Demystifying the Soai Reaction

Soumitra V. Athavale,[a] Adam Simon,[b] Kendall N. Houk,[b] and Scott E. Denmark[a]*

[a]*Roger Adams Laboratory, Department of Chemistry, University of Illinois, Urbana, Illinois, 61801, United States.* [b] *Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States*

*sdenmark@illinois.edu

Abstract:

The extraordinary Soai reaction has profoundly impacted chemists' perspective of chiral symmetry breaking, absolute asymmetric synthesis and its role in the origin of biological homochirality. Herein, we describe the unprecedented observation of asymmetry amplifying autocatalysis in the alkylation of 5-(trimethylsilylethynyl)pyridine-3-carbaldehyde using diisopropylzinc. Kinetic studies with a "Trojan-horse" substrate and spectroscopic analysis of a series of zinc-alkoxides that incorporate specific structural mutations reveal a 'pyridine-assisted cube escape'. The new cluster functions as a catalyst that activates the '*floor-to-floor*' bound aldehyde and poises a coordinated diisopropylzinc moiety for alkyl group transfer. Transition-state models leading to both the homochiral and heterochiral products were validated by density functional theory calculations. Moreover, experimental and computational analysis of the heterochiral complex provides a definitive explanation for the non-linear behavior of this system. Our deconstruction of the Soai system contributes substantially to understanding the mechanism of this transformation that has stood as a longstanding challenge in chemistry.

Introduction

The autocatalytic, asymmetry amplifying alkylation of pyrimidine-5-carbaldehydes with diisopropylzinc, crowned as the 'Soai reaction', occupies a venerable position in organic chemistry (Figure 1a).[1,2] The transformation is peerless in its efficiency for asymmetric autocatalysis and sensitivity to initial chiral imbalances, making it a *sui generis* example of a reaction predisposed to evolve toward homochirality. Enantiopure products are obtained in three reaction cycles even with calculated autocatalyst e.e. as low as 5×10^{-5} %.[3] In the absence of added catalyst, symmetry breaking can yield non-racemic products,[4-7] thus categorizing the transformation as an example of absolute asymmetric synthesis.[8] A large variety of chiral additives[1,9-15] and even circularly polarized light[16] and isotopic chirality[17-19] can influence the outcome of the reaction by biasing an initial imbalance toward one of the enantiomers. Soai's seminal discoveries have received widespread attention in diverse chemical fields and have revived discussions regarding absolute asymmetric synthesis, chiral symmetry breaking, and the origin of biological homochirality.[20-23]

Since the discovery of the reaction in 1995, an evolving mechanistic understanding has motivated the current hypothesis of a square-macrocycle-square (SMS) tetramer, product-alkoxide aggregate (Figure 1b) as the likely autocatalyst. Initial kinetic studies suggested a dimeric (2001)[24] alkoxide aggregate as the catalyst. A revised proposal (2003)[24] invoking a dimeric alkoxide catalyst and two substrate molecules leading to a tetrameric transition state was subsequently modeled with DFT calculations.[25] Meanwhile, dimer catalysis with a trimeric transition state model was also proposed (Soai, 2003).[26] Reaction simulations (2005, Buhse) demonstrated compatibility with both monomeric and dimeric catalysts, suggesting inherent limitations of kinetic analysis.[27,28] The first spectroscopic studies on the zinc alkoxide were
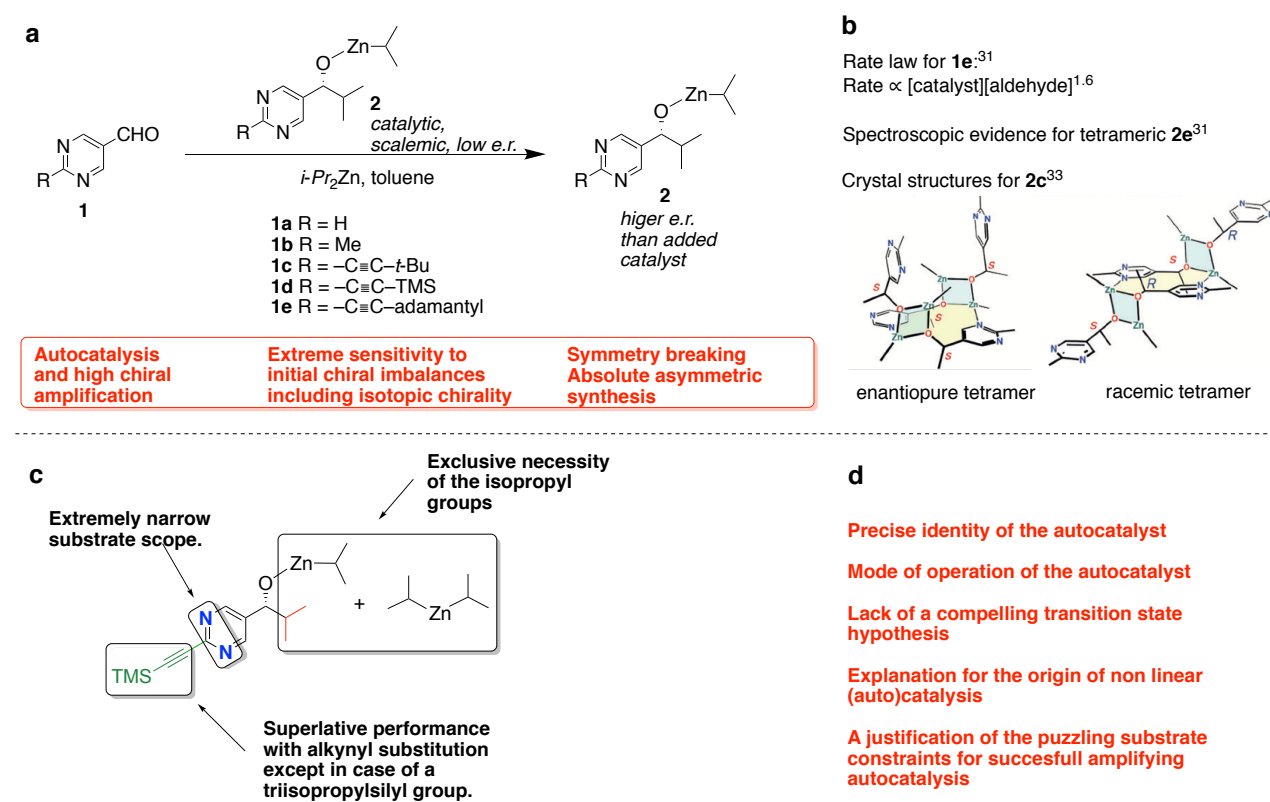

a.

1a R = H
1b R = Me
1c R = –C≡C–*t*-Bu
1d R = –C≡C–TMS
1e R = –C≡C–adamantyl

catalytic, scalemic, low e.r.

i-Pr₂Zn, toluene

2 *higer e.r. than added catalyst*

Autocatalysis and high chiral amplification

Extreme sensitivity to initial chiral imbalances including isotopic chirality

Symmetry breaking Absolute asymmetric synthesis

b.

Rate law for **1e**:[31]
Rate ∝ [catalyst][aldehyde]$^{1.6}$

Spectroscopic evidence for tetrameric **2e**[31]

Crystal structures for **2c**[33]

enantiopure tetramer   racemic tetramer

c.

Extremely narrow substrate scope.

Exclusive necessity of the isopropyl groups

Superlative performance with alkynyl substitution except in case of a triisopropylsilyl group.

d.

Precise identity of the autocatalyst

Mode of operation of the autocatalyst

Lack of a compelling transition state hypothesis

Explanation for the origin of non linear (auto)catalysis

A justification of the puzzling substrate constraints for succesfull amplifying autocatalysis

Figure 1: **a.** The Soai reaction system with its unique characteristics. **b.** Current state of structural and mechanistic understanding. **c.** Structural idiosyncrasies. **d.** Unsolved challenges and open questions.

consistent with a dimeric structure but were limited in scope because of the challenges associated with solubility and peak broadening in toluene, which is the reaction solvent (Brown, 2004).[29] In 2007, Klankermayer *et al.* suggested a square-macrocycle-square (SMS) tetramer as a possible end product species.[30] In 2010, Brown and Blackmond reported an intriguing inverse dependence of reaction rate on temperature and provided further support for the SMS tetramer as a viable species in the reaction cycle on the basis of further kinetic and spectroscopic investigations[31] as well as a subsequent DFT study on these tetramers by Gridnev.[32] Finally in 2015, Soai disclosed the first X-ray crystal structures (Figure 1b) of the enantiopure and racemic alkoxides which exist as tetramers or oligomers depending on crystallization conditions.[33]

The remarkable robustness of the Soai reaction is rivaled by the severe structural constraints on the reactants for successful asymmetric autocatalysis (Figure 1c).[34] Although systematic studies with other dialkylzinc reagents are lacking, asymmetric amplification is observed exclusively with diisopropylzinc. 2-Alkynylpyrimidine-5-carbaldehydes[35] remain the workhorse substrates and only scant reports with other competent aldehydes have been presented.[36,37]. Pyridine-3-carbaldehyde, for example does not demonstrate amplifying autocatalysis.[38] Note that an inspection of the crystal structure (see Figure 1b) does not reveal the significance of the second pyrimidine nitrogen in assembling the macrocyclic connectivity nor the role played by the diisopropylzinc bound to this nitrogen with respect to catalysis. Similarly, the defining contributions of the alkyl groups towards product aggregation are not apparent. Moreover, the reasons for superlative performance of the alkynyl-substituted substrates remain unclear. Another idiosyncrasy is the bewildering failure of 2-(triisopropylsilylethynyl)pyrimidine-5-carbaldehyde to display amplifying autocatalysis.[35] For a reaction demonstrating such striking characteristics, the extremely narrow substrate scope is perhaps unsurprising. Nevertheless, a fundamental rationalization of these structural constraints that allow exclusive access to the requisite catalytic intermediates remains obscure highlighting our lack of understanding of this reaction. Although the consensus of a tetrameric alkoxide catalyst represents a significant milestone in the identification of product aggregation, this knowledge has not enabled the formulation of a compelling transition state hypothesis nor provided deeper insights into the structural contributions from various parts of the molecule that give rise to asymmetric autocatalysis.

To summarize, the astonishing characteristics of this transformation pose a set of unique mechanistic questions that have challenged investigators for more than two decades (Figure 1d). Principle among these is: (1) an elucidation of the precise identity and *modus operandi* of the autocatalyst, (2) a transition state structure rationalizing the basis of asymmetric autocatalysis, (3) an explanation for the origin of positive, non-linear (auto)catalysis, and, (4) a justification of the puzzling, restrictive and idiosyncratic substrate requirements that allow successful amplifying autocatalysis. A holistic comprehension of the reaction remains an open challenge with fundamental implications on our understanding of chiral symmetry breaking and absolute asymmetric synthesis. Our work addresses these questions (*vida supra)* that lie at the heart of unmasking the mechanism of the Soai reaction.

Discovery of a New Amplifying Autocatalytic System
The first stage of our investigation was to deconstruct the Soai system and understand the contributions to catalysis and asymmetric amplification from its various components: (1) the dialkylzinc reagent, (2) the carbinol substituent, (3) the aromatic core and (4) the alkynyl substituent. Orienting studies began with the well-known pyrimidine-5-carbaldehyde **1d** and

different dialkylzinc reagents.[39] The next series of experiments probed the role of the aromatic core, which led us to the startling discovery[40] that *5-(trimethylsilylethynyl)pyridine-3-carbaldehyde (3) displayed efficient amplifying asymmetric autocatalysis in the reaction with diisopropylzinc* (Figure 2a). Product carbinol **4**, when included in this reaction as a scalemic catalyst, afforded newly-generated product with higher enantioenrichment.[41,42] The reaction is homogenous even at high concentrations and is conveniently monitored by tracking aldehyde consumption using *in-situ* IR spectroscopy.

The transformation displays all the characteristics of an amplifying, autocatalytic process (Figure 2b). Without added catalyst, sigmoidal aldehyde consumption, with an initial induction period, followed by a rapid reaction is observed, characteristic of autocatalysis (entry 1, blue dashed curve). Inclusion of scalemic product carbinol **(4)** eliminates the induction period, leading to the rapid production of product with higher enantioenrichment than the initially added autocatalyst (entry 2, red dashed curve). At lower concentrations, the background reaction without added product is sluggish and inefficient (entry 3, solid blue curve). However, inclusion of either enantioenriched (entry 4, solid red curve) or racemic (entry 5, solid green curve) carbinol **(4)** results in accelerated reactions and efficient product generation. In the case of the enantioenriched autocatalyst (entry 4), highly enantioenriched product is obtained whereas the racemic autocatalyst (entry 5) yields racemic product. Crucially, the latter reaction is much slower than the former indicating the lower catalytic activity of the racemic autocatalyst species, further highlighting the non-linear effect. *These observations establish that, just as in the case in the iconic Soai pyrimidine-system, non-linear amplifying autocatalysis is also operative in the reaction of 3 and diisopropylzinc*.

The new system also shares the other non-obvious, idiosyncratic characteristics typical to the Soai reaction. In the new pyridine system, an inverse rate profile with respect to temperature was noted, with a rate maximum at ~0 °C (Fig. S4-8). The analogous reaction of **3** with diethylzinc did not result in asymmetric amplification or rate enhancement (Figure 2c, entries 4 and 5). Remarkably, the reaction of diisopropylzinc with 5-(triisopropylsilylethynyl)pyridine-3-carbaldehyde was sluggish and demonstrated neither asymmetric amplification nor autocatalytic rate enhancement (Fig. S4-7*). Such striking similarities of the two systems clearly demonstrate that one of the nitrogen atoms in the pyrimidine ring is dispensable in the Soai reaction system.*

Catalyst Structure-Activity Relationships

To evaluate contributions of the alkyl group on the carbinol center in the catalyst and the O-Zn-alkyl group, a qualitative comparison of catalysis by four different zinc alkoxides (**PyII, PyEE, PyIE** and **PyEI)** was made by *in-situ* IR monitoring of dialkylzinc additions to **3** in catalyzed and uncatalyzed reactions (Figure 2c) (nomenclature: *Py/Ph* = pyridine/phenyl indicating the aromatic core, *I/E* = isopropyl/ethyl indicating the carbinol alkyl group and *I/E* = isopropyl/ethyl indicating the alkylzinc group). In all cases, the corresponding scalemic carbinols were used for generating the zinc alkoxides (see Supplementary Information for details) and the product enantiomeric composition was analyzed to assess any non-linear effects. Under these conditions, the uncatalyzed reactions (entries 1 and 4) are sluggish, show low conversion and ultimately stall. Inclusion of either **PyEE** (entry 5) or **PyIE** (entry 6) shows no appreciable rate enhancement. In the case of **PyEE**, the product carbinol has a lower e.r. than the added catalyst whereas a minor positive non-linear effect is seen in case of **PyIE**. Entry 2 represents the autocatalytic reaction with diisopropylzinc and reiterates the high catalytic efficiency and strong non-linear behavior of catalyst **PyII**. Finally, with alkoxide **PyEI** (entry 3), a sigmoidal aldehyde
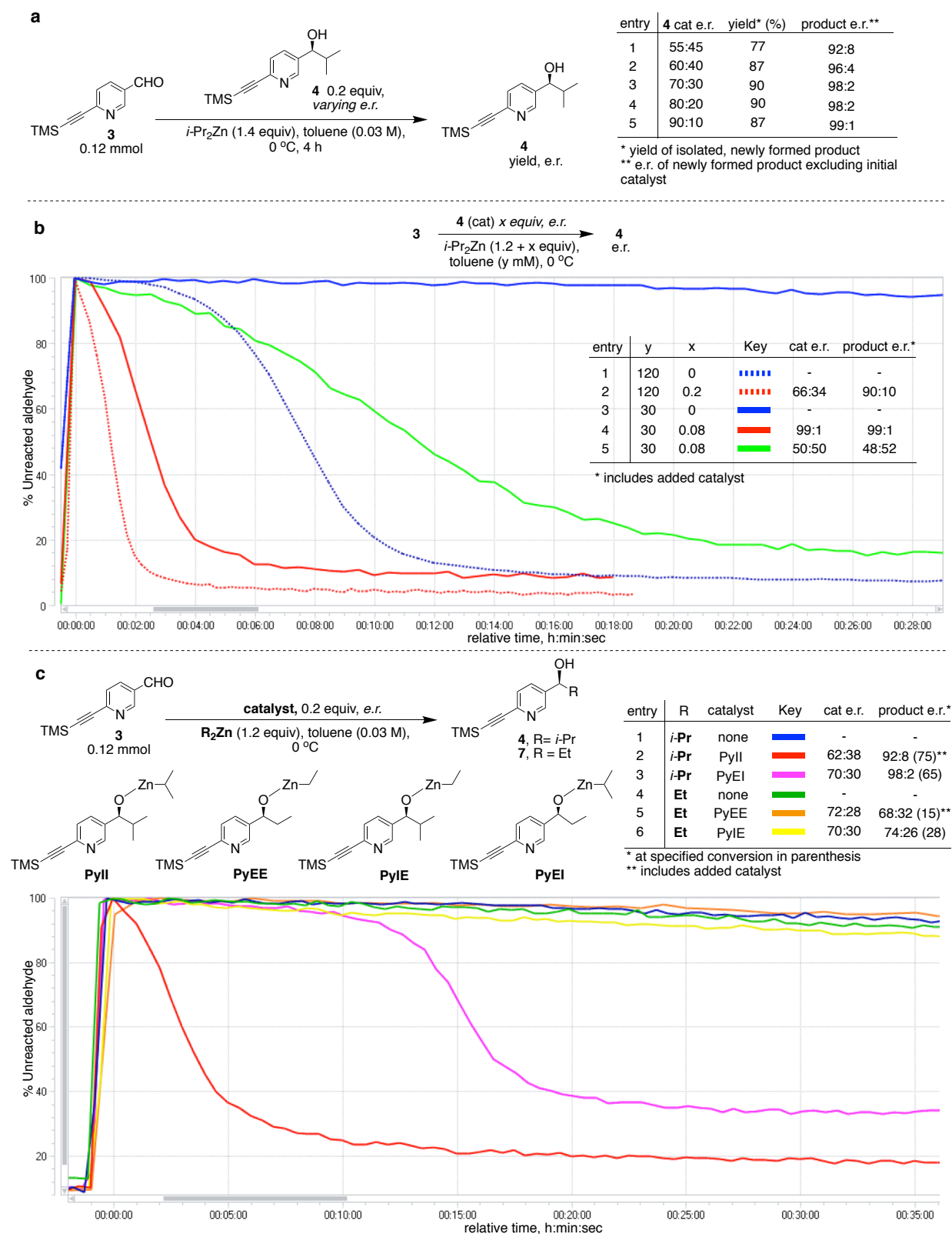
Figure 2: **a.** Asymmetric autocatalysis with chiral amplification in alkylation of **3** with diisopropylzinc. **b.** Observation of sigmoidal aldehyde consumption profiles and non-linear autocatalysis upon monitoring aldehyde consumption by *in-situ* IR spectroscopy **c** Qualitative comparison of catalytic activities and non-linear effect of alkoxides **Py*XX*** in dialkylzinc addition to **3** by ReactIR monitoring of aldehyde consumption

decay with high final product e.r. is achieved. The absence of an initial rate enhancement is noteworthy. This behavior is interpreted to arise from a low catalytic activity of **PyEI** giving rise to an initial production of **PyII** followed by strong asymmetric autocatalysis from a buildup of **PyII**. Taken together, these observations indicate that in this alkoxide series, only **PyII** possesses the requisite structural requirements for efficient, amplifying catalysis and replacement of either or both isopropyl groups results in alkoxides (**PyEE, PyIE** and **PyEI**) that possess markedly inferior catalytic efficiencies.

Spectroscopy of Zinc Alkoxides and the Structural Uniqueness of PyII

To gain insights into the structural requirements for efficient catalysis, occurring only in case of **PyII,** a spectroscopic study to characterize the solution state structures of phenyl (**PhEE, PhEI, PhIE** and **PhII**) and pyridyl (**PyEE, PyEI, PyIE** and **PyII**) zinc alkoxides in toluene was initiated (Figure 3). The high solubility of these species along with their stability in toluene for extended periods represented a significant advantage over their pyrimidine analogs and facilitated detailed 1D and 2D NMR spectroscopic characterization.

In general, for all alkoxides, spectra obtained from racemic samples showed greater multiplicity of signals and additional peaks compared to spectra obtained from enantioenriched samples, proving that none of these alkoxides exist as monomers (Supplementary Information, Section 3). Figure 3a presents the upfield region of the eight, enantioenriched alkoxides to allow comparison of important spectral characteristics. To assist this comparison, alkoxides differing by only a nitrogen atom in the aromatic core are assigned the same color and are arranged successively. Different colors highlight four possible combinations of the carbinol-alkyl and O-Zn-alkyl groups. For the enantioenriched phenyl alkoxides **PhEE** (spectrum 1), **PhEI** (spectrum 3) and **PhIE** (spectrum 5), a single set of peaks corresponding to a predominant species was observed (Supplementary Information, Section 3 for complete spectral comparisons). The signature characteristics of this species are the sharp, clearly resolved peaks with only minor perturbation over a wide temperature range (-20 – 50 ^{o}C, Supplementary Information, Section 3) and the atypical *upfield* chemical shifts of the alkylzinc protons H(7) and H(8) (Figure 3a, black boxes) in comparison to the free dialkylzinc protons H(9) and H(10).

For alkoxide **PhEI**, a remarkable quadruplet splitting pattern in the racemic aggregate is observed in comparison to the enantiopure aggregate (Figure 3b). Such a pattern is consistent with a symmetric, cubic tetramer structure, which finds ample precedent.[43,44] The racemic cubic tetramer can exist as a mixture of three diastereomers depending on the configurations at the four constituent alkoxide stereogenic centers: (1) chiral (D_2), (2) meso (S_4) and (3) heterochiral (C_1) (Figure 3b). On the basis of the statistical populations of these three species, the ideal 1:3:3:1 peak pattern seen here provides compelling evidence for the structure of **PhEI** as a racemic, cubic tetramer.

The atypical upfield shift of the alkylzinc protons is believed to arise from shielding by the aryl groups and is also observed in the ethylzinc alkoxide of (*S*)-1-phenylethanol (Fig. S3-96), the methylzinc alkoxide of which has been proposed to form a cubic tetramer.[45] On the basis of these observations, species **PhEE, PhEI** and **PhIE** are believed to form cubic tetramers (species **a**, Figure 3a) in solution and the upfield chemical shifts of protons H(7) and H(8) (black boxes) are a diagnostic feature of this structure.

The spectrum of pyridylzinc alkoxide **PyEE** (spectrum 2) shows a predominant species matching the spectral characteristics of a cubic tetramer. Clearly, presence of the pyridine nitrogen in **PyEE** does not disrupt the natural preference to form a cubic tetramer. Increasing the
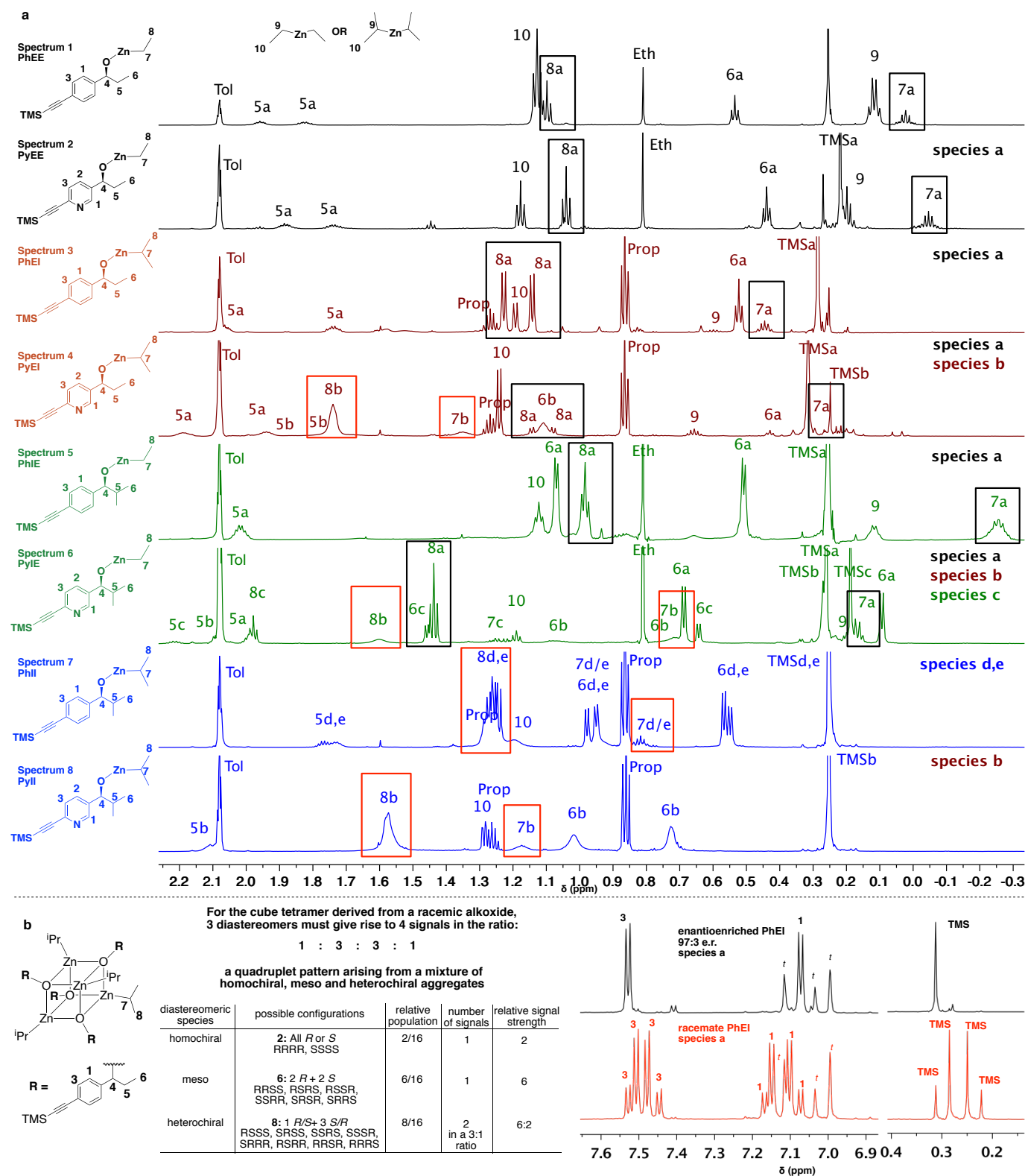
Figure 3: Partial ^{1}H-NMR spectra (750 MHz) of zinc alkoxides in toluene-d_8 at 23 [0]C. *tol* = residual toluene, *prop* = propane, *eth* = ethane. **a.** Enantioenriched zinc alkoxides with signature chemical shifts for species **a** (black boxed) and species **b-d** (red boxed) **b.** Diastereomeric species possible in a racemic cubic tetramer with their relative peak distributions. ^{1}H NMR of the aryl and TMS region for racemic **PhEI** displays this predicted quadruplet pattern.

bulk of the alkyl groups in the pyridyl series in species **PyEI** (spectrum 4) and **PyIE** (spectrum 6) still maintained the cubic tetramer assembly (species **a**, black boxed diagnostic peaks) but also showed the presence of other species (species **b**, **c**), the peaks of which are labeled with the suffix 'b' or 'c' in Figure 3a (*c.f.* H(**7b**), H(**6c**)). Broad peaks with *downfield* shifts of protons H(7) and H(8) (red boxes) characterize species **b**. A third species (**c**) with sharp peaks and downfield alkylzinc protons can be identified in the spectrum of **PyIE** (spectrum 6). In the spectra for the bulkiest derivatives of both phenyl and pyridyl complexes **PhII** (spectrum 7) and **PyII** (spectrum 8) respectively, diagnostic peaks for the cubic tetramer are completely absent. Two components can be detected in the spectrum of **PhII** (species **d**, **e**) whereas the single species with broad peaks seen in the spectrum of **PyII** resembles species **b** (in the spectra of **PyEI** and **PyIE**) in chemical shift patterns (red boxed). Species **b** and **c** also show diagnostic peak differences in the aryl region in comparison to species **a** (Fig. S3-48).

Diffusion Ordered Spectroscopy (DOSY) studies of the alkoxides for molecular weight estimations (Supplementary Materials, Section 3.11, Tables S3-4, S3-5) established that all of these compounds except **PhII** are tetrameric and are unchanged over a temperature range of 0 to 50 ^{0} C, highlighting the stability of these aggregates. DOSY analysis showed that **PhII** exists as a mixture of trimers. This evidence provides further support to the hypothesis that species **a** are cubic tetramers and reveals that species **b** and **c** represent distinct aggregate structures that are also tetrameric.

Rationalizing the effect of structure on alkoxide aggregation (Figure 4) begins with the assumption that a cubic tetramer is the preferred constitution for these alkylzinc alkoxides. However, replacement of ethyl groups around the cube core by bulkier isopropyl groups results in destabilization of the cube structure by steric repulsion to the point at which the cube is no longer viable as in cases of **PhII** and **PyII**. A weakened cube can be further disrupted by a pyridine nitrogen through coordination to the zinc alkyls to assist in the formation of alternative tetrameric aggregates. Indeed, addition of pyridine to solutions of **PhEI, PhIE** and **PhII** results in disruption of the aggregate structure whereas **PhEE** remains unaffected (Figure 4a) (Fig. S3-51). With a pyridine nitrogen available for coordination to zinc, alkoxides **PyEE, PyEI, PyIE** and **PyII** are poised to *escape the cube* only if the core is sufficiently weakened. The resulting continuum of structures displayed by these alkoxides represents this phenomenon. Whereas **PyEE** maintains a cubic tetramer constitution because its core is sufficiently stable to resist pyridine disruption, **PyEI** and **PyIE** exhibit formation of other tetrameric species owing to partial cube escape facilitated by nitrogen coordination. Finally, zinc alkoxide **PyII** represents the end point with complete cube escape and access to an exclusive, alternative tetrameric aggregate structure. The simplest way in which such a cube-escape can occur is an intramolecular coordination of the pyridyl nitrogen to *expand* the cube to a square-macrocycle-square (SMS) connectivity (Figure 4b). This connectivity is identical to the crystal structure described by Soai. *Clearly, only a single nitrogen is required in the aromatic core to build this tetramer.* In light of the striking similarity of **PyII** to the structure and properties of the original Soai system **2d**, as well as the rationale for the assembly of the SMS tetramer, the proposal that its identity is analogous to that found in the recently disclosed X-ray crystal structure of **2c** (the 'Soai tetramer') seems compelling (Figure 4c).
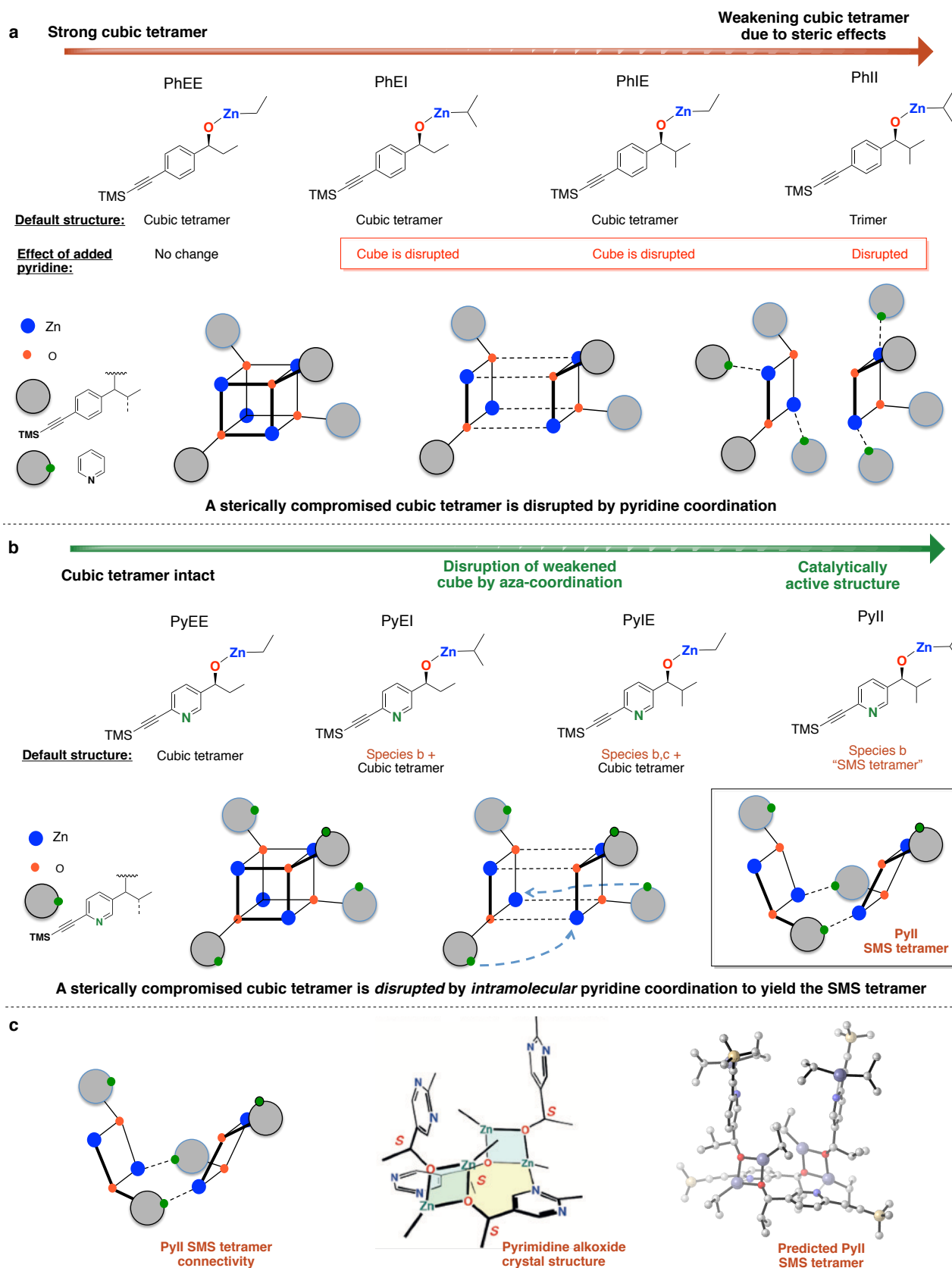
a Strong cubic tetramer — Weakening cubic tetramer due to steric effects

PhEE — PhEI — PhIE — PhII

Default structure: Cubic tetramer — Cubic tetramer — Cubic tetramer — Trimer

Effect of added pyridine: No change — Cube is disrupted — Cube is disrupted — Disrupted

Zn / O

A sterically compromised cubic tetramer is disrupted by pyridine coordination

b Cubic tetramer intact — Disruption of weakened cube by aza-coordination — Catalytically active structure

PyEE — PyEI — PyIE — PyII

Default structure: Cubic tetramer — Species b + Cubic tetramer — Species b,c + Cubic tetramer — Species b "SMS tetramer"

Zn / O

PyII SMS tetramer

A sterically compromised cubic tetramer is *disrupted* by *intramolecular* pyridine coordination to yield the SMS tetramer

c

PyII SMS tetramer connectivity — Pyrimidine alkoxide crystal structure — Predicted PyII SMS tetramer

Figure 4: **a.** Sterically weakened **PhXX** cubic tetramers are disrupted by pyridine. **b.** A conceptually similar 'cube escape' occurs in the **PyXX** alkoxides with intramolecular pyridine coordination yielding the SMS tetramer. **c.** The predicted identity of the cube-escaped **PyII** SMS tetramer by analogy with the Soai pyrimidine alkoxide crystal structure

Reactant Constraints of PyII and Kinetics with a Trojan-Horse Substrate

We discovered that **PyII** can maintain its critical attributes of enantioselective catalysis and positive non-linear effect in the non-autocatalytic alkylation of pyridine-3-aldehyde (**5**) with diisopropylzinc (Figure 5, entries 1 and 2). However, the aldehyde decay profile is reminiscent of a conventional catalyst-substrate reaction because the newly formed product contributes minimally to catalysis (Fig S4-6)[38]. This reaction, in which pyridine-3-aldehyde acts as a surrogate, '*Trojan-Horse*' substrate, is thus representative of the alkyl transfer step in the autocatalytic reaction, and was employed to provide insights into the mode of action of the **PyII** catalyst. *In-situ* IR monitoring of aldehyde consumption was employed for a comprehensive initial rate kinetic analysis of the diisopropylzinc alkylation of **5** catalyzed by **PyII** (Figure 5b, see Supplementary Information, section 2 for details). The initial rate displayed a first order dependence on catalyst concentration and zeroth order dependence on both **5** and diisopropylzinc, indicating that saturation binding of both reactants to the catalyst aggregate precedes the rate limiting alkyl transfer. We believe that the decrease in conversion rate over time arises from catalyst poisoning and/or product inhibition by the product alkoxide that can mimic the **PyII** monomer and incorporate into the **PyII** aggregate. Arrhenius activation parameters were determined assuming a tetrameric catalyst aggregate. *A significant observation which highlights the substrate specificity of the PyII catalyst was the poor reactivity of benzaldehyde 8 and 4-(trimethylsilylethynyl)benzaldehyde 10 under identical catalytic conditions (entries 3 and 4).*

Substrate Binding and Catalysis

For the hypothesis that **PyII** emulates the 'Soai tetramer' to be valid, it must also provide a clear explanation for the observed saturation kinetic behavior (with **5**) and production of homochiral products from both **3** and **5**. Catalysis must be facilitated by a reactant-binding model that would account for activation of the carbonyl group as well as an incoming diisopropylzinc
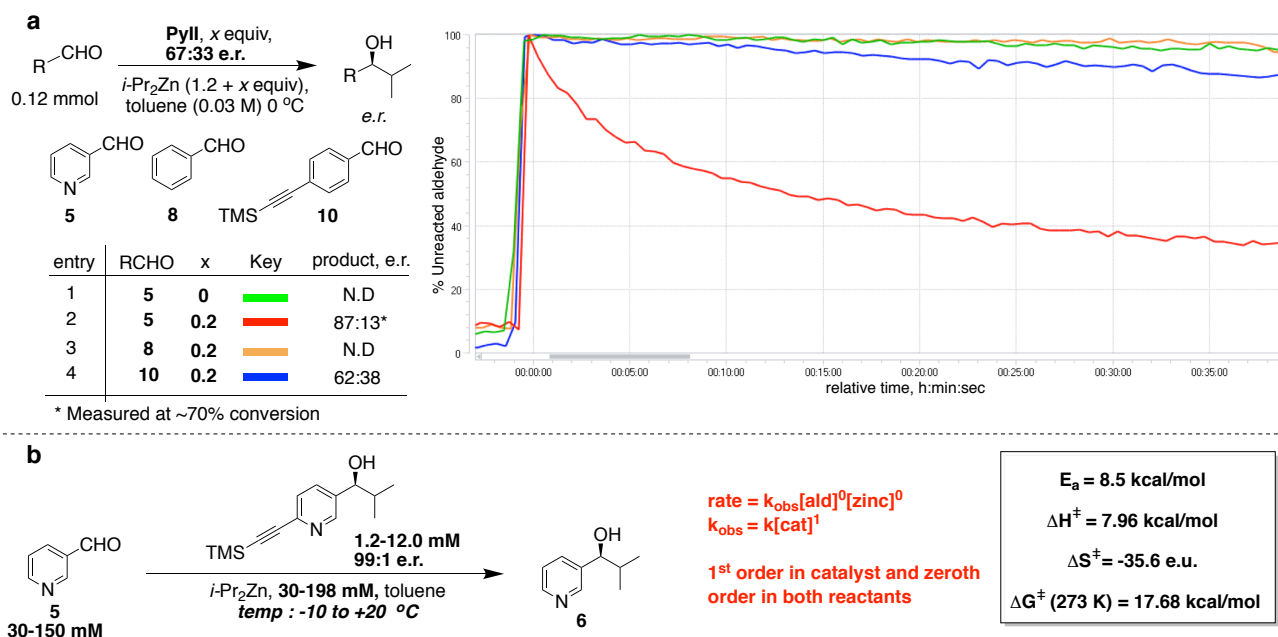

Figure 5: **a.** *In-situ* IR monitoring profiles for aldehyde consumption in diisopropylzinc alkylations catalyzed by scalemic **PyII**. **b.** Rate law and Arrhenius parameters for the non-autocatalytic reaction as measured using initial rate kinetic analysis.

prior to alkyl transfer. The critical demonstration of poor reactivity of aldehydes **8** and **10** (Figure 5a) indicates that pyridine-3-carbaldehydes are a privileged substrate class for the SMS tetramer owing to nitrogen coordination to a favorable site on the catalyst. Independent spectroscopic studies revealed that: (1) the dialkylzinc binds to the pyridine nitrogens of **PyII** (Fig. S3-50) as well as substrates **3** and **5** (Fig. S3-56), (2) the carbonyl group in **5** has an inherently poor affinity to diisopropylzinc and binding is exclusive to the ring nitrogen with diagnostic chemical shift changes (Figs. S3-55, S3-56) and, (3) the carbonyl group in **5** *can* coordinate to an unsaturated zinc in the alkoxide core (Fig. S3-53).

In light of these considerations, the energetic preferences of two proposed models (*floor-to-foor* and *floor-to-arm)* describing reactant docking prior to the alkyl transfer were interrogated with density functional theory (DFT) calculations. The crystal structure for **2e** was employed as the starting geometry for the homochiral **PyII** tetramer after the replacement of pyrimidine cores and *tert*-butyl groups by pyridines and trimethylsilyl groups, respectively. Computational evaluation of the binding of aldehyde **5** to homochiral **PyII** tetramer reveals a preference for the *floor-to-floor* model (**int-5a**, Figure 6a) (for the floor-to-arm model calculations, refer Supplementary Information, Section 5). Alkyl group transfer by diisopropylzinc is primed for delivery to the *Si* face of the aldehyde in **int-5a,** which would lead to the observed (*S*) stereoisomer. Conversely, the *Re* face is accessible only through a rotation of both the substrate and one arm of the tetramer as shown in **int-5b**, which is unfavorable by 5.9 kcal/mol. The stereocontrolling transition structures **TS-5a** and **TS-5b** arising from **int-5a** and **int-5b**, and having activation free energy barriers of 12.5 kcal/mol and 14.7 kcal/mol respectively are shown in Figure 6a. This energy difference is in agreement with the observed stereoselectivity of **PyII** with pyridine-3-carbaldehyde **5** (95:5 e.r., $\Delta\Delta G^{\ddagger}$ = 1.7 kcal/mol, see Table S2-1, Exp 5a-c). Although of no direct stereochemical consequence, it is interesting to note that the isopropyl group transfers (S_E2 reactions) occur with "inversion of configuration" at the methine carbon (*vide infra*).

The transition structures and minima for the autocatalytic reaction of aldehyde **3** and tetramer **PyII** were also investigated computationally (Figure 6b). The ground state geometry **int-3a** is quite similar to **int-5a**, N.B. the bulky aldehyde **3** is conformationally restricted in the tetramer. The lowest-energy transition structure **TS-3a** leads to the alcohol (*S*)-**4**, and the next higher energy transition structure **TS-3b** (4.4 kcal/mol) leads to the (*R*)-**4** product. This large difference in energy comports with an enantioselective autocatalytic mechanism. **TS-3a** is analogous to **TS-5a**, in which facile alkyl transfer (with "inversion") takes place from diisopropylzinc in the ground state geometry **int-3a**. Contrariwise, **TS-3b**, represents an S_E2-type mechanism with "retention of configuration" at the methine of the transferring isopropyl group. This higher-energy pathway occurs because the left arm of the tetramer does not rotate for alkyl transfer to the *Re*-face of the aldehyde as in **int-5b**. This conformation is inaccessible for aldehyde **3** because to arrange this substrate to receive the isopropyl group would require that the trimethylsilylethynyl substituent be in the same space as the opposite arm (N.B., the C5 position of aldehyde **5** in both **int-5b** and **TS-5b**, which corresponds to the trimethylsilylethynyl-substituted position of **3**). A similar type of transition structure was located for aldehyde 5, i.e. without the trimethylsilyethynyl group (**TS-5e**, Fig S5-5), which was 5.2 kcal/mol higher in energy. This observation reveals that the constrained S_E2 substitution with retention of configuration (**TS-3b** and **TS-5e**) is highly disfavored compared to the invertive S_E2 pathway seen in **TS-5a** and **TS-3a**. Thus, the stereoselectivity arises from the steric environment of the
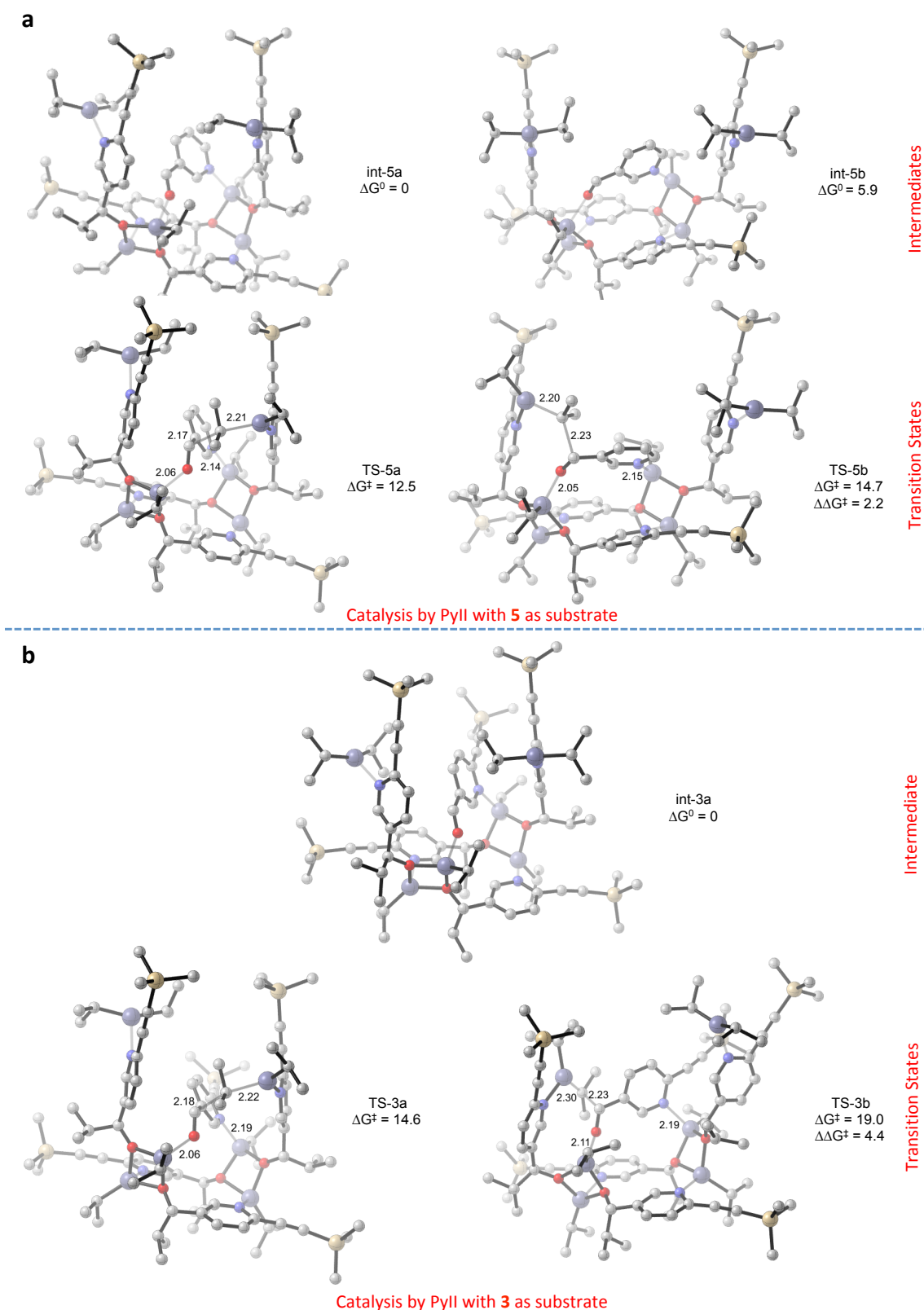
a

int-5a
$\Delta G^0 = 0$

int-5b
$\Delta G^0 = 5.9$

Intermediates

TS-5a
$\Delta G^{\ddagger} = 12.5$

TS-5b
$\Delta G^{\ddagger} = 14.7$
$\Delta\Delta G^{\ddagger} = 2.2$

Transition States

Catalysis by PyII with **5** as substrate

b

int-3a
$\Delta G^0 = 0$

Intermediate

TS-3a
$\Delta G^{\ddagger} = 14.6$

TS-3b
$\Delta G^{\ddagger} = 19.0$
$\Delta\Delta G^{\ddagger} = 4.4$

Transition States

Catalysis by PyII with **3** as substrate

Figure 6: DFT studies for floor-to-floor substrate docking and alkyl transfer by the **PyII** tetramer. Transition structures **TS-3a–3b** and **TS-5a–5b**, and optimized minima **int-3a** and **int-5a–5b** calculated at the M06-2X/def2-TZVPP–SMD (toluene)//B3LYP/6-31G(d) level of theory. Energies reported in kcal/mol. Purple, zinc; yellow, silicon; red, oxygen; blue, nitrogen; gray, carbon. Hydrogens are hidden for clarity. ΔG^0, free energy of activation. $\Delta\Delta G^{\ddagger}$, relative free energy activation barrier.

chiral tetramer, which favors *Si*-face addition to the aldehyde when bound in the preferred *floor-to-floor* arrangement.

Finally, the catalytic competence of the heterochiral (racemic) **PyII** tetramer was computationally investigated (Figure 7). The heterochiral aggregate was modeled according to the Soai crystal structure for racemic **2c** (Figure 1). The heterochiral tetramer is energetically preferred to the homochiral **PyII** tetramer by 2.1 kcal/mol (G_{rel} blue). From inspection of this structure, it is immediately apparent that substrate binding in a two-point, *floor-to-floor* arrangement is impossible because the *requisite, unsaturated zinc atoms are positioned on opposite sides of the macrocycle*. Indeed, the only stable binding mode located for substrate **3** to the heterochiral **PyII** tetramer was a single-point binding leading to **int-rac** (Figure 7a red, ΔG^0 = 14.8 kcal/mol). This complex is highly disfavored in comparison to the *'floor-to-floor'* binding to the homochiral **PyII** tetramer leading to **int-3a** (Figure 7b red, ΔG^0 = 2.9 kcal/mol). Interestingly, the activation barrier for the alkyl transfer **int-rac** to **TS-rac** ($\Delta G^{\ddagger}$ = 14.8 kcal/mol) is nearly identical to the barrier for the homochiral assembly **int-3a** to **TS-3a** ($\Delta G^{\ddagger}$ = 14.6 kcal/mol). Thus, the heterochiral **PyII** tetramer is catalytically incompetent not because the alkyl transfer is unfavorable, but rather because substrate binding is energetically inaccessible (Figure 7c). *Therefore, because the heterochiral **PyII** tetramer is energetically preferred and kinetically incompetent compared to the homochiral **PyII** tetramer ($\Delta G_{het}^{\ddagger} - \Delta G_{homo}^{\ddagger}$ = 12.1 kcal/mol), a mechanism for non-linearity in the asymmetric autocatalytic action of the homochiral **PyII** SMS tetramer is now established.*

Summary

The Soai reaction represents a fortuitous confluence of three phenomena in catalysis: autocatalysis, enantioselective catalysis and non-linear effects (asymmetric amplification). The pyridine-based autocatalytic system **3** provides a conceptual simplification because omission of a nitrogen atom in the aromatic core of the pyrimidine system in turn translates to four fewer coordinating sites on the SMS tetramer. The new system is amenable to detailed kinetic and spectroscopic investigations and presents decisive answers to critical mechanistic questions highlighted at the beginning of this thesis. The experimentally verified pyridine-coordination-assisted 'cube escape' model provides a compelling hypothesis to explain how the SMS macrocycle tetramer is a natural evolutionary endpoint of the pyridinyl (and presumably pyrimidyl) alkoxides. Conversely, the reaction of diethylzinc generates a catalytically inactive cubic tetramer product and fails to display asymmetric autocatalysis. The ring nitrogen atom with a 1,3-relationship to the alkoxide oxygen is indispensable for the assembly of the catalyst by bridging two dimeric units and providing binding sites for reactants on the catalyst arms. Complementarily, this 1,3-relationship in the substrate facilitates a two-point binding mode for aldehyde activation. In other words, we suggest that the SMS tetramer allows catalysis with an 'expansion' of the cubic tetramer, made possible by replacing two Zn-O cube bonds by the aromatic pyridine linker (Figure 4). Such a structural change provides two unsaturated zinc atoms for reactant coordination, at a distance defined by the aromatic pyridine linkers, whose lengths precisely match the requisite dimensions for a 'floor-to-floor' substrate binding! The resulting bound complex is poised to produce the homochiral product alkoxide due to the constrained geometry of the tetramer, which favors alkyl delivery from only one diisopropylzinc-bound arm. This model also elucidates how inclusion of the alkynyl substituent naturally improves selectivity by further disfavoring the competing pathway through an unfavorable steric interaction. Likewise, benzaldehydes, incapable of such a two-point binding are incompetent
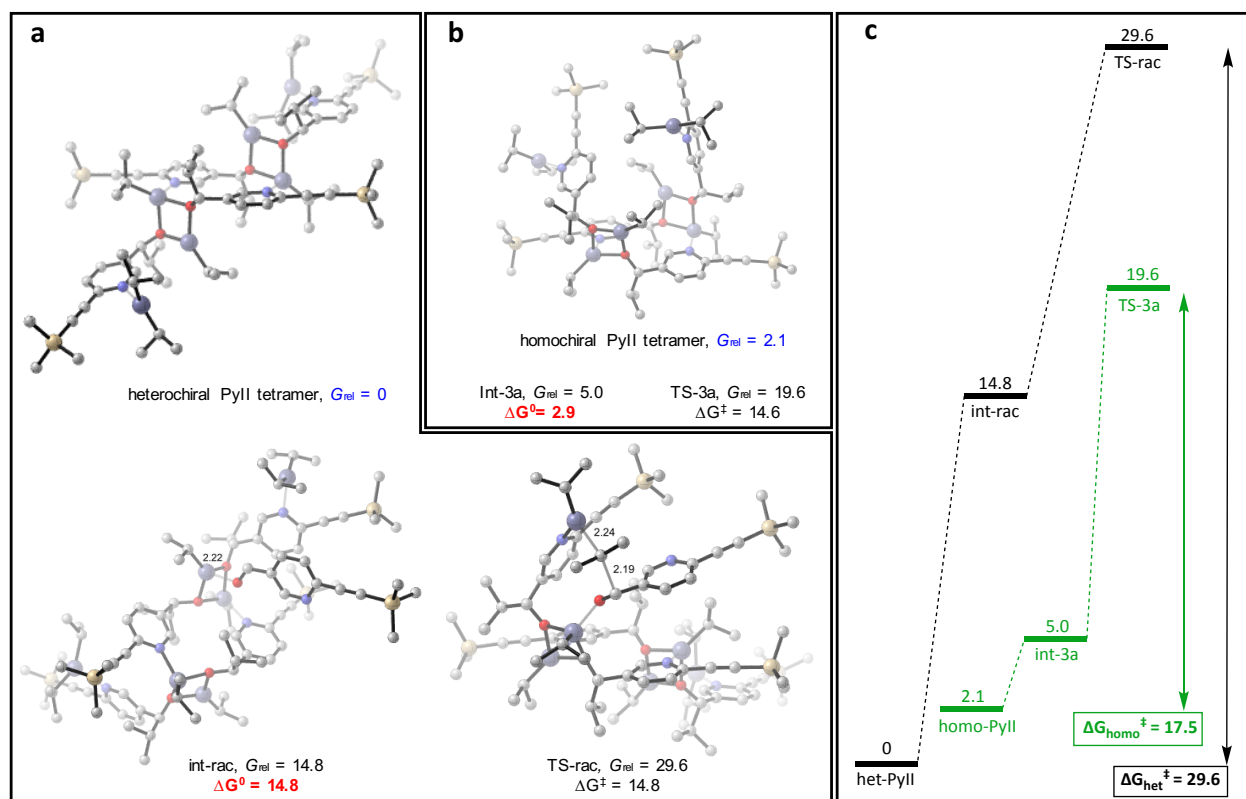
Figure 7. DFT studies for comparison of catalysis by the heterochiral (panel **a** and black graphic in panel **c**) and homochiral (panel **b** and green graphic in panel **c**) **PyII** tetramers calculated at the M06-2X/def2-TZVPP–SMD (toluene)//B3LYP/6-31G(d) level of theory. Panel **c** summarizes this comparison in the form of a relative energy diagram. Energies reported in kcal/mol. Purple, zinc; yellow, silicon; red, oxygen; blue, nitrogen; gray, carbon. Hydrogens are hidden for clarity. ΔG^0, free energy of binding. $\Delta G^{\ddagger}$, free energy activation barrier.

substrates. The geometry of the energetically preferred heterochiral racemic tetramer precludes such a *floor-to-floor* binding, thus providing a simple resolution for the origin of non-linearity. In conclusion, these results contribute to addressing longstanding mechanistic questions regarding the extraordinary Soai reaction, provide an opportunity to further broaden the reaction scope to other substituted pyridine-3-carbaldehydes, and will serve as a platform for further studies and explorations in the fascinating area of asymmetric autocatalysis.

Acknowledgments

Author contributions

S.V.A conceptualized the project, designed and performed chemistry experiments, analyzed data and wrote the manuscript. S.E.D. secured funding, supervised, analyzed data and revised the manuscript. A.S. performed DFT calculations under the supervision of K.N.H. All authors contributed in assembling the final draft of the manuscript.

Supplementary Information
Materials and Correspondence
Sections 1-5
Figures S1-1 to S5-8
Tables S2-1 to S5-29

References

1   Soai, K., Kawasaki, T. & Matsumoto, A. Asymmetric Autocatalysis of Pyrimidyl Alkanol and Its Application to the Study on the Origin of Homochirality. *Acc. Chem. Res.* **47**, 3643-3654, doi:10.1021/ar5003208 (2014).

2   Soai, K., Shibata, T., Morioka, H. & Choji, K. Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. *Nature* **378**, 767-768 (1995).

3   Sato, I., Urabe, H., Ishiguro, S., Shibata, T. & Soai, K. Amplification of Chirality from Extremely Low to Greater than 99.5 % ee by Asymmetric Autocatalysis. *Angewandte Chemie International Edition* **42**, 315-317, doi:10.1002/anie.200390105 (2003).

4   Soai, K. *et al.* Asymmetric synthesis of pyrimidyl alkanol without adding chiral substances by the addition of diisopropylzinc to pyrimidine-5-carbaldehyde in conjunction with asymmetric autocatalysis. *Tetrahedron: Asymmetry* **14**, 185-188, doi:http://dx.doi.org/10.1016/S0957-4166(02)00791-7 (2003).

5   Singleton, D. A. & Vo, L. K. Enantioselective Synthesis without Discrete Optically Active Additives. *J. Am. Chem. Soc.* **124**, 10010-10011, doi:10.1021/ja027129o (2002).

6   Singleton, D. A. & Vo, L. K. A Few Molecules Can Control the Enantiomeric Outcome. Evidence Supporting Absolute Asymmetric Synthesis Using the Soai Asymmetric Autocatalysis. *Org. Lett.* **5**, 4337-4339, doi:10.1021/ol035605p (2003).

7   Soai K., Shibata T., Kowata Y.: Japan Kokai Tokkyo Koho JP 1997 9-268179. Application date: February 1 and April 18, 1996.

8   Mislow, K. Absolute Asymmetric Synthesis: A Commentary. *Collect. Czech. Chem. Commun.* **68**, 849-864 (2003).

9   Sato, I. *et al.* Highly enantioselective synthesis of organic compound using right- and left-handed helical silica. *Tetrahedron Lett.* **44**, 721-724, doi:http://dx.doi.org/10.1016/S0040-4039(02)02707-7 (2003).

10   Lutz, F., Sato, I. & Soai, K. The Asymmetric Power of Chiral Ligands Determined by Competitive Asymmetric Autocatalysis. *Org. Lett.* **6**, 1613-1616, doi:10.1021/ol049559k (2004).

11   Kawasaki, T. *et al.* Highly enantioselective asymmetric autocatalysis using chiral ruthenium complex-ion-exchanged synthetic hectorite as a chiral initiator. *Organic & Biomolecular Chemistry* **7**, 1073-1075, doi:10.1039/B823282B (2009).

12   Sato, I. *et al.* Determination of absolute configurations of amino acids by asymmetric autocatalysis of 2-alkynylpyrimidyl alkanol as a chiral sensor. *J. Organomet. Chem.* **692**, 1783-1787, doi:http://dx.doi.org/10.1016/j.jorganchem.2006.11.042 (2007).

13   Kawasaki, T. *et al.* Asymmetric Amplification Using Chiral Cocrystals Formed from Achiral Organic Molecules by Asymmetric Autocatalysis. *Angew. Chem. Int. Ed.* **44**, 2774-2777, doi:10.1002/anie.200462963 (2005).

14   Sato, I. *et al.* Asymmetric Induction by Helical Hydrocarbons: [6]- and [5]Helicenes. *Angew. Chem. Int. Ed.* **40**, 1096-1098, doi:10.1002/1521-3773(20010316)40:6<1096::AID-ANIE10960>3.0.CO;2-K (2001).

15   Kawasaki, T. *et al.* Chiral Discrimination of Cryptochiral Saturated Quaternary and Tertiary Hydrocarbons by Asymmetric Autocatalysis. *J. Am. Chem. Soc.* **128**, 6032-6033, doi:10.1021/ja061429e (2006).

16   Kawasaki, T. *et al.* Enantioselective Synthesis of Near Enantiopure Compound by Asymmetric Autocatalysis Triggered by Asymmetric Photolysis with Circularly Polarized Light. *J. Am. Chem. Soc.* **127**, 3274-3275, doi:10.1021/ja0422108 (2005).

17   Kawasaki, T. *et al.* Autocatalysis Triggered by Carbon Isotope (13C/12C) Chirality. *Science* **324**, 492-495, doi:10.1126/science.1170322 (2009).

18   Kawasaki, T. *et al.* Asymmetric Autocatalysis: Triggered by Chiral Isotopomer Arising from Oxygen Isotope Substitution. *Angew. Chem. Int. Ed.* **50**, 8131-8133, doi:10.1002/anie.201102263 (2011).

19   Matsumoto, A. *et al.* Asymmetric Induction by a Nitrogen 14N/15N Isotopomer in Conjunction with Asymmetric Autocatalysis. *Angew. Chem. Int. Ed.* **55**, 15246-15249, doi:doi:10.1002/anie.201608955 (2016).

20   Blackmond, D. G. The Origin of Biological Homochirality. *Cold Spring Harbor Perspectives in Biology* **2**, a002147, doi:10.1101/cshperspect.a002147 (2010).

21   Flügel, R. M. *Chirality and Life: A Short Introduction to the Early Phases of Chemical Evolution.* 1 edn, (Springer-Verlag Berlin Heidelberg, 2011).

22   Weissbuch, I. & Lahav, M. Crystalline Architectures as Templates of Relevance to the Origins of Homochirality. *Chem. Rev.* **111**, 3236-3267, doi:10.1021/cr1002479 (2011).

23   Cintas, P. in *Top. Curr. Chem.*   (Springer-Verlag Berlin Heidelberg, 2013).

24   Blackmond, D. G., McMillan, C. R., Ramdeehul, S., Schorm, A. & Brown, J. M. Origins of Asymmetric Amplification in Autocatalytic Alkylzinc Additions. *J. Am. Chem. Soc.* **123**, 10103-10104, doi:10.1021/ja0165133 (2001).

25   Ercolani, G. & Schiaffino, L. Putting the Mechanism of the Soai Reaction to the Test: DFT Study of the Role of Aldehyde and Dialkylzinc Structure. *The Journal of Organic Chemistry* **76**, 2619-2626, doi:10.1021/jo102525t (2011).

26   Sato, I. *et al.* Relationship between the time, yield, and enantiomeric excess of asymmetric autocatalysis of chiral 2-alkynyl-5-pyrimidyl alkanol with amplification of enantiomeric excess. *Tetrahedron: Asymmetry* **14**, 975-979, doi:http://dx.doi.org/10.1016/S0957-4166(03)00164-2 (2003).

27   Islas, J. R. *et al.* Mirror-symmetry breaking in the Soai reaction: A kinetic understanding. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 13743-13748, doi:10.1073/pnas.0503171102 (2005).

28   Buhse, T., Noble-Terán, M. E., Cruz, J.-M., Micheau, J.-C. & Coudret, C. in *Advances in Asymmetric Autocatalysis and Related Topics*   (eds Gyula Pályi, Róbert Kurdi, & Claudia Zucchi) 71-110 (Academic Press, 2017).

29   Gridnev, I. D., Serafimov, J. M. & Brown, J. M. Solution Structure and Reagent Binding of the Zinc Alkoxide Catalyst in the Soai Asymmetric Autocatalytic Reaction. *Angew. Chem. Int. Ed.* **43**, 4884-4887, doi:10.1002/anie.200353572 (2004).

30   Klankermayer, J., Gridnev, I. D. & Brown, J. M. Role of the isopropyl group in asymmetric autocatalytic zinc alkylations. *Chem. Commun.*, 3151-3153, doi:10.1039/B705978G (2007).

31   Quaranta, M., Gehring, T., Odell, B., Brown, J. M. & Blackmond, D. G. Unusual Inverse Temperature Dependence on Reaction Rate in the Asymmetric Autocatalytic Alkylation of Pyrimidyl Aldehydes. *J. Am. Chem. Soc.* **132**, 15104-15107, doi:10.1021/ja103204w (2010).

32   Gridnev, I. D. & Vorobiev, A. K. Quantification of Sophisticated Equilibria in the Reaction Pool and Amplifying Catalytic Cycle of the Soai Reaction. *ACS Catalysis* **2**, 2137-2149, doi:10.1021/cs300497h (2012).

33   Matsumoto, A. *et al.* Crystal Structure of the Isopropylzinc Alkoxide of Pyrimidyl Alkanol: Mechanistic Insights for Asymmetric Autocatalysis with Amplification of Enantiomeric Excess. *Angew. Chem. Int. Ed.* **54**, 15218-15221, doi:10.1002/anie.201508036 (2015).

34   Gehring, T., Busch, M., Schlageter, M. & Weingand, D. A concise summary of experimental facts about the Soai reaction. *Chirality* **22**, E173-E182, doi:10.1002/chir.20849 (2010).

35   Shibata, T., Yonekubo, S. & Soai, K. Practically Perfect Asymmetric Autocatalysis with (2-Alkynyl-5-pyrimidyl)alkanols. *Angew. Chem. Int. Ed.* **38**, 659-661, doi:10.1002/(SICI)1521-3773(19990301)38:5<659::AID-ANIE659>3.0.CO;2-P (1999).

36   Tanji, S. *et al.* Asymmetric autocatalysis of 5-carbamoyl-3-pyridyl alkanols with amplification of enantiomeric excess. *Tetrahedron: Asymmetry* **11**, 4249-4253, doi:http://dx.doi.org/10.1016/S0957-4166(00)00420-1 (2000).

37   Shibata, T., Choji, K., Hayase, T., Aizu, Y. & Soai, K. Asymmetric autocatalytic reaction of 3-quinolylalkanol with amplification of enantiomeric excess. *Chem. Commun.*, 1235-1236, doi:10.1039/CC9960001235 (1996).

38   Soai, K., Niwa, S. & Hori, H. Asymmetric self-catalytic reaction. Self-production of chiral 1-(3-pyridyl)alkanols as chiral self-catalysts in the enantioselective addition of dialkylzinc reagents to pyridine-3-carbaldehyde. *J. Chem. Soc., Chem. Commun.*, 982-983, doi:10.1039/C39900000982 (1990).

39   These results will be disclosed in a full account of this work.

40   Till now, the pyrimidine core was implicitly assumed to be a necessity for efficient amplifying autocatalysis (see Figure 1c). Hence this result with the analogous pyridine system was unexpected.

41   Romagnoli, C., Sieng, B. & Amedjkouh, M. Asymmetric Amplification Coupling Enantioselective Autocatalysis and Asymmetric Induction for Alkylation of Azaaryl Aldehydes. *Eur. J. Org. Chem.* **2015**, 4087-4092, doi:10.1002/ejoc.201500508 (2015).

42   A catalyzed diisopropylzinc addition to 3 was reported in ref. 41. This study claimed that the reaction of diisopropylzinc and 5-(t-butylylethynyl)pyridine-3-carbaldehyde showed chiral erosion. However, under our reaction conditions, we see amplifying autocatalysis with this substrate as well.

43   Noltes, J. G. & Boersma, J. Investigations on organozinc compounds IX. Coordination chemistry of organozinc compounds RZnX: organozinc derivatives of tert-butanol, some phenols, diethylhydroxylamine and some oximes. *J. Organomet. Chem.* **12**, 425-431, doi:http://dx.doi.org/10.1016/S0022-328X(00)88695-7 (1968).

44   Jana, S., Berger, R. J. F., Fröhlich, R., Pape, T. & Mitzel, N. W. Oxygenation of Simple Zinc Alkyls: Surprising Dependence of Product Distributions on the Alkyl Substituents and the Presence of Water. *Inorg. Chem.* **46**, 4293-4297, doi:10.1021/ic062438r (2007).

45   Kitamura, M., Okada, S., Suga, S. & Noyori, R. Enantioselective addition of dialkylzincs to aldehydes promoted by chiral amino alcohols. Mechanism and nonlinear effect. *J. Am. Chem. Soc.* **111**, 4028-4036, doi:10.1021/ja00193a040 (1989).