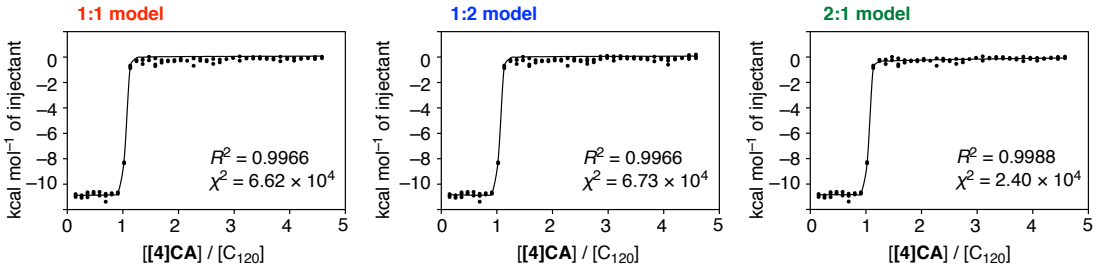


Figure 3.3. Fitting analysis of **[4]CC**+C_{120}



model	*N*	*k*	K_1 (M^{-1})	K_2 (M^{-1})	ΔH_1 (kcal mol^{-1})	ΔH_2 (kcal mol^{-1})	*SS*	AIC	w_i
1:1	117	2	3.47×10^7	–	–10.9	–	7.62×10^6	1301	4.8×10^{-26}
1:2	117	4	3.55×10^7	7.89×10^2	–10.9	+0.8	7.60×10^6	1305	7.5×10^{-27}
2:1	117	4	6.44×10^7	1.49×10^3	–10.8	–2.9	2.72×10^6	1184	1

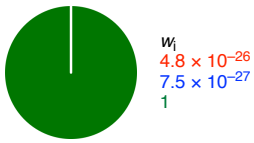
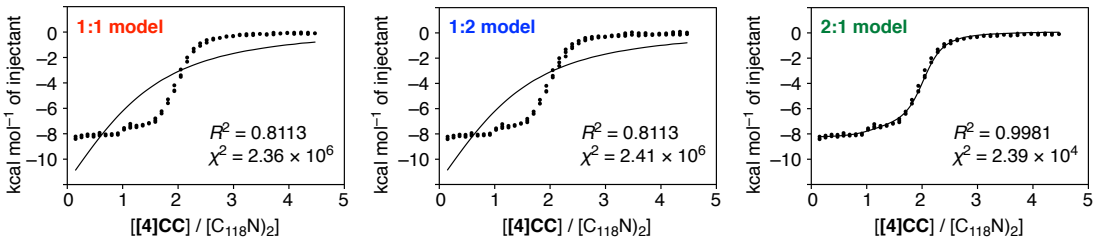


Figure 3.4. Fitting analysis of **[4]CA**+C_{120}



model	*N*	*k*	K_1 (M^{-1})	K_2 (M^{-1})	ΔH_1 (kcal mol^{-1})	ΔH_2 (kcal mol^{-1})	*SS*	AIC	w_i
1:1	114	2	1.17×10^4	–	–21.8	–	2.65×10^8	1675	4.9×10^{-114}
1:2	114	4	1.17×10^4	1.56×10^{-2}	–21.8	+6.0	2.65×10^8	1679	6.6×10^{-115}
2:1	114	4	7.01×10^6	3.20×10^5	–8.2	–7.6	2.63×10^6	1153	1

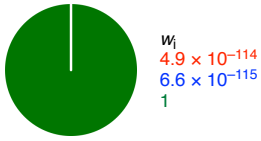


Figure 3.5. Fitting analysis of **[4]CC**+$C_{118}N_2$

Table 3.1 Modern analyses of ITC titration experiments of host-guest complexes

	[4]CC + C_{120} [a]	**[4]CA** + C_{120} [a]	**[4]CC** + $C_{118}N_2$ [a]
R^2	**1:1** = 0.8790	**1:1** = 0.9966	**1:1** = 0.8113
	1:2 = 0.8790	**1:2** = 0.9966	**1:2** = 0.8113
	2:1 = 0.9938	**2:1** = 0.9988	**2:1** = 0.9981
χ^2	**1:1** = 1.16×10^6	**1:1** = 6.62×10^4	**1:1** = 2.36×10^6
	1:2 = 1.19×10^6	**1:2** = 6.73×10^4	**1:2** = 2.41×10^6
	2:1 = 6.12×10^4	**2:1** = 2.40×10^4	**2:1** = 2.39×10^4
AIC	**1:1** = 1637	**1:1** = 1301	**1:1** = 1675
	1:2 = 1640	**1:2** = 1305	**1:2** = 1679
	2:1 = 1294	**2:1** = 1184	**2:1** = 1153
w_i	**1:1** = 3.5×10^{-75}	**1:1** = 4.8×10^{-26}	**1:1** = 4.9×10^{-114}
	1:2 = 4.7×10^{-76}	**1:2** = 7.5×10^{-27}	**1:2** = 6.6×10^{-115}
	2:1 = 1 [b]	**2:1** = 1 [b]	**2:1** = 1 [b]
K_n [c]	$K_1 = 1.16 \times 10^7\ M^{-1}$	$K_1 = 6.44 \times 10^7\ M^{-1}$	$K_1 = 7.01 \times 10^6\ M^{-1}$
	$K_2 = 1.30 \times 10^5\ M^{-1}$	$K_2 = 1.49 \times 10^3\ M^{-1}$	$K_2 = 3.20 \times 10^5\ M^{-1}$

[a] Titration experiments in *ortho*-dichlorobenzene (*o*DCB) were originally performed and reported in ref. 18 (**[4]CC**+C_{120}), 19 (**[4]CA**+C_{120}) and 20 (**[4]CC**+$C_{118}N_2$), and the raw titration data were reanalyzed in this study. [b] The best models had corresponding crystal structures (**[4]CC**+C_{120}: CCDC1452474; **[4]CA**+C_{120}: CCDC1573517; **[4]CC**+$C_{118}N_2$: CCDC1996110). [c] The K_n values for the best models (**2:1**) are shown.

The modern analyses revealed minor yet important features of the dumbbell-and-cylinder associations. The changes in association constants and thermodynamic parameters are summarized in Figure 3.6. The association constants were not severely affected and showed minor changes in log K values. However, the thermodynamic parameters behind these constants deviated from the classical analyses. Thus, although the enthalpy terms were essentially not affected, the entropy terms changed considerably. When the entropy terms are derived from slopes of fitting curves of ITC,[24] it is reasonable to observe such changes by the alternation of the fitting method. Importantly, modern analyses showed that the second-stage association of **[4]CA** with C_{120} was mainly driven by enthalpy. In a previous study, the entropy-driven nature of the association of

the second cylinder was a perplexing phenomenon (ΔH = –1.7 kcal mol^{-1}, $-T\Delta S$ = –3.1 kcal mol^{-1}; 298 K), but the present modern analyses revealed that the enthalpy was also a major contributor to the second association (ΔH = –2.9 kcal mol^{-1}, $-T\Delta S$ = –1.4 kcal mol^{-1}; 298 K).

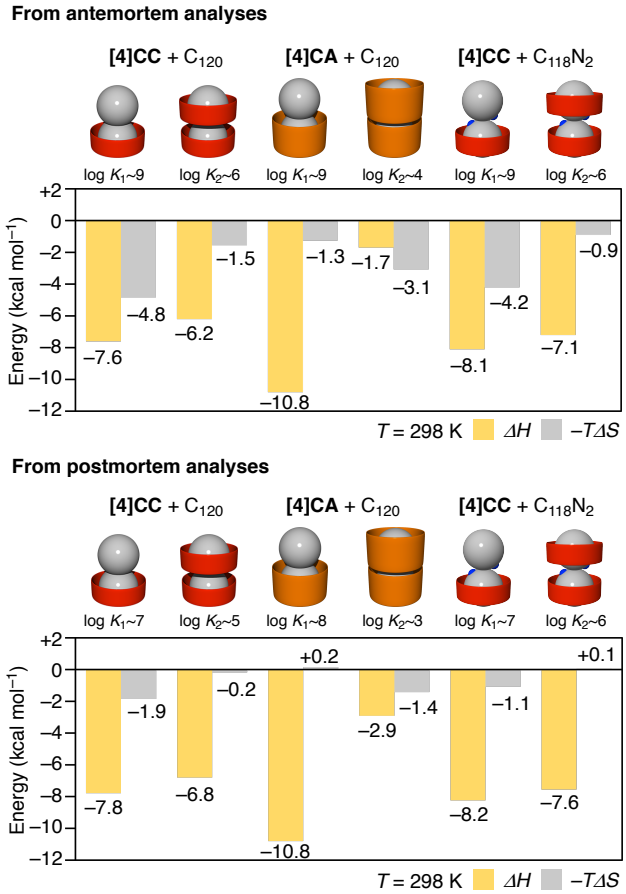


Figure 3.6. Association of the dumbbell-and-cylinder complexes before and after the modern analyses (*o*DCB). The classical data are taken from preceding papers.[18,19,20] The modern analyses by AIC showed that **2:1** is the best model for these systems. See Table 3.1.

3.4 [4]CC+X-COR

In Chapter 2, the stoichiometries of **[4]CC**+X-**COR** were determined by the *F*-test. Here, AIC analyses were also performed for supramolecular complexes of **[4]CC**+X-**COR**. As shown in Figure 3.7, the smallest AIC values were recorded for the **1:2** model in all the cases. The Akaike weight, the w_i values, was almost 1 for the **1:2** model. Therefore, these results suggest, with high confidence, that the **1:2** model is the best model among the three (**1:1**, **1:2** and **2:1**). In conclusion, the **1:2** model, which is consistent with the results of the *F*-test and the crystal structures, was also obtained as the best model by AIC analyses.

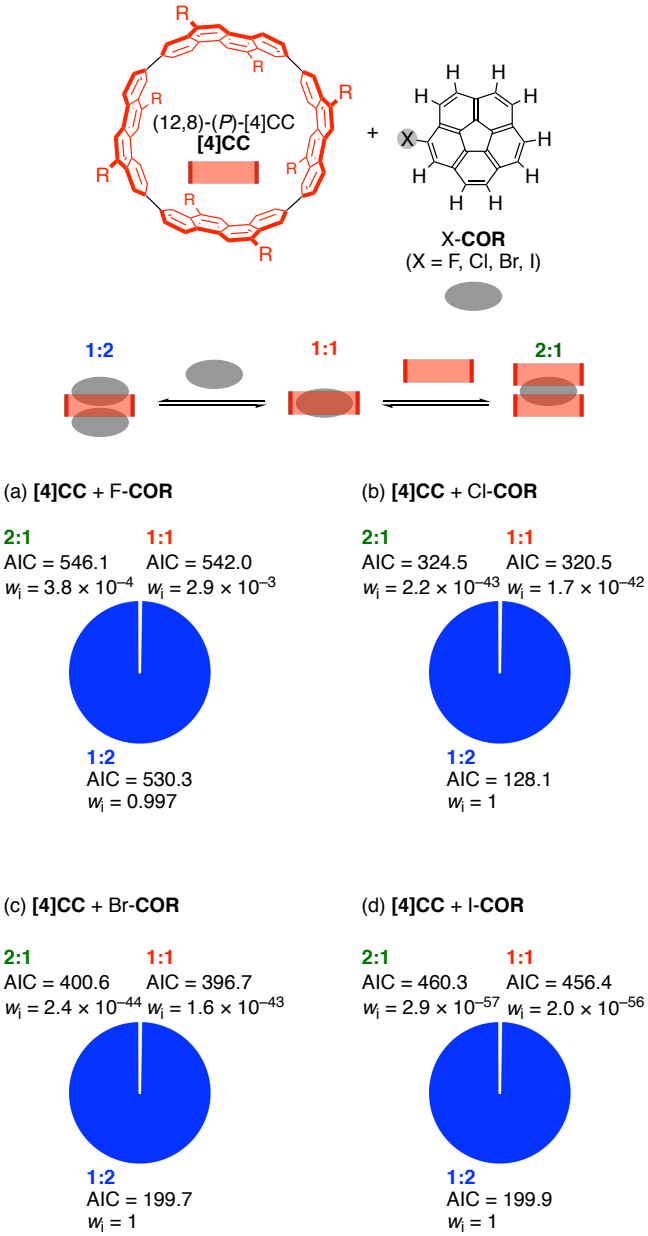


Figure 3.7. The results of the AIC analyses for **[4]CC**+X-**COR**. (a) **[4]CC**+F-**COR**. (b) **[4]CC**+Cl-**COR**. (c) **[4]CC**+Br-**COR**. (d) **[4]CC**+I-**COR**.

3.5 [4]CC+C_{70}

Next, an interesting case for the modern AIC analyses was examined. The system is composed of **[4]CC** and C_{70} as the host and guest, which has been treated as **1:1** in previous classical analyses.[25] As described above, ITC investigations with C_{120} and $C_{118}N_2$ revealed that the more complex **1:2** was the best model, and it was envisioned that C_{70} with a "shorter" oval shape should serve as an ideal test case to examine the versatilities of the AIC method to verify simpler fitting models. As shown in Figure 3.8, ITC titrations were performed in triplicate in *o*-dichlorobenzene (*o*DCB) at 298 K. Three isotherms were globally fitted with three models (**1:1**, **2:1** and **1:2**),[23] and the fitting qualities were first evaluated for GOF by R^2. The R^2 values were 0.9842 for the three models (Figure 3.9), which demonstrated sufficient GOF levels of the fitting.[26] Finally, the fitting models were compared by AIC. The AIC scores were 679.9, 684.0 and 683.9 for **1:1**, **1:2** and **2:1**, respectively, and the highest w_i value of 0.79 for **1:1** confirmed **1:1** as the best model with the highest weight of evidence.[27] The association constants and thermodynamic parameters were consequently derived from the fitting analysis to afford $K_1 = 1.96 \times 10^7$ M^{-1}, $\Delta H_1 = -9.89$ kcal mol^{-1} and $-T\Delta S_1 = -0.05$ kcal mol^{-1} (298 K).[28] The results show that the simpler models with fewer fitting parameters can be selected as the best model by AIC over the more complicated models with inherent superiority for fitting.[7]

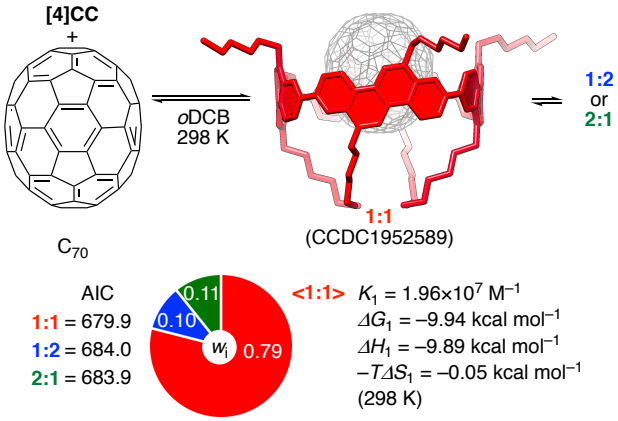
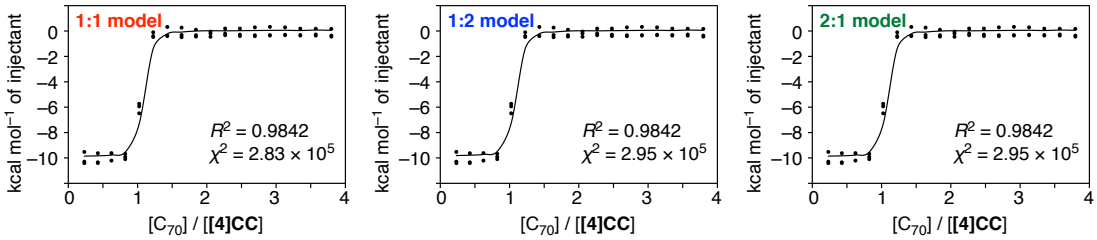


Figure 3.8. Modern AIC analyses of **[4]CC**+C_{70} systems.



model	*N*	*k*	K_1 (M^{-1})	K_2 (M^{-1})	ΔH_1 (kcal mol^{-1})	ΔH_2 (kcal mol^{-1})	*SS*	AIC	w_i
1:1	54	2	1.96×10^7	–	–9.89	–	1.47×10^7	679.9	0.79
1:2	54	4	1.92×10^7	3.98×10^{-1}	–9.84	–0.27	1.47×10^7	684.0	0.10
2:1	54	4	1.97×10^7	4.25×10^1	–9.89	–0.12	1.47×10^7	683.9	0.11

Figure 3.9. Fitting analysis of **[4]CC**+C_{70}

For comparison, an attempt to determine the stoichiometries were also tried to by the *F*-test, which failed. The results of the *F*-test are summarized in Figure 3.10. As written in the introduction, the **1:1**, **1:2** and **2:1** models cannot be compared by the *F*-test. Therefore, the comparison of the **1:1** and **1:2** models was first performed, and then the comparison of the **1:1** and **2:1** models was performed. In the case of the dumbbell-and-cylinder systems, **[4]CC**+C_{120}, **[4]CA**+C_{120} and **[4]CC**+$C_{118}N_2$, a **2:1** model was suggested in comparison to the **1:1** and **2:1** models. On the other hand, **1:2** was not necessarily suggested when **1:1** and **1:2** were compared, which indicated that no conclusion can be obtained by the *F*-test. Then, in the case of the complex **[4]CC**+C_{70}, both the **1:2** and **2:1** models were not necessarily suggested in the *F*-test. Therefore, the best model cannot be determined by the *F*-test, which again emphasizes the superiority of AIC analysis over the *F*-test.

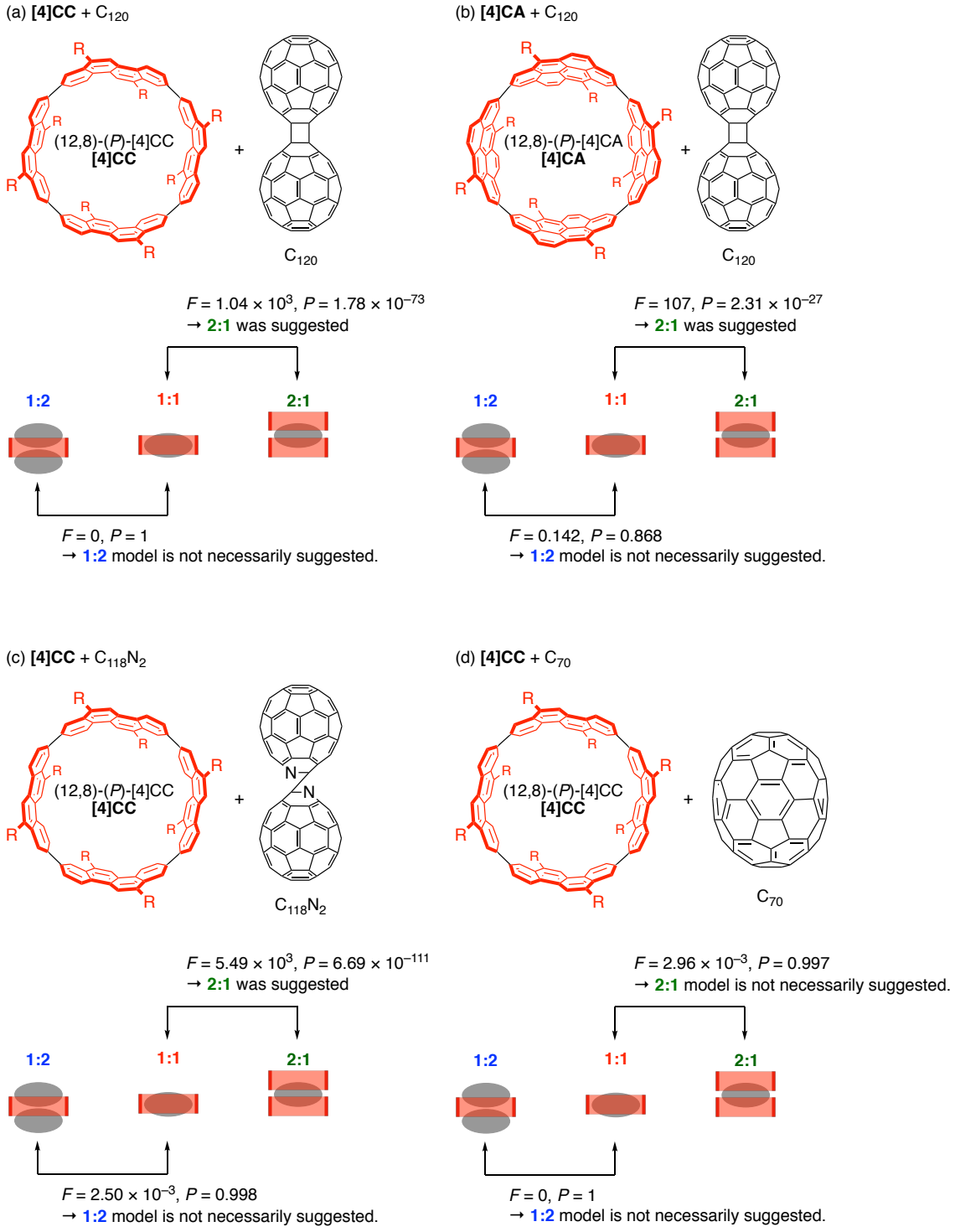


Figure 3.10. The result of the *F*-test. (a) **[4]CC**+C_{120}. (b) **[4]CA**+C_{120}. (c) **[4]CA**+$C_{118}N_2$. (d) **[4]CC**+C_{70}.

3.6 [4]CC + pyrene

第 6 節

本節については，５年以内に雑誌等で刊行予定のため非公開.

第 6 節

本節については，５年以内に雑誌等で刊行予定のため非公開.

3.7 Conclusion

In summary, the AIC and its weight (w_i) were introduced as a method for model selection after the death of the Job plot. Unlike preceding modern analyses with statistical null hypothesis testing (*F*-test/*P*-value),[5,17,29] these information-theoretic measures can be applied to models with nested and unnested relationships,[7] and **1:1**, **1:2** and **2:1** equilibrium models were compared for the dumbbell-and-cylinder complexes. By introducing the AIC measures, mysterious entropy contributions to the **[4]CA**+C_{120} system in previous classical analyses were resolved.[19] In addition, we performed the AIC analyses of **[4]CC**+X-**COR** systems to confirm its applicability.[17] Normally, in fitting analyses, the model with more parameters results in a better fit.[5,7] Therefore, fitting evaluations with GOF cannot offer any conclusion, which inevitably necessitates the introduction of a method for model selections. In this study, the AIC/w_i analyses were also applied to the **[4]CC**+C_{70} system and led to the conclusion of a simpler 1:1 model. Thus, the result demonstrates that AIC can serve as an Occam's razor for model selection in host-guest chemistry using the weight of evidence as a quantitative measure. Cylindrical hosts with cylindrical nanotube structures should readily accommodate multiple guests,[14] and AIC analyses should be particularly beneficial to study their host-guest chemistry in the future.[13] The information-theory approach with AIC has been applied to biological and medical sciences,[7,30] which is gradually expanding its reach to chemical science in fields such as analytical, pharmaceutical or environmental chemistry.[31] It is believed that chemists in the field of supramolecular chemistry can also find methods and analyses in the present study useful and instructive for investigations of their systems.

3.8 Experimental section

General

ITC experiments were performed by using a Malvern MicroCal iTC200 microcalorimeter.

Methods

Fitting

Procedures of fitting with equations are briefly described for 2:1 model.[3] The fitting equations of 1:1 and 1:2 model are shown in Chapter 2.

2:1. The binding constant K_1 and K_2 can be expressed by using concentration of a host ([H]), a guest ([G]), a 1:1 complex ([HG]) and a 2:1 complex ([H_2G]) as follows:

$$K_1 = \frac{[HG]}{[H][G]} \quad \text{(Eq. 3.3)}$$

$$K_2 = \frac{[H_2G]}{[HG][H]} \quad \text{(Eq. 3.4).}$$

In the 2:1 system, the physical property (Y) that is related to the absolute concentration of each species can be represented as

$$Y = Y_H[H] + Y_G[G] + Y_{HG}[HG] + Y_{H_2G}[H_2G] \quad \text{(Eq. 3.5),}$$

where Y_H, Y_G, Y_{HG} and Y_{H2G} are mole-dependent physical properties of pure H, pure G, pure HG and pure H_2G respectively. When pure G does not contribute to the physical property, one can simplify the equation as follows:

$$Y = Y_H[H] + Y_{HG}[HG] + Y_{H_2G}[H_2G] \quad \text{(Eq. 3.6).}$$

The change of the physical property (ΔY) is expressed as a subtraction of the initial value (Y_0) from Y and thus is derived as

$$\Delta Y = Y - Y_0 = \left(Y_H[H] + Y_{HG}[HG] + Y_{H_2G}[H_2G]\right) - Y_H\left([H] + [HG] + [H_2G]\right)$$

$$= \left(Y_{HG} - Y_H\right)[HG] + \left(Y_{H_2G} - Y_H\right)[H_2G] = Y_{\Delta HG}[HG] + Y_{\Delta H_2G}[H_2G] \quad \text{(Eq. 3.7).}$$

[HG] and [HG_2] are represented as the following equations:

$$\left[\mathrm{HG}\right]=\frac{\left[\mathrm{G}\right]_0 K_1\left[\mathrm{H}\right]}{1+K_1\left[\mathrm{H}\right]+K_1K_2\left[\mathrm{H}\right]^2} \qquad \text{(Eq. 3.8).}$$

$$\left[\mathrm{H_2G}\right]=K_2\left[\mathrm{H}\right]\left[\mathrm{HG}\right]=\frac{\left[\mathrm{G}\right]_0 K_1K_2\left[\mathrm{H}\right]^2}{1+K_1\left[\mathrm{H}\right]+K_1K_2\left[\mathrm{H}\right]^2} \qquad \text{(Eq. 3.9).}$$

Using (eq. 3.21), (eq. 3.22) and (eq. 3.23), one can obtain

$$\Delta Y=\frac{Y_{\Delta\mathrm{HG}}\left[\mathrm{G}\right]_0 K_1\left[\mathrm{H}\right]+Y_{\Delta\mathrm{H_2G}}\left[\mathrm{G}\right]_0 K_1K_2\left[\mathrm{H}\right]^2}{1+K_1\left[\mathrm{H}\right]+K_1K_2\left[\mathrm{H}\right]^2} \quad \text{(Eq. 3.10).}$$

In this equation, the concentration of H during a titration experiment is obtained by solving the following cubic equation,

$$K_1K_2\left[\mathrm{H}\right]^3+\left(2K_1K_2\left[\mathrm{G}\right]_0-K_1K_2\left[\mathrm{H}\right]_0+K_1\right)\left[\mathrm{H}\right]^2+\left(K_1\left[\mathrm{G}\right]_0-K_1\left[\mathrm{H}\right]_0+1\right)\left[\mathrm{H}\right]-\left[\mathrm{H}\right]_0=0$$

(Eq. 3.11).

In the calorimetry, quantity of heat (Q) becomes the corresponding physical property. Thus, the ITC titration data were fitted to an equation,

$$Q=\frac{\Delta H_1V\left[\mathrm{G}\right]_0 K_1\left[\mathrm{H}\right]+\left(\Delta H_1+\Delta H_2\right)V\left[\mathrm{G}\right]_0 K_1K_2\left[\mathrm{H}\right]^2}{1+K_1\left[\mathrm{H}\right]+K_1K_2\left[\mathrm{H}\right]^2} \quad \text{(Eq. 3.12),}$$

where V is the volume of solution, and ΔH_1 and ΔH_2 are the enthalpy term in the first and second association, respectively. In this non-linear fitting analysis, we used the Solver program in Microsoft Excel to obtain the fitting parameters, ΔH_1, ΔH_2, K_1 and K_2 with R^2 and χ^2 for evaluations of GOF.

Global fitting with [4]CC+C_{120}, [4]CA+C_{120} and [4]CC+$C_{118}N_2$

The ITC titration experiments of dumbbell-and-cylinder systems were performed in triplicate in the preceding works for **[4]CC**+C_{120},[18] **[4]CA**+C_{120}[19] and **[4]CC**+$C_{118}N_2$ in *o*-dichlorobenzene (*o*DCB).[20] In previous studies, each titration dataset was analyzed by using one model that was separately decided by the Job plot. In this study, the datasets were combined, and the whole data were subjected to global fitting of each dumbbell-and-cylinder system.[23] The combined data were fitted by three models (1:1, 1:2 and 2:1) in a respective manner, and the results were evaluated with R^2 and χ^2 for GOF. The fitting data for all the models were compared by AIC and w_i.

Titration experiments with [4]CC+C_{70}

The **[4]CC**+C_{70} system was previously investigated by combining two titration methods (fluorescence quenching and ITC) and the Job plot for the classical analyses.[25] In this study, the ITC titration experiments of were performed in triplicate for the modern analyses. A solution of C_{70} in *o*DCB (0.400 mM) was added to a solution of **[4]CC** in *o*DCB (0.0200 mM in a microcalorimeter cell) at 298 K by using the automated titration mode via a syringe. The whole data were fitted by three models (1:1, 1:2 and 2:1) in a respective manner, and the results were evaluated with R^2 and χ^2 for GOF. The fitting data for all the models were compared by AIC and w_i.

3.9 References

1. J.-M. Lehn, *Supramolecular Chemistry: Concepts and Perspectives*, VCH, Weinheim, 1995.
2. C. L. D. Gibb and B. C. Gibb in *Supramolecular Chemistry: From Molecules to Nanomaterials*, ed. P. A. Gale and J. W. Steed, Wiley, Chichester, 2012, vol. 1, pp 45-65.
3. (a) P. Thordarson, *Chem Soc. Rev.* **2011**, *40*, 1305-1323; (b) P. Thordarson in *Supramolecular Chemistry: From Molecules to Nanomaterials*, ed. P. A. Gale and J. W. Steed, Wiley, Chichester, 2012, vol. 2, pp 239-274.
4. (a) P. Job, *Ann. Chim.* **1928**, *9*, 113-203; (b) C. Y. Huang, *Meth. Enzymol.* **1982**, *87*, 509-525; (c) J. S. Renny, L. L. Tomasevich, E. H. Tallmadge and D. B. Collum, *Angew. Chem. Int. Ed.* **2013**, *52*, 11998-12013.
5. (a) F. Ulatowski, K. Dąbrowa, T. Bałakier, J. Jurczakjoc, *J. Org. Chem.* **2016**, *81*, 1746-1756; (b) D. B. Hibbert, P. Thordarson, *Chem. Commun.* **2016**, *52*, 12792-12805.
6. The equations in Figure 3.1 apply to mole-dependent ΔY. For properties that depend on mole fractions such as chemical shift differences ($\Delta\delta$), the coefficients (eg. $Y_{\Delta HG2}$) should have an additional multiplicity of 2. For example, see eqn (32) of ref. 3a. Note also that some equations in the references have minor errors and need careful examinations before use.
7. (a) K. P. Burnham, D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, Springer, New York, 2nd edn, 2002; (b) H. Motulsky, A. Christopoulos, *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*, Oxford Univ. Press, Oxford, 2004.
8. G. Currell, Scientific Data Analysis, Oxford Univ. Press, Oxford, 2015.
9. (a) D. H. Johnson, *J. Wildl. Manag*. **1999**, *63*, 763-772; (b) D. R. Anderson, K. P. Burnham, W. L. Thompson, *J. Wildl. Manag.* **2000**, *64*, 912-923.
10. (a) H. Akaike, *IEEE Trans. Automat. Contr.* **1974**, *AC19*, 716-723; (b) H. Akaike, *Psychometrika*, **1987**, *52*, 317-332.
11. Z. Sun, T. Matsuno, H. Isobe, *Bull. Chem. Soc. Jpn*. **2018**, *91*, 907-921.
12. K. Ikemoto, T. M. Fukunaga, H. Isobe, *Proc. Jpn. Acad. B*, **2022**, *98*, 379-400.

13. (a) B. W. Smith, M. Monthioux, D. E. Luzzi, *Nature*, **1998**, *396*, 323-324; (b) T. Okazaki, Y. Iizumi, S. Okubo, H. Kataura, Z. Liu, L. Suenaga, Y. Tahara, M. Yudasaka, S. Okada, S. Iijima, *Angew. Chem. Int. Ed.* **2011**, *50*, 4853-4857; (c) M. Koshino, T. Tanaka, N. Solin, K. Suenaga, H. Isobe, E. Nakamura, *Science*, **2007**, *316*, 853; (d) M. Koshino, Y. Niimi, E. Nakamura, H. Kataura, T. Okazaki, K. Suenaga, S. Iijima, *Nat. Chem*. **2010**, *2*, 117-124; (e) I. V. Krive, R. I. Shekhter, M. Jonson, *Low Temp. Phys.* **2006**, *32*, 1171-1194; (f) S. Iijima, *Physica B*, **2002**, *323*, 1-5.

14. T. Matsuno, S. Sato, H. Isobe in *Comprehensive Supramolecular Chemistry II*, ed. J. L. Atwood, G. W. Gockel, L. J. Barbour, Elsevier, Amsterdam 2017, vol. 3, pp 311-328.

15. J. W. Steed, J. L. Atwood and P. A. Gale in *Supramolecular Chemistry: From Molecules to Nanomaterials*, ed. P. A. Gale, J. W. Steed, Wiley, Chichester, 2012, vol. 1, pp 3-7.

16. T. Matsuno, S. Terasaki, K. Kogashi, R. Katsuno, H. Isobe, *Nat. Commun*. **2021**, *12*, 5062.

17. T. Matsuno, K. Takahashi, K. Ikemoto, H. Isobe, *Chem. Asian. J.* **2022**, *17*, e202200076.

18. T. Matsuno, S. Sato, A. Yokoyama, S. Kamata, H. Isobe, *Angew. Chem. Int. Ed.* **2016**, *55*, 15339-15343.

19. T. Matsuno, S. Kamata, S. Sato, A. Yokoyama, P. Sarkar, H. Isobe, *Angew. Chem. Int. Ed.* **2017**, *56*, 15020-15024.

20. T. Matsuno, M. Someya, S. Sato, S. Maeda, H. Isobe, *Angew. Chem. Int. Ed.* **2020**, *59*, 14570-14576.

21. (a) K. Tashiro, Y. Hirabayashi, T. Aida, K. Saigo, K. Fujiwara, K. Komatsu, S. Sakamoto, K. Yamaguchi, *J. Am. Chem. Soc*. **2002**, *124*, 12086-12087; (b) H. Sato, K. Tashiro, H. Shinmori, A. Osuka, Y. Murata, K. Komatsu, T. Aida, *J. Am. Chem. Soc.* **2005**, *127*, 13086-13087; (c) J. Rio, S. Beeck, G. Rotas, S. Ahles, D. Jacquemin, N. Tagmatarchis, C. Ewels, H. A. Wegner, *Angew. Chem. Int. Ed.* **2018**, *57*, 6930-6934; (d) A. Stergiou, J. Rio, J. H. Griwatz, D. Arčon, H. A. Wegner, C. P. Ewels, N. Tagmatarchis, *Angew. Chem. Int. Ed.* **2019**, *58*, 17745-17750; (e) Y.-Y. Xu, H.- R. Tian, S.-H. Li, Z.-C. Chen, Y.-R. Yao, S.-S. Wang, X. Zhang, Z.-Z. Zhu, S.-L. Deng, Q.

Zhang, S. Yang, S.-Y. Xie, R.-B. Huang and L.-S. Zheng, *Nat. Commun.*, **2019**, *10*, 485; (f) C. García- Simón, C. Colomban, Y. A. Çetin, A. Gimeno, M. Pujals, E. Ubasart, C. Fuertes-Espinosa, K. Asad, N. Chronakis, M. Costas, J. Jiménez-Barbero, F. Feixas, X. Ribas, *J. Am. Chem. Soc.* **2020**, *142*, 16051-16063; (g) Y. Tanuma, A. Stergiou, A. B. Bobnar, M. Gaboardi, J. Rio, J. Volkmann, H. A. Wegner, N. Tagmatarchis, C. P. Ewels, D. Arčon, *Nanoscale*, **2021**, *13*, 19946-19955; (h) W. Wang, F. Hanindita, Y. Hamamoto, Y. Li, S. Ito, *Nat. Commun.* **2022**, *13*, 1498.

22. The global fitting analyses of two-stage associations require two events to be detected during a single sequence of titration. For the two associations of the present system to be detected, the fluorescence quenching titration is not suitable, because the first quenching constant (k_{HG}) and the second quenching constant (k_{HG2}) are zero. See also ref. 3.
23. J. M. Beechem, *Meth. Enzymol.* **1992**, *210*, 37-54.
24. F. P. Schmidtchen in *Supramolecular Chemistry: From Molecules to Nanomaterials*, ed. P. A. Gale, J. W. Steed, Wiley, Chichester, 2012, vol. 2, pp 275-296.
25. (a) S. Hitosugi, R. Iizuka, T. Yamasaki, R. Zhang, Y. Murata, H. Isobe, *Org. Lett.* **2013**, *15*, 3199-3201; (b) T. Matsuno, S. Sato, R. Iizuka, H. Isobe, *Chem. Sci.* **2015**, *6*, 909-916; (c) T. Matsuno, Y. Nakai, Y. Maniwa, M. Someya, S. Sato, H. Isobe, *Chem. Asian J.* **2020**, *15*, 273- 278.
26. The fitting quality does not necessarily correlate with the model significance (ref. 5,7,8).
27. A 1:1 complex was observed in a single crystal of the **[4]CC**+C_{70} system (CCDC1952589). See also ref. 25.
28. The previous antemortem analyses afforded comparable values (ref. 25): $\log K_1 = 9.6$, $\Delta H_1 = -10.1$ kcal mol^{-1} and $-T\Delta S_1 = +0.298$ kcal mol^{-1}.
29. A care should be paid in evaluations of *P*-values from the statistical null hypothesis testing. See ref. 7 and 9.
30. Although AIC is not popular in the field of chemistry yet, this measure is versatile and has been applied to various types of model selections. The original paper (ref. 10a) has been cited >33000 times (Web of Science, as of Dec. 2022).
31. (a) J. Xu, Y. Wang, G. Wang, C. Huang, H. Hao, Q. Yin, *J. Chem. Thermodyn.* **2016**,

92, 12-20; (b) H. Taira, H. Nagase, T. Endo, H. Ueda, *J. Incl. Phenom. Macrocycl. Chem.* **2006**, *56*, 23-28; (c) M. Hanselmann, M. Kirchner, B. Y. Renard, E. R. Amstalden, K. Glunde, R. M. A. Heeren, F. A. Hamprecht, *Anal. Chem.* **2008**, *80*, 9649-9658; (d) E. E. Finney, R. G. Finke, *Chem. Mater.* **2009**, *21*, 4468-4479.

Chapter 4. Summary

In this thesis, supramolecular complexes of cylindrical hosts with multiple guests are surveyed by statistical analyses as a unique showcase for highlighting the importance of the stoichiometry determination of the complexes after the death of the Job plot. As methods relying on statistical analyses, the *F*-test and an information-theory approach of Akaike's information criterion (AIC) are used for the determination of the stoichiometry of supramolecular complexes.

In **Chapter 1**, the general introduction about the stoichiometry determination of the supramolecular complexes was discussed. The current situation is summarized in that the Job plot analysis, one of the conventional methods, is now outdated owing to its applicability. This chapter also described the emerging field of supramolecular complexes of cylindrical molecules, which allows for the encapsulation of multiple guests.

In **Chapter 2**, the associations between **[4]CC** and a series of halogenated corannulenes (X-**COR**) were examined (Figure 4.1). The formation of the supramolecular complexes was confirmed by ^{1}H NMR spectroscopy in the solution phase. First, Job plot analysis using the chemical shift change of ^{1}H NMR spectra was conducted. According to the Job plot, a 1:1 association for F-**COR** and a 1:2 association for Cl-**COR**, Br-**COR** and I-**COR** were indicated. The discrepancy in the difference in stoichiometries was questioned, so the *F*-test was used to test the ITC data as an alternative for the Job plot. As a result, the stoichiometry of the complex was determined to be 1:2 in all cases, which was unequivocally determined by X-ray crystallographic analysis. Interestingly, it was found that the cooperativity of the associations, the relation between the first and second association, was changed depending on the size of the halogen substituents of the guests. Negative cooperativity was revealed for F-**COR**, which indicated that the second association was disturbed by the first association. On the other hand, in the case of other guests, such as Cl-**COR**, Br-**COR** and I-**COR**, cooperativity was positive, which showed that the second association was facilitated by the first association. This unique difference in cooperativity was attributed to the structural differences of the complexes. Thus, when the halogen atom is small enough for the **[4]CC** host such as F-**COR** to wrap around the rim, hydrogen atoms can make good contact with the cylinder wall for CH-π interactions.

However, when the halogen atom becomes larger, the first X-**COR** guest cannot be fully wrapped by the **[4]CC** host and is tilted to lose some of the CH-π hydrogen bonds. Then a space in the cylinder remains to accommodate the second X-**COR** guest via CH-π contacts. Therefore, the positive cooperativity was activated by the size mismatch between **[4]CC** and X-**COR**, which could be accompanied by the inclination of the guest in the cylinder.

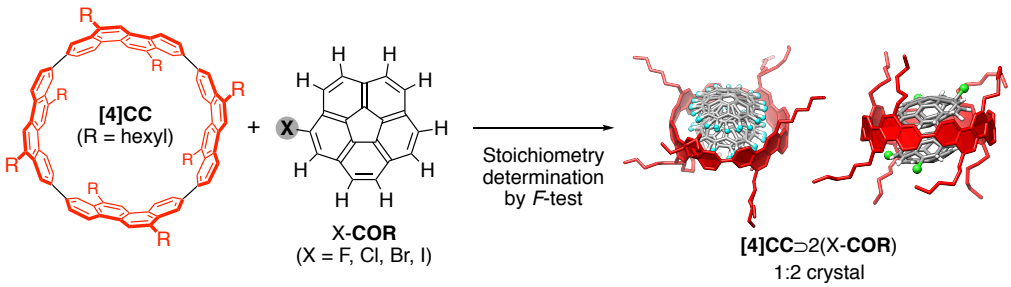


Figure 4.1. The association between **[4]CC** and X-**COR**s described in Chapter 2.

In **Chapter 3**, much more complicated supramolecular systems of cylindrical hosts are investigated. As a few testcases, supramolecular complexes of cylindrical hosts and C_{120}, which were revealed to show a 2:1 complexation by crystallographic analysis in the preceding study, were examined using the statistical analysis method to determine the stoichiometry (Figure 4.2). It was found that the *F*-test faced a limited applicability in these complicated supramolecular systems that encompass, for example, 1:2 and 2:1 models. Thus, AIC is used in the stoichiometry determination of the supramolecular complexes in this study. Among the possible three models (1:1, 1:2 and 2:1 models), AIC analysis concluded that the 2:1 model was the best model for **[4]CC**+C_{120}. Two other systems, **[4]CA**+C_{120} and **[4]CC**+$C_{118}N_2$, were also analyzed by AIC. As a result, both dumbbell-and-cylinder complexes confirmed that the 2:1 model was the best model for these systems. As an interesting test case for the 1:1 complexation, the complex of **[4]CC** and C_{70} was examined by AIC, which also resulted in the conclusion that the 1:1 model was the best model. This result showed that the simpler models with fewer fitting parameters can be selected as the best model by AIC over the more complicated models with inherent superiority for fitting.

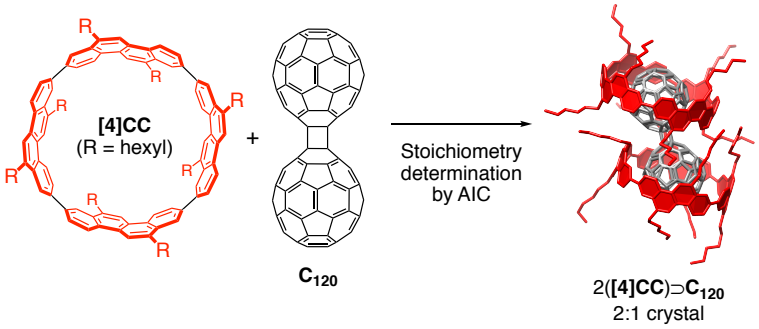


Figure 4.2. Application of AIC to the dumbbell-and-cylinder systems.

Overall, this work demonstrates the usefulness of AIC as a method to determine the stoichiometry of host-guest complexes. AIC analysis has been widely applied to biological and medicinal chemistry, but to the best of my knowledge, it has not been applied to host-guest chemistry. Cylindrical hosts can in principle accommodate multiple guests, and AIC analyses should be beneficial for studying host-guest chemistry. Using precise stoichiometry determination, an in-depth understanding of the inner space of cylindrical hosts could be elucidated, as exemplified by the unique cooperativity found in this work. The impact of this work not only focused on cylindrical molecules but also covers a variety of fields of supramolecular chemistry. In this sense, this work introducing AIC should be a versatile remedy after the death of the Job plot.

List of Publications

" Activation of positive cooperativity by size-mismatch assembly via inclination of guests in a single-site receptor"
T. Matsuno, K. Takahashi, K. Ikemoto, H. Isobe, *Chem. Asian. J.* **2022**, *17*, e202200076.
(Chapter 2)

"Introduction of Akaike's information criterion after the death of the Job plot"
K. Ikemoto, K. Takahashi, T. Ozawa, H. Isobe, *submitted*
(Chapter 3)

Other publications

Not applicable

Acknowledgments

I would like to express my deep gratitude to Professor Hiroyuki Isobe for providing the environment and facilities and guiding me through five years of my academic career full of difficulties. Without any of these, my studies, including this thesis, would not have been possible.

I must express my sincere gratitude to Dr Taisuke Matsuno, not only for his scientific support in teaching me experimental skills, knowledge and methods but also for his guidance in my own life and moral support. In particular, through discussions and detailed advice, my study was able to proceed smoothly. I was able to develop scientific skills, experiments and presentation and grow spiritually thanks to these teachings and support.

I would like to thank Dr Koki Ikemoto for supporting me in the final year, although we were originally in different groups. Thanks to his support and advice I was able to continue my studies. Without his help, I would not have been able to complete my Ph.D. thesis.

I would also like to thank all the current and former members of the Isobe laboratory. They have supported me in direct and indirect ways so that I could enjoy my research life in the lab. I was able to continue my research for five years because the time I spent with them was comfortable.

I thank SPring-8 and KEK for providing access to their synchrotron radiation facilities, without which the crystal structures in my study would not be available.

I thank the Japan Chemical Industry Association for the financial support that has made my research life fulfilling.

Finally, I would like to express my special gratitude to my family for their financial and psychological support. Without their support, I would not be able to afford my studies. Thank you very much.

Tokyo, Japan

December 19th, 2022

Kanato Takahashi

Department of Chemistry

The University of Tokyo