# SUPPORTING INFORMATION

## Self-Duplicating Amplification in a Dynamic Combinatorial Library

Shengguang Xu and Nicolas Giuseppone*

## Contents

***Discussion of individual reactions and of the origin of the kinetic and thermodynamic observations.***

**Table A.** Values of initial rates ($V_0$) and equilibrium constants (K) for the condensation reactions leading to imines $Al_{(1-3)}Am_{(1,2)}$. The data were determined by $^1H$ NMR at 22°C and in various deuterated solvent compositions.

| Entry | Imine[a] | CDCl₃ / DMSO$_{d6}$ volume ratio[a] | Initial rate ($V_0$ / mM.h$^{-1}$)[c] | Equilibrium constant (K)[d] |
|---|---|---|---|---|
| 1 | $Al_1Am_1$ | 1:0 | 5.9 | 300 |
| 2 | $Al_1Am_1$ | 1:0[b] | $5.4\ 10^{-2}$ | n.d. |
| 3 | $Al_1Am_1$ | 1:0.1 | $6.5\ 10^{-1}$ | n.d. |
| 4 | $Al_1Am_1$ | 1:0.4 | $4.2\ 10^{-1}$ | n.d. |
| 5 | $Al_1Am_1$ | 0:1 | $2.0\ 10^{-1}$ | n.d. |
| 6 | $Al_1Am_2$ | 1:0 | 1.8 | 50 |
| 7 | $Al_2Am_1$ | 1:0 | $4.8\ 10^{-1}$ | 17 |
| 8 | $Al_2Am_2$ | 1:0 | $3.5\ 10^{-1}$ | 10 |
| 9 | $Al_3Am_1$ | 1:0 | $5.3\ 10^{-1}$ | 20 |
| 10 | $Al_3Am_2$ | 1:0 | $3.8\ 10^{-1}$ | 12 |

[a] All reactions (except entry 2) were performed at a concentration of 15 mM for each starting material $Al_{(1-3)}$ and $Am_{(1,2)}$; [b] concentration of 2 mM for each starting materials $Al_1$ and $Am_1$; [c] Determined by plotting the product concentration *vs.* time graphics and by using the "tangent" tool of the software "origin 7.5" at t = 0; [d] Defined as the concentration ratio $[Al_{(1-3)}Am_{(1,2)}]^2$ / $[Al_{(1-3)}] [Am_{(1,2)}]$, assuming that $[Al_{(1-3)}Am_{(1,2)}] = [H_2O]$ which is not measurable by NMR.

By looking at the thermodynamics of the individual reactions in a solution of $CDCl_3$ at 22°C (Table A), it appears that $Al_1Am_1$, as expected, presents a much higher stability constant (K = 300, entry 1) compared to all the other imines. However, a slight difference exists between, on one hand, $Al_2Am_1$, $Al_2Am_2$, $Al_3Am_1$, and $Al_3Aml_2$ ($10 \leq K \leq 20$, entries 7-10) and, on the other hand, $Al_1Am_2$ (K = 50, entry 6). This classification of the six compounds in three sets can also be made by looking at the kinetic parameters in the same conditions (initial concentrations of 15 mM for each starting material $Al_{(1-3)}$ and $Am_{(1,2)}$). The formation of $Al_1Am_1$ ($V_0 = 59\ 10^{-1}$ mM.h$^{-1}$, entry 1) is more than three times faster than the formation of $Al_1Am_2$ ($18\ 10^{-1}$ mM.h$^{-1}$, entry 6), and about fourteen times faster than the formation of the 4 other imines ($3.5\ 10^{-1}$ mM.h$^{-1} \leq V_0 \leq 5.3\ 10^{-1}$ mM.h$^{-1}$, entries 7-10). We assume that the middle values of K and $V_0$ for compound $Al_1Am_2$ are the consequence of the formation of a dimer $[Al_1Am_2]_2$ as the presence of a remaining N-H bond on the adenine group, but with a much lower stability than $[Al_1Am_1]_2$ due to a higher steric hindrance of the benzyl group. The presence of broad resonance signals and that of a chemical shift of 10.5 ppm for the imide proton in $^1H$ NMR are in correlation with such a weak and slow exchange process between $Al_1Am_2$ and $[Al_1Am_2]_2$ (see NMR data in the supporting information material below). Moreover, in the reaction producing $Al_1Am_1$, the presence of increasing volume ratios of $DMSO_{d6}$ as co-solvent – that is capable of disrupting hydrogen bonds – leads to a decrease of the corresponding initial rates to a value close to the ones observed for the set of imines $Al_{(2,3)}Am_{(1,2)}$ (entries 1,3-5).

Although these comparative analyses with protected and unprotected recognition groups, in various solvents, and together with the characteristic chemical shifts in $^1H$ NMR, indicate that the higher equilibrium and rate constants for the formation of $Al_1Am_1$ are supposed to be related to the presence of homodimer $[Al_1Am_1]_2$, these amplifications can be the consequences of several mechanistic channels.[1] This is why we investigated if a formal termolecular autocatalytic pathway (i.e. acceleration of the reaction in the presence of the product and sigmoid concentration/time profile) is detectable in this system. This question was also important because of the controversial discussions between Rebek, Menger and Reinhoudt about the initial amide-type self-replicator designed by Rebek in which the amide bond itself can catalyze the reaction.[2] In our case, we do not obtain sigmoid time-dependent profile in pure $CDCl_3$, but the system displays an extremely slight exponential growth in the first 10% of the reaction by decreasing the product inhibition effect in a $CDCl_3:DMSO_{d6}$ (1:0.4) solution (Figure B). This observation was correlated by the fact that the product $Al_1Am_1$ catalyzes the condensation of $Al_1$ and $Am_1$, as is shown in Figure B, which describes the time course of the reaction upon the addition of initial amounts of product. Indeed, for the

following initial concentrations of $Al_1Am_1$ : 0 M, 1.20 mM, 3.45 mM, and 7.50 mM; we measured the corresponding initial rates : 4.15 $10^{-1}$ mM.h$^{-1}$, 7.59 $10^{-1}$ mM.h$^{-1}$, 9.30 $10^{-1}$ mM.h$^{-1}$, and 10.7 $10^{-1}$ mM.h$^{-1}$, respectively. These results show that the adenosine/Kemp's imide based replicator is able to weakly catalyze its own formation, although it importantly does not contain an amide bond in our case but an isosteric imine one.[2 b,c] However, the much higher initial rate of formation of $Al_1Am_1$ in pure chloroform (5.9 mM.h$^{-1}$) compared to non-dimeric imines $Al_{(2,3)}Am_{(1,2)}$ (3.5 $10^{-1}$ to 5.3 $10^{-1}$ mM.h$^{-1}$) is assumed to require the presence of different concomitant catalytic termolecular and bimolecular channels (in particular including the intramolecular $Al_1/Am_1$ condensation that is likely to be the kinetically determining channel), which is in agreement with the general conclusions of Reinhoudt's studies on Rebek's previous self-replicator.[2 d]
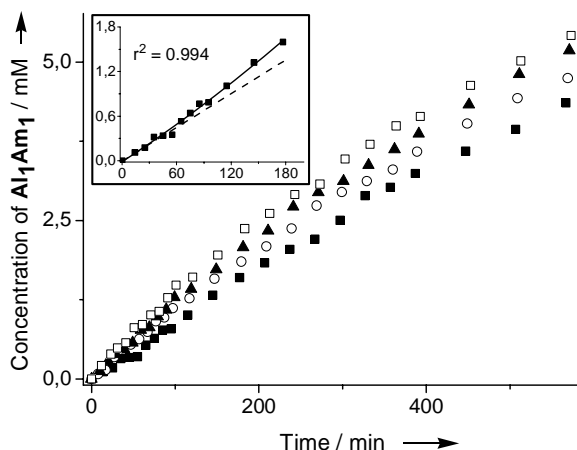


**Figure B.** Evolution with time of the formation of compound $Al_1Am_1$ (beginning of the reaction in a CDCl$_3$:DMSO$_{d6}$ / 1:0.4 volume ratio), as a function of the presence of various initial amounts of product $Al_1Am_1$. $[Al_1]_i = [Am_1]_i = 15$ mM; (■) : No initial presence of product $Al_1Am_1$; (○) : initial addition of 8 mol% of product $Al_1Am_1$; (▲) : initial addition of 23 mol% of product $Al_1Am_1$; and (□) : initial addition of 50 mol% of product $Al_1Am_1$. In the expanded window (*top left*), the kinetic evolution shows an exponential growth at the very beginning of the reaction when no product is initially added. All the reactions were performed at the same time and from the same initial solution of $Al_1$ and $Am_1$. Measured by $^1$HNMR at 22°C (uncertainty: ± 0.1 mM).

References for the above discussion:

[1] (a) R. J. Pearson, E. Kassianidis, A. M. Z. Slawin, D. Philp, *Org. Biomol. Chem.* **2004**, *2*, 3434; (b) R. J. Pearson, E. Kassianidis, A. M. Z. Slawin, D. Philp, *Chem. Eur. J.* **2006**, *12*, 6829.
[2] (a) V. Rotello, J.-I. Hong, J. Rebek, *J. Am. Chem. Soc.* **1991**, *113*, 9422; (b) F. M. Menger, A. V. Eliseev, N. A. Khanjin, *J. Am. Chem. Soc.* **1994**, *116*, 3613; (c) F. M. Menger, A. V. Eliseev, N. A. Khanjin, M. J. Sherrod, *J. Org. Chem.* **1995**, *60*, 2870; (d) D. N. Reinhoudt, D. M. Rudkevich, F. de Jong, *J. Am. Chem. Soc.* **1996**, *118*, 6880.

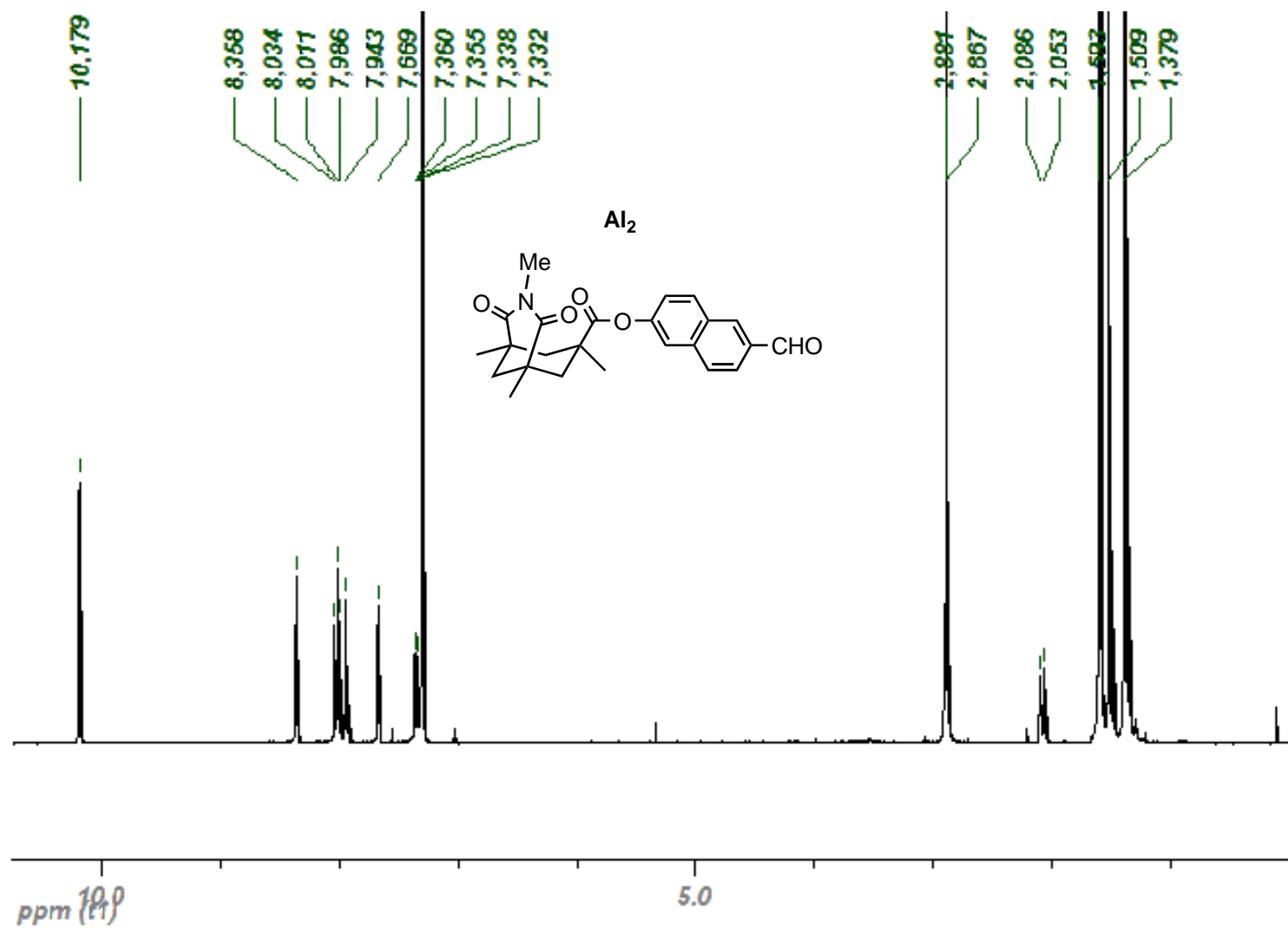**_Discussion of the terms self-duplication and self-replication._**

By examining the literature, there appear some discrepancies in the use of the term self-replication.

1/ Here we understand the term "self-replication" as discussed by G. von Kiedrowski (V. Patzke, G. von Kiedrowski, _ARKIVOC_ **2007**, 29), i.e. related to auto-catalytic systems – in which the catalysis is the result of an association between the final product and its two components in a termolecular complex – displaying a sigmoid concentration-time profile and following a square root rate law.

2/ Here we understand the term "self-duplication" as a general property for a system to thermodynamically or kinetically (or both) favour its own formation.

S5

$^1$H NMR of **AI$_2$**

# S7

10,187

8,377
8,073
8,051
8,020
7,999
7,944
7,922
7,680
7,675
7,401
7,395
7,379
7,373
7,292

2,414

1,592

**AI$_3$** AcO⟶naphthalene⟶CHO

ppm (t1)   10.0

5.0

# S8

$^1$H NMR of **Am$_2$**

8,433
7,864
7,420
7,402
7,389
7,372
7,353
7,332
7,315
7,292
6,052
6,044
5,523
5,516
5,507
5,500
5,077
5,068
5,061
5,052
4,898
4,293
4,280
3,060
3,049
3,019

**Am$_2$**

ppm (t1)   9.0   8.0   7.0   6.0   5.0   4.0   3.0   2.0   1.0   0.0

# S10

Directly from the condensation of **Al₁** and **Am₁** (no isolation)

Peak labels: 11,177 · 10,050 · 8,048 · 7,636 · 7,384 · 7,194 · 5,818 · 5,812 · 5,804 · 5,798 · 5,744 · 5,738 · 5,255 · 5,242 · 4,904 · 3,989 · 3,653 · 3,648 · 3,620 · 3,614 · 2,907 · 2,871 · 2,783 · 2,748 · 2,102 · 2,037 · 2,001

**Al₁Am₁**

ppm (f1)     10.0          5.0          0.0

# S11

Directly from the condensation of **Al₁** and **Am₂** (no isolation)

Peak labels: 10,033; 8,095; 7,741; 7,683; 7,544; 7,621; 7,234; 7,214; 7,195; 7,172; 7,099; 7,094; 7,027; 7,071; 6,134; 5,515; 5,510; 5,507; 5,499; 5,486; 5,117; 4,656; 4,650; 3,791; 3,77; 3,764; 3,771; 3,767; 2,812; 2,796; 2,780; 2,763; 2,747

**Al₁Am₂**

ppm (f1)   10.0                    5.0                        0.0

# S12

Directly from the condensation of **Al$_2$** and **Am$_1$** (no isolation)

$^1$H NMR of **Al$_2$Am$_1$**



Al$_2$Am$_1$

10,055 8,214 8,163 7,805 7,735 7,717 7,619 7,616 7,683 7,439 7,434 7,172 7,139 7,134 7,117 7,112 6,019 6,012 5,492 5,229 5,113 4,677 3,825 3,816 1,925

ppm (f1) 10.0    5.0    0.0

# S13

Directly from the condensation of **Al₂** and **Am₂** (no isolation)



¹H NMR of **Al₂Am₂**

Al₂Am₂

Peaks: 10.080, 8.326, 8.202, 7.801, 7.732, 7.664, 7.482, 6.073, 6.057, 5.950, 5.498, 5.492, 5.483, 5.477, 5.136, 5.130, 5.121, 5.115, 4.677, 4.671, 4.665, 3.865, 3.854, 1.981, 1.948

ppm (f1) 10.0          5.0          0.0

# S14

Directly from the condensation of **Al₃** and **Am₁** (no isolation)

$^1$H NMR of **Al₃Am₁**

10,066

8,218
7,792
7,559
6,000
5,994
5,551
5,545
5,536
5,530
5,198
5,138
5,133
5,123
5,118
4,709
4,704
3,812
3,803
2,070

**Al₃Am₁**

ppm (f1) 10.0          5.0          0.0

# S15

Directly from the condensation of **AI₃** and **Am₂** (no isolation)

$^1$H NMR of **AI₃Am₂**



AI₃Am₂

Peak labels: 10,074 | 8,187 | 7,783 | 7,750 | 7,630 | 7,468 | 7,288 | 7,284 | 7,226 | 7,215 | 7,181 | 7,156 | 6,055 | 6,049 | 5,509 | 5,494 | 5,132 | 5,126 | 5,117 | 4,675 | 4,670 | 4,664 | 4,654 | 3,855 | 3,844 | 2,079

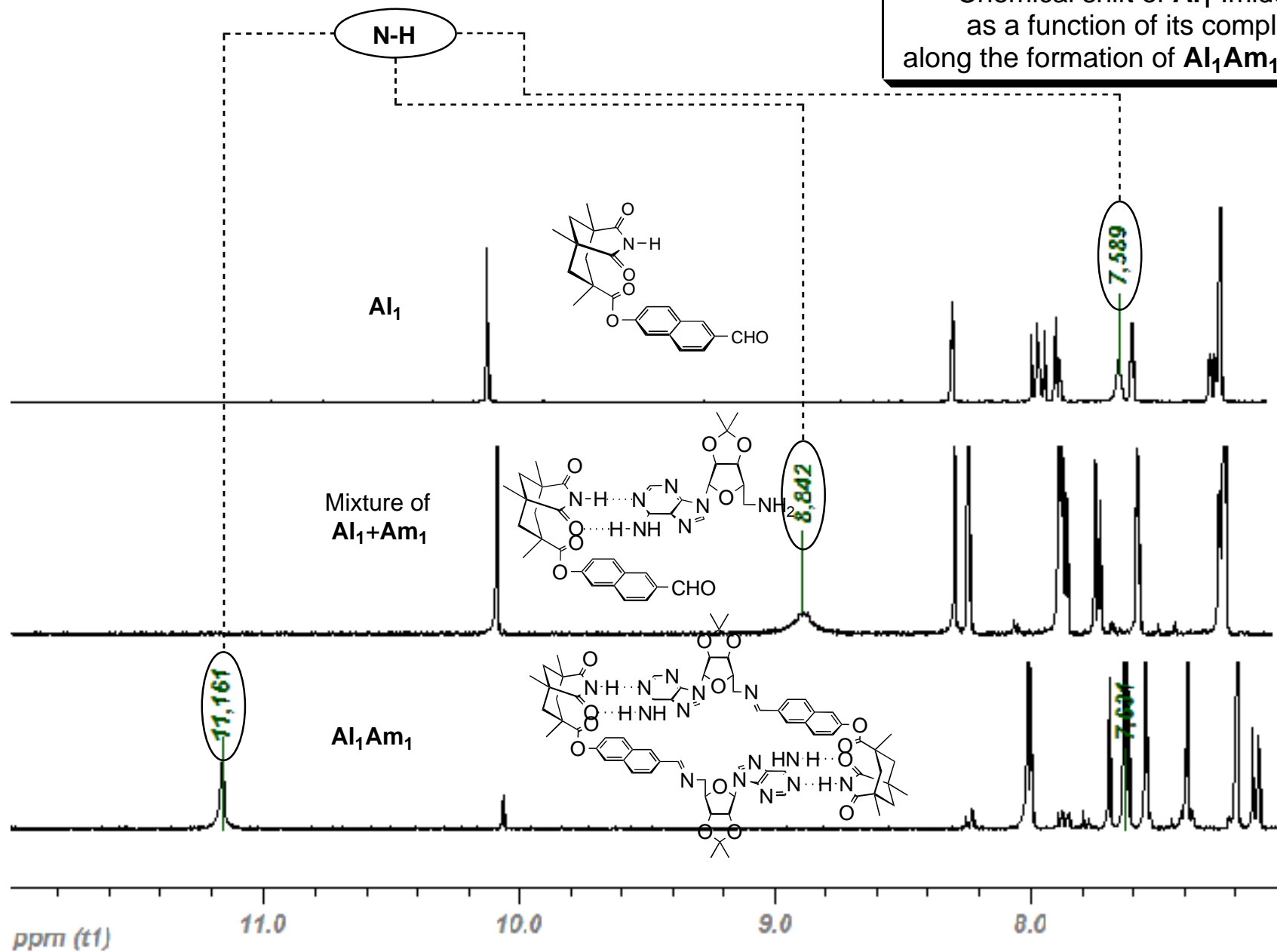ppm (f1)  10.0          5.0          0.0

S16



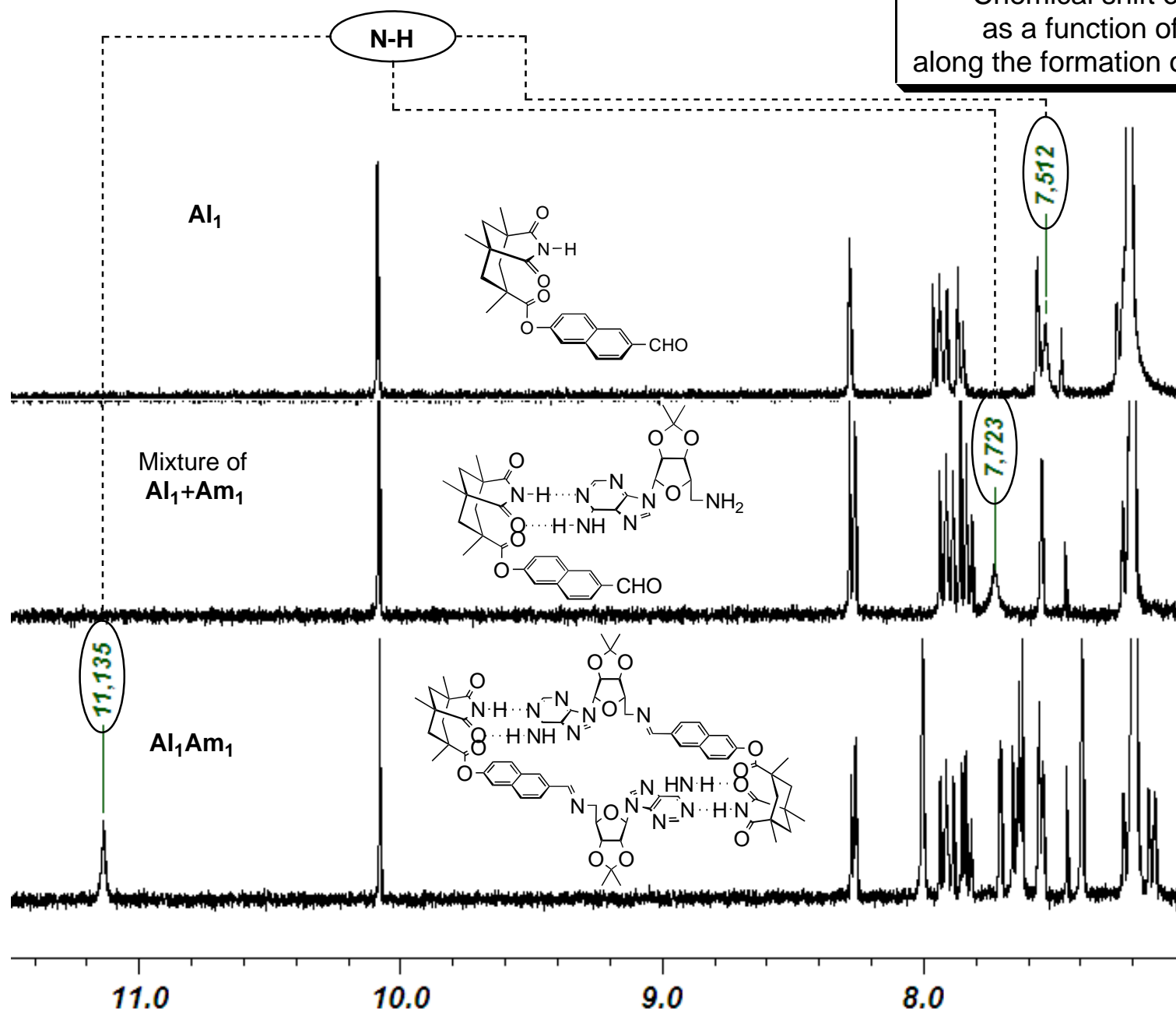Mixture of 11 compounds for **DCL2** at t = 16 h
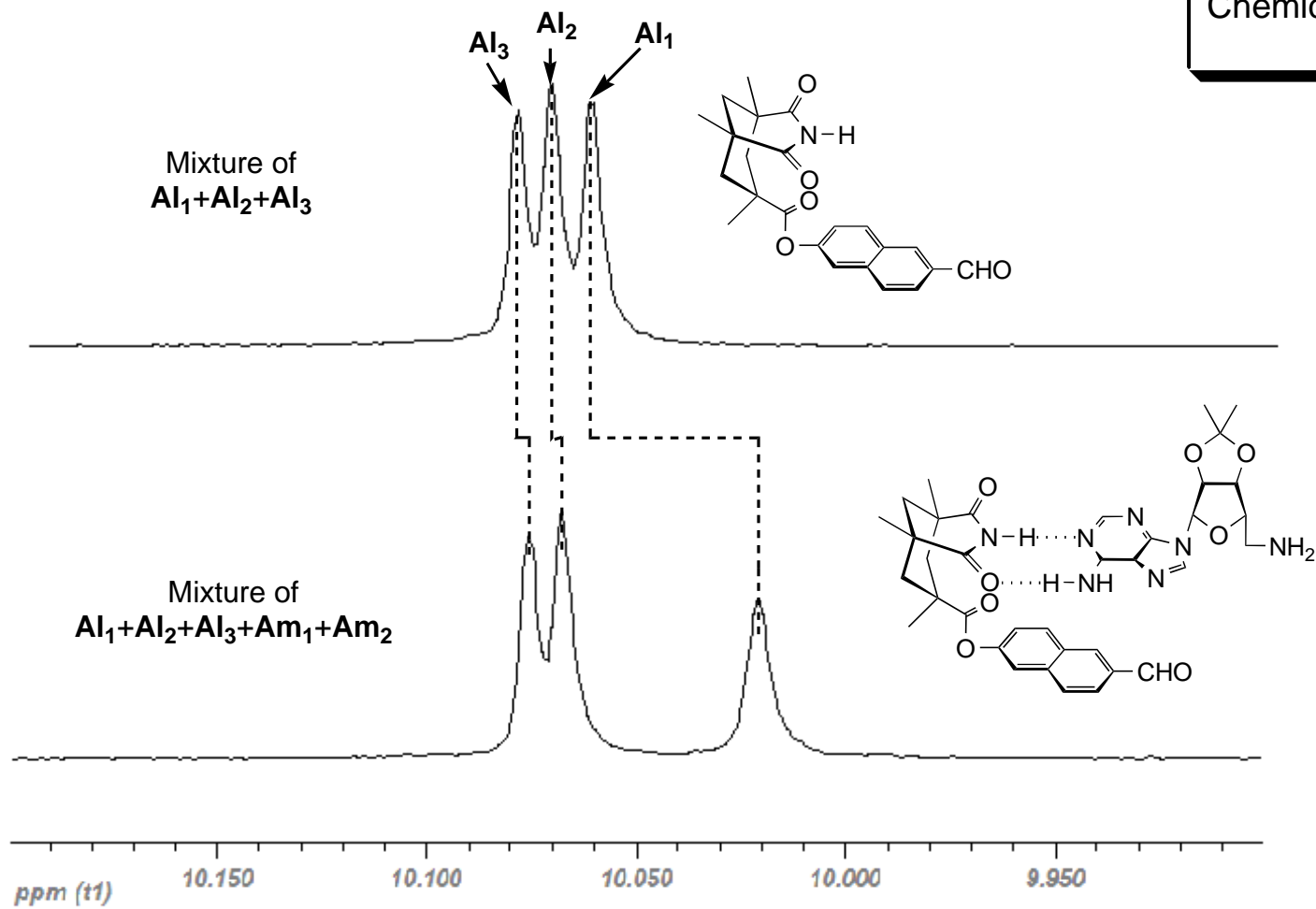
ppm (t1)   10.0   5.0

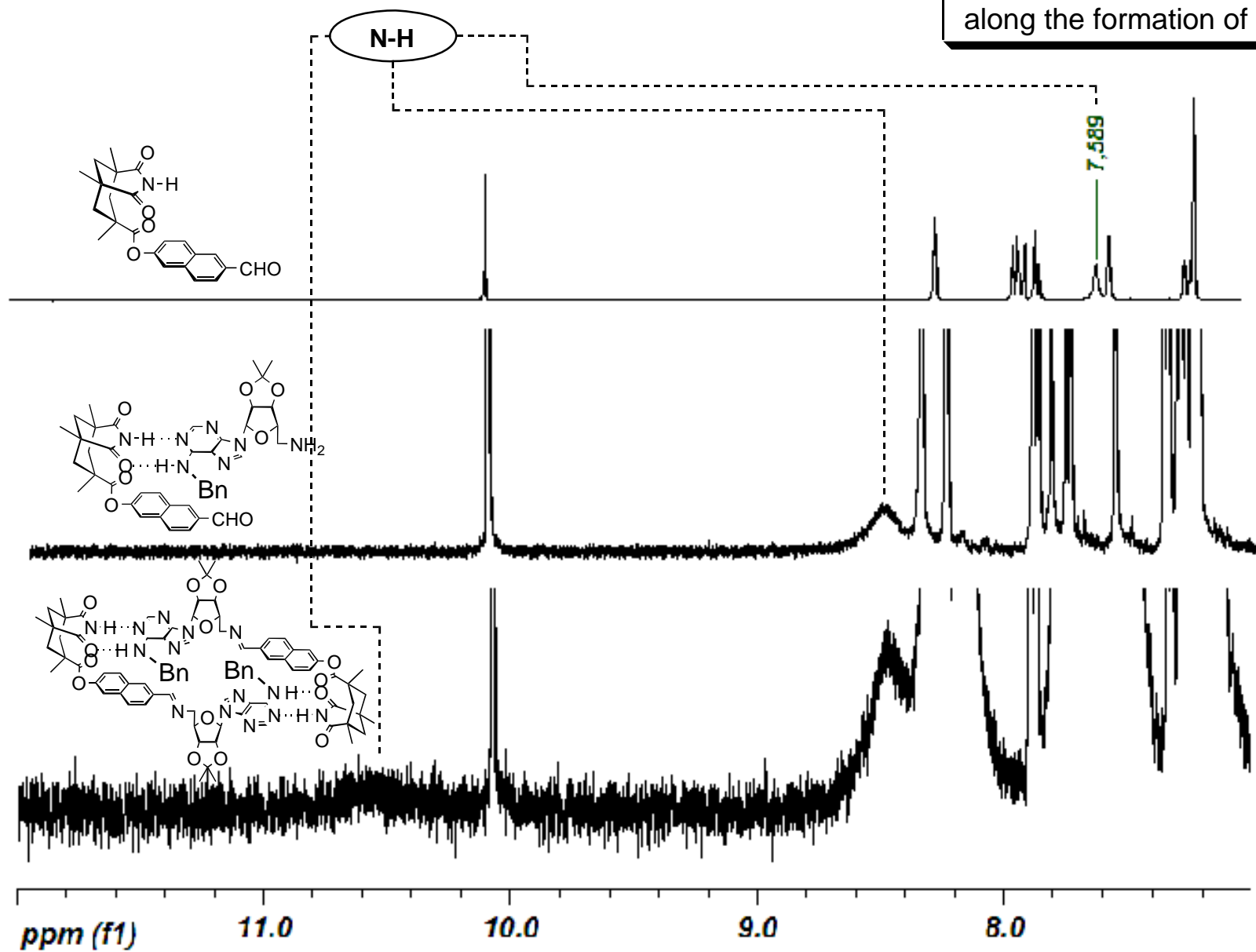Chemical shift of **Al₁** Imide proton as a function of its complexation along the formation of **Al₁Am₁** (c = 15mM)

Chemical shift of **Al₁** Imide proton as a function of its complexation along the formation of **Al₁Am₁** (c = 1.4mM)

Chemical Shift of **AI₁** in presence of **Am₁** in **CDL₂**

Mixture of
**AI₁+AI₂+AI₃**

Mixture of
**AI₁+AI₂+AI₃+Am₁+Am₂**

Chemical shift of $Al_1$ Imide proton as a function of its complexation along the formation of $Al_1Am_2$ (c = 15mM)

N-H

7,589

ppm (f1)   11.0   10.0   9.0   8.0

Service de spectrometrie de masse - Institut de Chimie - Strasbourg - UMR 7177 CNRS / ULP

**Analysis Info**

| | | | |
|---|---|---|---|
| Analysis Name | O3783RBG.d | Acquisition Date | 8/1/2007 2:30:08 PM |
| Method | esi low pos.m | Operator | RBG |
| Sample Name | XU3 | Instrument | micrOTOF |
| Comment | | | |

**Acquisition Parameter**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Source Type | ESI | Capillary | 4500 V | Nebulizer | 0.3 Bar | Corona | 244 nA |
| Ion Polarity | Positive | Set Capillary Exit | 120.0 V | Dry Gas | 5.0 l/min | Set Hexapole RF | 150.0 V |
| Scan Range | n/a | Set Skimmer 1 | 50.0 V | Dry Heater | 200 °C | APCI Heater | 517 °C |

+MS, 1.3-2.2min #(88-145)

**Dimer+H⁺**

1363.6020

1364.6183

1365.6066

1366.6036

**ESI-TOF** Exact Mass Spectroscopy
of homodimer **[Al$_1$Am$_1$]$_2$**

Theoretical
Exact Mass: 1363,58947

$H^+$

**Dimer+Li⁺**

1369.6217

1370.6216

1371.6125

Bruker Daltonics DataAnalysis 3.1          printed:     8/2/2007     10:27:41 AM          Page 1 of 1
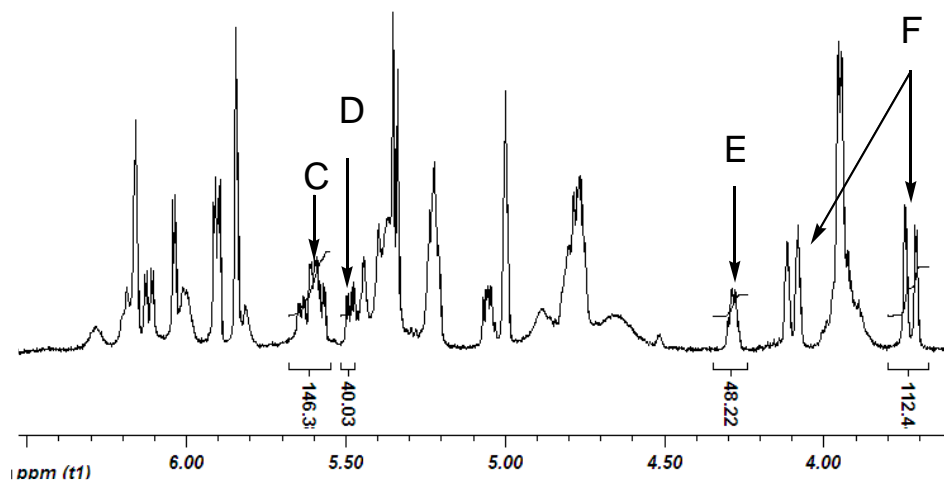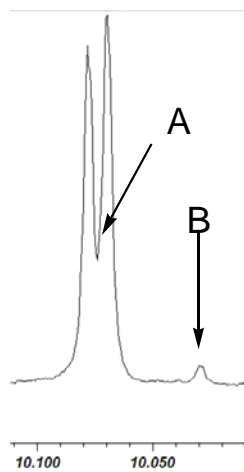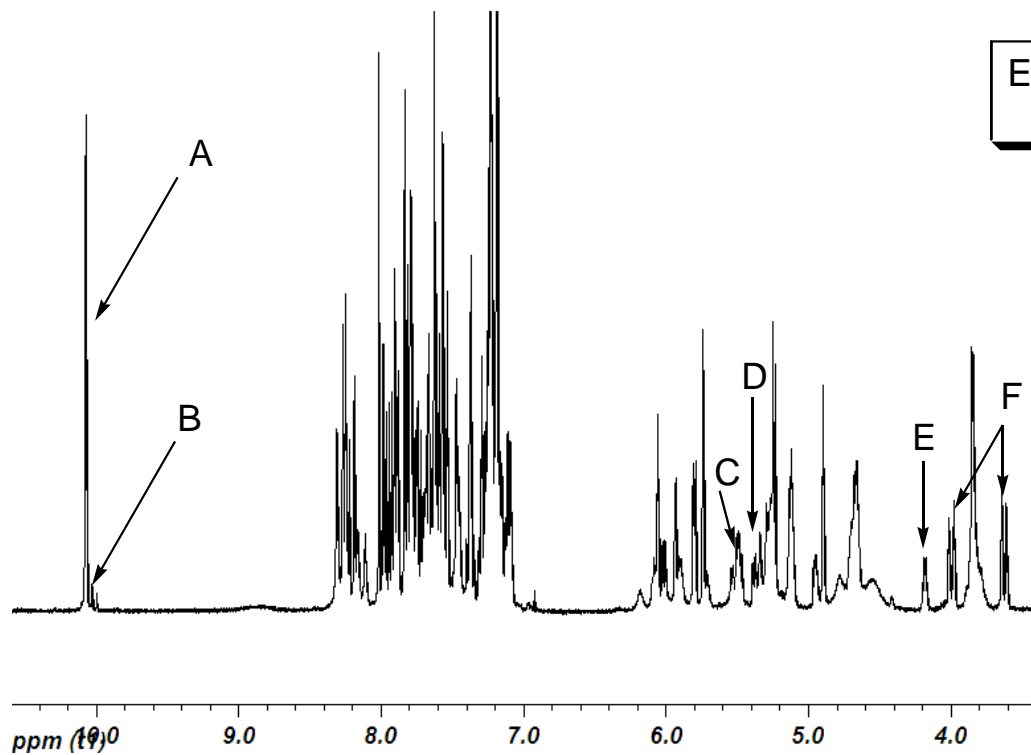
S22

Superimposition of the time course of the 6 individual reactions producing $Al_{(1-3)}Am_{(1,2)}$. Determination of equilibration constants and initial rates of reaction (c = 15 mM) (see Table 1)

Initial rates:

- ■ $V_0Al_1Am_1$ = 5.9mM/h
- ● $V_0Al_1Am_2$ = 1.8mM/h
- ▲ $V_0Al_2Am_1$ = 0.48mM/h
- ▼ $V_0Al_2Am_2$ = 0.35mM/h
- ◆ $V_0Al_3Am_1$ = 0.53mM/h
- ◀ $V_0Al_3Am_2$ = 0.38mM/h

Equilibrium constants:

- ■ $KAl_1Am_1$ = 300
- ● $KAl_1Am_2$ = 50
- ▲ $KAl_2Am_1$ = 17
- ▼ $KAl_2Am_2$ = 10
- ◆ $KAl_3Am_1$ = 20
- ◀ $KAl_3Am_2$ = 12

# S23



Extended regions of $^1$H NMR spectrum for **DCL2** at t = 16 h

1/ Attribution of the resonance signals from spectra of individual compounds
A:  $\mathbf{Al_2 + Al_3}$
B:  $\mathbf{Al_1}$
C:  $\mathbf{Al_1Am_2 + Al_2Am_1 + Al_2Am_2 + Al_3Am_1 + Al_3Am_2}$
D:  $\mathbf{Am_2}$
E:  $\mathbf{Am_1 + Am_2}$
F:  $\mathbf{Al_1Am_1}$

> Protocole for the determination of the concentration
> of constituents $\mathbf{Al_{(1\text{-}3)}Am_{(1,2)}}$ in **DCL2** and **DCL3** by deconvolution of the $^1$H NMR spectra

2/ Because of the constant ratio $\mathbf{Al_2/Al_3}$ all over the course of the reaction, we assume that:
$\mathbf{Al_2 = Al_3}$ at any time

3/ Because of the very close values of initial rates and equilibrium constants for $\mathbf{Al_2Am_1}$, $\mathbf{Al_3Am_1}$, $\mathbf{Al_2Am_2}$, and $\mathbf{Al_3Am_2}$, in individual reactions as well as in **DCL1**, and because of (a) the antagonistic relationship between $\mathbf{Al_1Am_1}$ and $\mathbf{Al_2Am_1}$; $\mathbf{Al_3Am_2}$, and (b) the agonistic relationship between $\mathbf{Al_1Am_1}$ and $\mathbf{Al_2Am_2}$; $\mathbf{Al_3Am_2}$, we assume that:
$\mathbf{Al_2Am_1 = Al_3Am_1}$ and
$\mathbf{Al_2Am_2 = Al_3Am_2}$ at any time

4/ Thus, it comes that:
$\mathbf{Al_1Am_1} = F$
$\mathbf{Al_1Am_2} = [(C+E+F) / 2] - B - F$
$\mathbf{Al_2Am_1} = \mathbf{Al_3Am_1} = [(C+E+F) / 2] - (E-D) - F$
$\mathbf{Al_2Am_2} = \mathbf{Al_3Am_2} = [(C+E+F) / 2] - D - \mathbf{Al_1Am_2}$

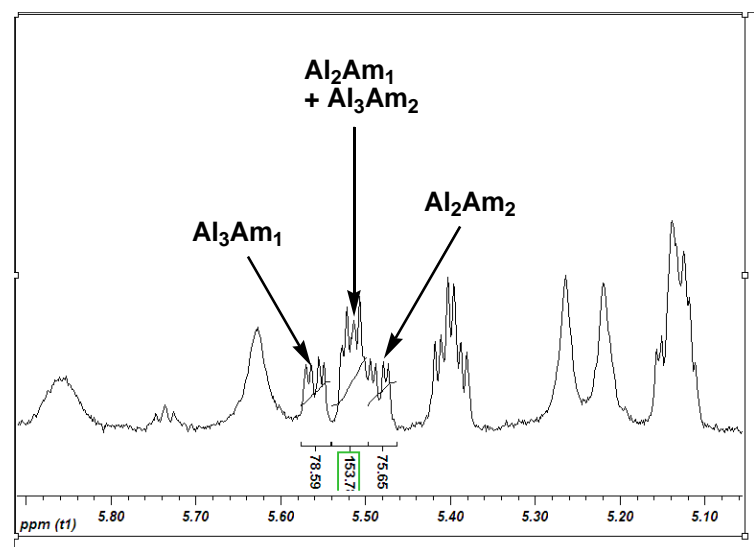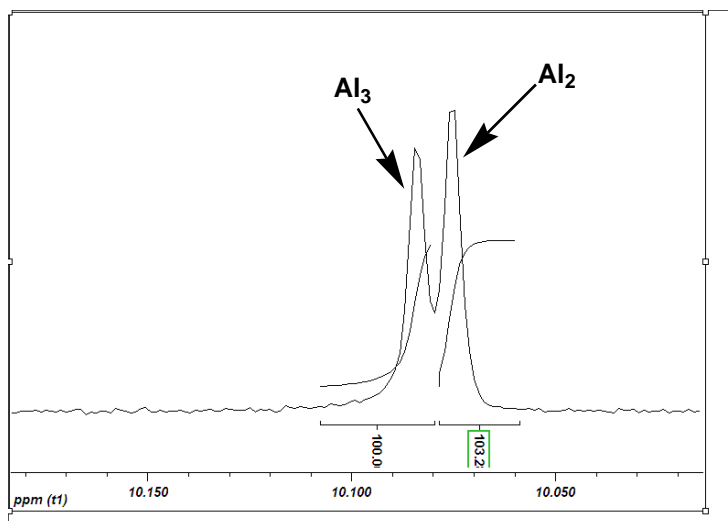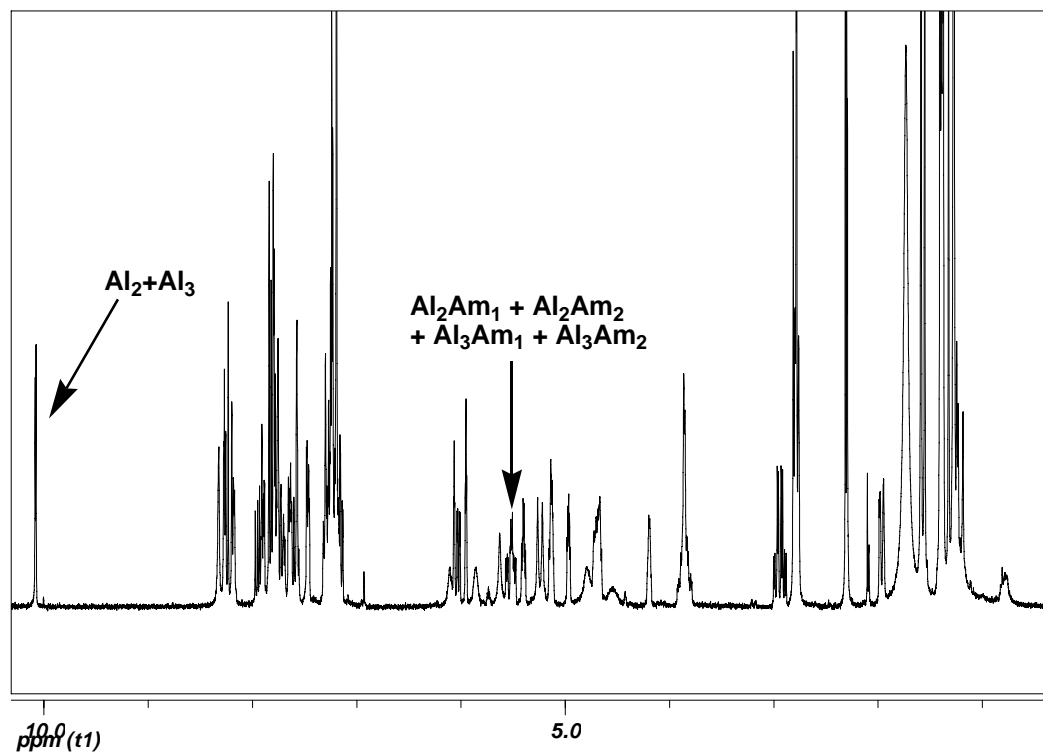$\mathbf{Al_1} = B$
$\mathbf{Al_2} = \mathbf{Al_3} = A / 2$
$\mathbf{Am_1} = E - D$
$\mathbf{Am_2} = D$

5/ For each set of values at a given time, we checked that the following equalities are respected:
$\mathbf{Al_1 + Al_1Am_1 + Al_1Am_2 = Al_2 + Al_2Am_1 + Al_2Am_2 = Al_3 + Al_3Am_1 + Al_3Am_2 = Am_1 + Al_1Am_1 + Al_2Am_1 + Al_3Am_1 = Am_2 + Al_1Am_2 + Al_2Am_2 + Al_3Am_2}$

Extended regions of $^1$H NMR spectrum for **DCL1** at t = 20 h

$Al_2+Al_3$

$Al_2Am_1 + Al_2Am_2 + Al_3Am_1 + Al_3Am_2$

10.0
ppm (t1)

5.0

$Al_3$

$Al_2$

10.150     10.100     10.050
ppm (t1)

100.0     103.2

$Al_2Am_1 + Al_3Am_2$

$Al_3Am_1$

$Al_2Am_2$

5.80   5.70   5.60   5.50   5.40   5.30   5.20   5.10
ppm (t1)

78.59   153.7   75.65

At beginning the 4 starting materials were mixed in an equimolar ratio, thus:

$Al_2Am_1 + Al_2Am_2 + Al_2 = Al_3Am_1 + Al_3Am_2 + Al_3$

at any time

From NMR integration data:

$Al_2 = Al_3$ and
$Al_3Am_1 = Al_2Am_2$ and
$Al_3Am_1 + Al_2Am_2 = Al_2Am_1 + Al_3Am_2$

at any time

Thus, it comes that :

$Al_2Am_1 = Al_2Am_2 = Al_3Am_1 = Al_3Am_2$

at any time

Protocole for the determination of the concentration of constituents $Al_{(2-3)}Am_{(1,2)}$ in **DCL1** by deconvolution of the $^1$H NMR spectra

Determination of the initial rates of reaction of constituents $Al_{(2-3)}Am_{(1,2)}$ in **DCL1** from $^1$H NMR spectra (c = 15 mM)



Figure: Concentration / mM vs Time / h. $V_0 = 0.45$ mM / h. Legend: □ $Al_2Am_1$, ○ $Al_2Am_2$, △ $Al_3Am_1$, ▼ $Al_3Am_2$.