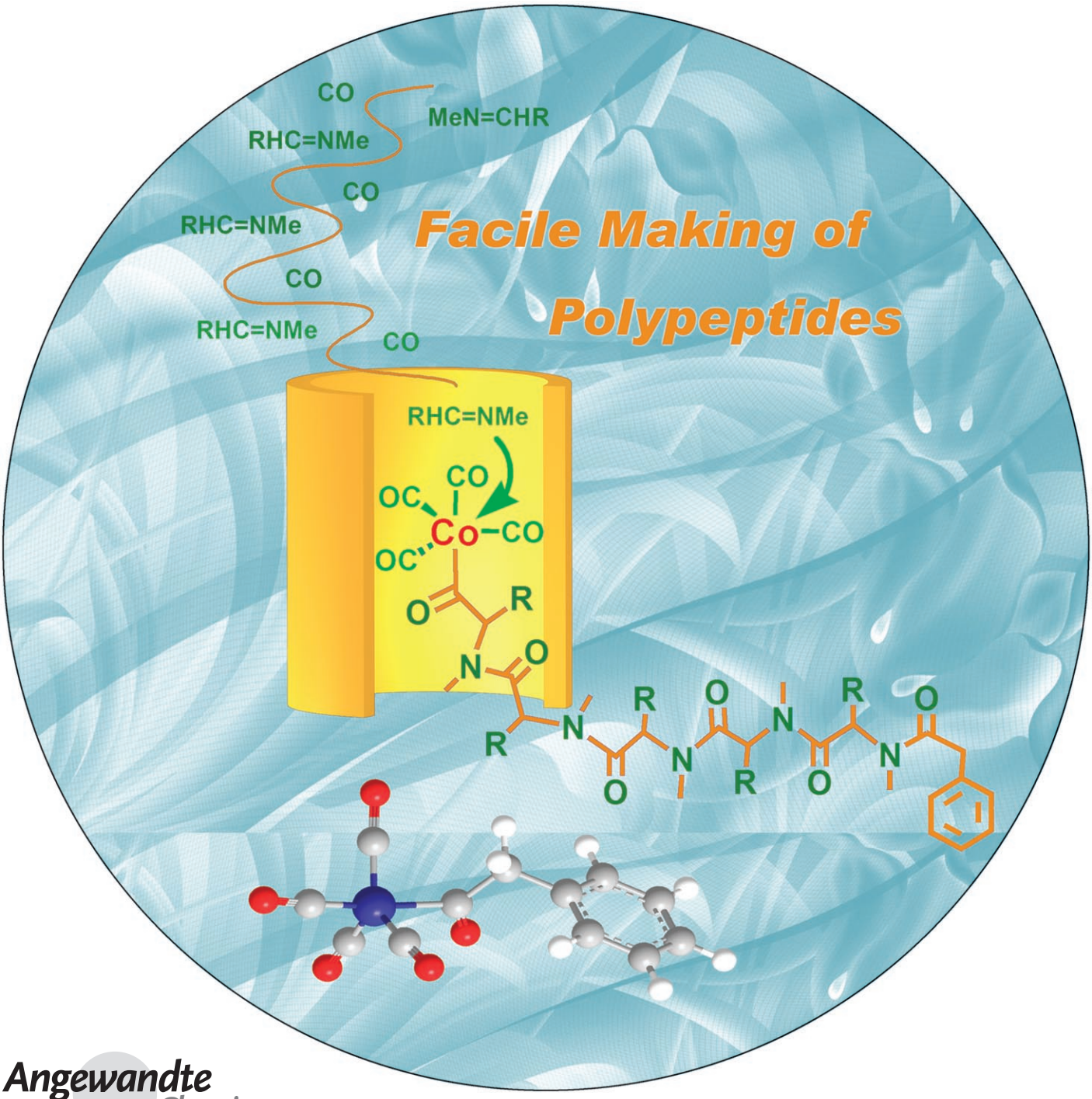
Polypeptide Synthesis

DOI: 10.1002/anie.200700646

Metal-Catalyzed Copolymerization of Imines and CO: A Non-Amino Acid Route to Polypeptides**

Huailin Sun, Jian Zhang, Qiuhua Liu, Lei Yu, and Jiangyu Zhao*

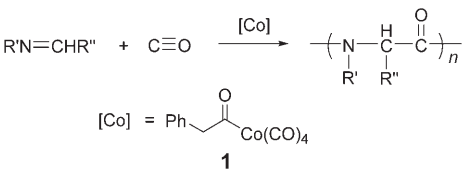


Polypeptides are an exceptionally significant class of biopolymers that not only are responsible for both the structure and function of most living things but also have broad applications in materials, catalysis, and pharmaceuticals. For a century, almost all studies on polypeptide synthesis have been based on the use of amino acids as starting materials, which requires tedious procedures for presynthesis of amino acids and subsequent activation of the highly stable carboxyl groups using stoichiometric amounts of special reagents to form peptide bonds.[1–3] Herein, we report a shortcut method that does not use amino acids but instead employs readily available imines and carbon monoxide (CO) as monomers that undergo metal-catalyzed alternating copolymerization to directly form polypeptides. We find that a simple acylcobalt complex can effectively catalyze this reaction to produce polypeptides of high molecular weights with low polydispersity. This efficient metal-catalyzed synthesis of polypeptides from inexpensive and plentiful starting materials makes large-scale production of the polypeptide materials possible in a fashion similar to Ziegler–Natta polymerization reactions.

Construction of polymers from small organic monomers represents an important transformation. Beginning with Ziegler and Natta's discoveries, metal catalysis has been recognized as an extremely efficient tool for such purposes. Recent breakthroughs achieved in homogeneous catalysis using well-defined metal catalysts have resulted in rapid progress in the study of various types of polymerization reactions, such as the insertion and the metathesis polymerization of alkenes and the copolymerization of olefins and CO.[4] Most of these reactions, however, involve the use of alkenes as monomers. In sharp contrast, similar reactions of imines are largely unexplored, even though imines constitute a huge class of organic compounds readily available from various aldehydes (or ketones) and appropriate amines. This phenomenon is usually attributed to the structural character of imines, that is, the lone pair of electrons on nitrogen that favors σ coordination to metals and hence prohibits the π complexation required for activation of the carbon–nitrogen double bond. Of particular note is the copolymerization of imines and CO, which has long been suggested to be a potentially attractive route to polypeptides but is difficult to realize owing to lack of appropriate catalysts.[5–8]

In 1998, Sen and co-workers[5] and Arndsten and co-workers[6] independently reported the first observation of imine insertion into acyl carbon–palladium bonds. This achievement constitutes a critical step toward metal-catalyzed copolymerization of imines and CO. However, desired subsequent insertion of CO into the resulting carbon–metal bond failed to occur. The adjacent amide carbonyl strongly coordinates to the metal center to form a stable five-membered metallacycle that prevents coordination of incoming CO. Tuning the ligands on palladium could partly solve the problem, resulting in CO insertion to produce a single amino acid unit.[7,9] However, facile elimination of metal moieties from the organic scaffold took place. As a result, even a simple dipeptide has never been obtained. Other metals, such as nickel[10] and manganese,[11] have also been tested, but none exhibited desired catalytic activity. Therefore, a catalyst that allows continuous insertion of imines and CO is still unknown.

We address this issue by employing cobalt rather than palladium or other previously used metals. Cobalt is chosen because it is one of the most frequently used catalysts for carbonylation reactions,[4] and it has recently been successfully applied to catalyze copolymerization of aziridines and CO.[12] It is now well-established that the active species in such catalysis is acylcobalt, the chemistry of which has been thoroughly studied.[13] Evidence also indicates that cobalt is less prone to coordinate to an adjacent carbonyl group, which will be critically important for the present reaction. Although a phosphine-substituted acylcobalt complex has failed to catalyze this reaction,[8] we find now that the simple acylcobalt complex **1** without a phosphine ligand can effectively catalyze the copolymerization of imines and CO under suitable conditions, giving rise to the desired polypeptides as shown in Equation (1).



1 (1)

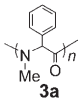
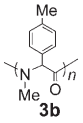
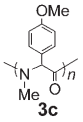
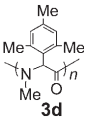
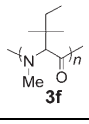
Catalyst **1** was synthesized according to the reported procedures.[12,13] Its ability to catalyze the copolymerization was first probed with the stable imine $MeN{=}CHC_6H_5$ (**2a**), which is available either commercially or from condensation of benzaldehyde and methylamine. The polymerization was performed in dioxane under 800 psi of CO pressure at 50 °C. After removing the solvent and washing with hexane, the corresponding polypeptide **3a** was obtained as a solid product (Table 1, entries 1 and 2). Analysis by ^{1}H NMR spectroscopy showed the presence of phenyl, *N*-methyl, and methine protons at the expected positions,[14–16] which were confirmed by ^{13}C and HSQC NMR spectroscopy. The methine proton exhibited a rather high chemical shift at about 6.50 ppm, which is in accord with previous observations.[15] The amide carbonyl that did not appear in the ^{1}H NMR spectrum was observable in ^{13}C NMR and IR spectra. The polypeptide structure was finally established by MALDI-TOF MS analysis. The mass spectrum exhibited a series of peaks with the same interval between them, which was exactly equal to the

[*] Prof. H. Sun, J. Zhang, Q. Liu, L. Yu, J. Zhao
Department of Chemistry, Nankai University
94 Weijin Road, Tianjin 300071 (China)
Fax: (+86) 22-2349-8132
E-mail: sunhl@nankai.edu.cn

[**] We thank Prof. J.-P. Cheng, Prof. J.-G. Tian, Prof. Q.-C. Wang, and Prof. Y.-X. Che for their help. The authors also thank Dr. D. J. Siegward and Prof. K. Kavallieratos for reviewing the manuscript, and Prof. X.-Q. Zhu for valuable discussion. This work was supported by NSFC (nos. 20372036 and 20421202), the Natural Science Foundation of Tianjin (no. 05YFJMJC06800) and the Ministry of Education of China (no. 03406).

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

Table 1: Cobalt-catalyzed copolymerization of imines and CO to produce polypeptides.

Entry	Imines[a]	M/C[b]	*t* [h]	Polypeptides	Yield [%]	$M_n \times 10^{-3}$[c]	PDI[d]
1	**2a**	11:1	6		73	2.4	1.16
2	**2a**	22:1	12		38	2.4	1.16
3	**2a**	11:1[e]	6		45	1.9	[g]
4	**2a**	11:1[f]	6	**3a**	48	1.7	[g]
5	**2b**	11:1	6		100	3.2	1.23
6	**2b**	22:1	12	**3b**	54	3.6	1.39
7	**2c**	11:1	12		100	4.1	1.28
8	**2c**	22:1	12	**3c**	59	4.1	1.22
9	**2d**	11:1	12		75	1.9	[g]
10	**2d**	22:1	12	**3d**	31	1.8	[g]
11	**2e**	22:1	12		100	8.2	1.29
12	**2e**	44:1	12		100	11.8	1.22
13	**2e**	66:1	12		100	18.5	1.18
14	**2e**	100:1	24		100	28.8	1.18
15	**2e**	100:1	12	**3e**	87	22.1	1.20
16	**2e**	100:1	6		36	11.0	1.27
17	**2f**	22:1	12		100	4.3	1.18
18	**2f**	44:1	24		100	8.0	1.15
19	**2f**	80:1	36		100	13.7	1.14
20	**2f**	100:1	48	**3f**	100	19.4	1.17

[a] Reactions were performed in dioxane at 50 °C under 800 psi of CO pressure. [b] Monomer-to-catalyst molecular ratio. [c] Number- and weight-average molecular weights (M_n and M_w, respectively) were determined by GPC relative to polystyrene standards in THF. [d] Polydispersity index, M_n/M_w. [e] Reaction in DME. [f] Reaction in benzene. [g] Not determined due to closeness to lower molecular-weight limits.

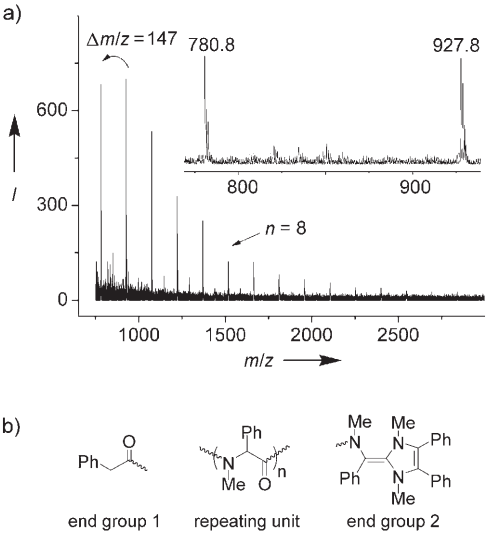


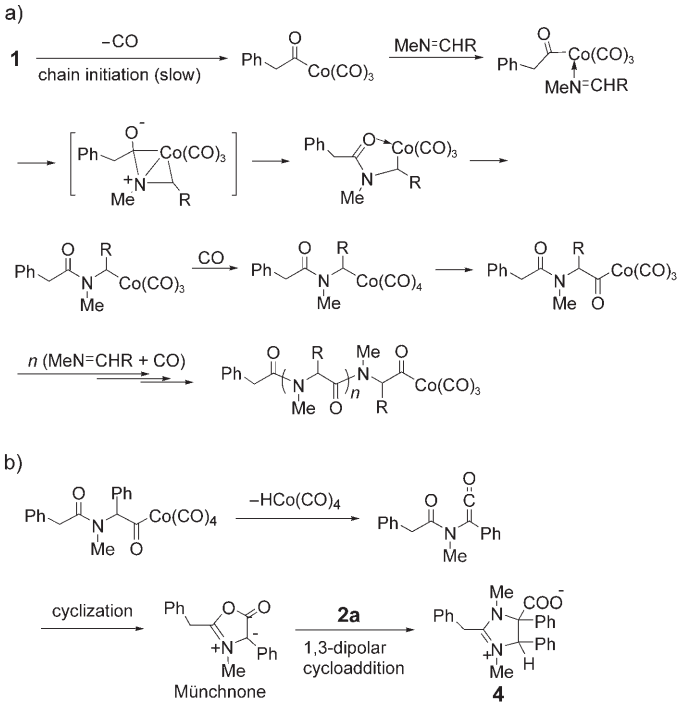
Figure 1. MALDI-TOF MS analysis of polypeptides. a) The spectrum of **3a** (Table 1, entry 1) obtained in the reflection mode with 2,5-dihydroxybenzoic acid (DHB) as the matrix. The inset is an expansion of the *m*/*z* region from 770 to 940. b) Tentative assignment of the end group structures.

mass of the repeating unit (Figure 1a). Hydrolysis of the polypeptide quantitatively produced *N*-methyl phenylglycine, which further confirms the structure of the polypeptide.[16]

The successful characterization of the polypeptide demonstrates that repeated insertion of imines and CO has occurred. Gel permeation chromatography (GPC) analysis indicated a number-average molecular weight (M_n) of about 2000 dalton. The reaction conditions were optimized. It was found that varying CO pressures from 600 to 1000 psi did not significantly change the yield of the polypeptide. Raising the temperature to 80 °C lead to reduced yields presumably owing to catalyst decomposition, and reaction at room temperature also resulted in decreased yields probably caused by chain initiation problems. Other solvents, such as dimethoxyethane (DME) and benzene, were unable to improve the reaction (Table 1, entries 3 and 4). It seems that less-polar solvents were not favored by the copolymerization.

These results indicated that rapid chain termination is a serious problem preventing formation of high-molecular-weight polypeptides. Since it was initially thought that the reaction might proceed via formation of *N*-acyliminium intermediates from reaction of acylcobalt with imines,[8] chain termination might occur through the well-known nucleophilic addition of adjacent phenyl groups to the cationic carbon atom of *N*-acyliminium.[17] However, the fact that the copolymerization could proceed in benzene suggests that this explanation is not the correct one. A plausible pathway that does not involve *N*-acyliminium intermediates is the so-called coordination mechanism, in which an imine is first coordinated to cobalt and then undergoes concerted intramolecular insertion into an acyl carbon–cobalt bond via a four-centered transient state, resulting in a five-membered metallacycle. A similar mechanism has been suggested by theoretical calculations for imine insertion into acyl carbon–palladium bonds.[18] However, the five-membered cobalt metallacycle must be less stable than that of palladium, just as expected, and hence allow CO coordination and insertion to produce the next generation of acylcobalt intermediates. These intermediates have a vacant coordination site at cobalt, which would be easily coordinated by imines to begin the next round of chain propagation (Scheme 1a).

The chain termination mechanism has been examined by means of end-group analysis. ^{1}H and ^{13}C NMR spectra showed the presence of the phenylacetyl end group inherited from the catalyst, but the second end groups were not detected. According to MALDI-TOF MS analysis, they were certainly not carboxyl groups derived from hydrolysis of the acylcobalt species[12] but were most likely consistent with the imidazoline derivative shown in Figure 1b. Attempts to isolate such end groups through hydrolysis of the polypeptides were not successful, but a small-molecule imidazoline **4** was obtained and completely characterized.[16] This compound is obviously a product of 1,3-dipolar cycloaddition of



Scheme 1. Possible mechanism of cobalt-catalyzed alternating copolymerization of imines and CO. a) Proposed mechanism for chain initiation by CO dissociation from catalyst **1**, which is a well-established slow process occurring on a time scale of hours,[12,13] and of chain propagation through imine coordination to cobalt followed by intramolecular migratory insertion. b) Proposed chain-termination pathway, as exemplified by the formation of imidazoline **4**. Further transformation of imidazolines to downstream derivatives might occur under appropriate conditions.

imines and mesoionic Münchnones.[7] The latter should have been formed through β-hydrogen elimination of the acylcobalt intermediate and subsequent cyclization of the resulting ketene (Scheme 1b). Similar β-hydrogen elimination to form ketenes has been found to occur extensively in palladium systems.[7,9] In the case of cobalt, however, this was the first time such a reaction has been observed. Nevertheless, the reverse process (addition of hydridocobalt to ketenes) has been reported.[19] Although further work is required for complete clarification of the chemistry involved here, it seems to be of little doubt that the second end group should be derived from the imidazolines after loss of carbon dioxide and hydrogen as a result of oxidation upon exposure to air.[16]

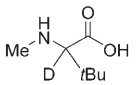
As β-hydrogen elimination is the key step of chain termination, the high acidity of the methine hydrogen atom, revealed by the large chemical shift of the proton, must be responsible for the facile chain termination during polymerization. Thus, when a methyl group was introduced onto the phenyl ring in the imine MeN=CHC_6H_4-*p*-Me (**2b**), a sudden increase in the yield and the molecular weight of polypeptide **3b** was observed (Table 1, entries 5 and 6). This improvement is obviously a result of the electron-donating effect of the methyl group, which retards chain termination by lowering the acidity of the hydrogen atom. Further increase of the electron-releasing ability using the *p*-methoxy group in imine MeN=CHC_6H_4-*p*-OMe (**2c**) resulted in polypeptide **3c** with $M_n > 4000$ dalton (Table 1, entries 7 and 8). Interestingly, even the very crowded mesityl imine MeN=CHC_6H_2-2,4,6-Me_3 (**2d**) provided polypeptide **3d**, albeit in lower molecular weights and yields (Table 1, entries 9 and 10).

High-molecular-weight polypeptides were finally obtained using alkyl in place of aryl imines. Owing to the low sensitivity of the reaction with respect to steric factors of the substituent, pivalaldehyde imine MeN=CH-*t*Bu (**2e**), which is very stable during prolonged storage, was used. The reaction was performed under the same conditions as for aryl imines. After removing the solvent, the desired polypeptide **3e** was obtained as a solid product in quantitative yield. Surprisingly, the quantitative conversion was maintained even at high monomer-to-catalyst (M/C) molecular ratios (Table 1, entries 11–14). The products were characterized by various spectroscopic methods.[16] Analysis by ^{1}H NMR spectroscopy showed that the methine protons had much smaller chemical shifts (less than 5.5 ppm) than in the aryl-substituent cases, indicative of dramatic decrease of the acidity of the hydrogen atom. This decrease must be the reason why chain termination has been largely suppressed. MALDI-TOF MS of the low-molecular-weight samples gave a series of polypeptide molecular-ion peaks (see Figure S35 in the Supporting Information) whose exact masses were consistent with Münchnone end groups rather than the imidazoline derivatives observed for the aryl polypeptides. This finding could be rationalized by the fact that the chain termination might occur after complete consumption of the imine, thus making 1,3-dipolar addition of the imine to Münchnone impossible.

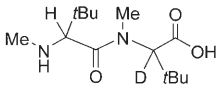
GPC analysis revealed a low polydispersity index (PDI) for the polypeptides. The molecular weights of the polypeptides increased roughly linearly with the increase of M/C molecular ratios (see Figure S44 in the Supporting Information); meanwhile, the PDIs improved gradually. Similar changes of the PDIs were also observed during copolymerization. This result is consistent with the slow chain initiation (Scheme 1a), which is incompatible with the relatively fast chain propagation. Replacing one of the methyl groups of the *tert*-butyl substituent in **2e** with an ethyl group in imine MeN=$CHC(Me)_2Et$ (**2f**) reduced the rate of polymerization, probably for steric reasons, while the quantitative transformation to polypeptide **3f** was also observed (Table 1, entries 17–19). Linear change of the molecular weights with the M/C ratios was demonstrated again (see Figure S45 in the Supporting Information). The PDIs of the polymers were obviously lower than before the introduction of the ethyl group. This finding may be explained by the lowered chain propagation rate that becomes more compatible with the slow chain initiation. Noting the behaviors of the copolymerization, including quantitative conversion, low PDIs, and linear increase of molecular weight, the tertiary alkyl imine systems seemingly meet most of the criteria for a living polymerization process, although attempts to synthesize block copolymers by adding a second imine monomer were not successful, probably because of the instability of the active catalytic species after consumption of the monomers.[20,21]

The success in modifying the tertiary alkyl group leaves room for further introduction of functional groups. Interestingly, the N- and C-substituted polypeptides obtained above

cannot, for steric reasons, be accessed by other means, including the most frequently used ring-opening polymerization of amino acid *N*-carboxyanhydrides.[22–25] These polypeptides, unlike the N-unsubstituted ones, are soluble in common organic solvents, such as THF and chloroform. One of the unique properties of these polymers is that they can be facilely degraded by trifluoroacetic acid (TFA).[16] Although TFA cleavage of small peptides has been reported recently,[26,27] polypeptide degradation by TFA has never been observed. The ultimate products were found to be a mixture of amino acid **5** and dipeptide **6**, for the case of the *tert*-butyl polypeptide **3e**, with regiospecific deuteration on the α-carbons when the degradation was performed in $[D_1]$TFA. This result is consistent with the mechanism involving cleavage of peptide bonds by adjacent carbonyl oxygen atoms to form Münchnone intermediates (see Scheme S3 in the Supporting Information).[28] Such unique degradation properties would be useful for applications of special functional materials.



5



6

We have shown that the highly desirable copolymerization of imines and CO has been realized through proper choice of a simple cobalt catalyst, demonstrating once again the ability of metal catalysis to manipulate organic transformations.[29] Of particular note are the ready availability and low cost of the starting materials as well as the atom-economic feature of the reaction process (see Scheme S1 in the Supporting Information), which render the reaction well-suited for large-scale production of the polypeptide materials. In fact, this method is the shortest possible route for chemical (or abiotic) synthesis of polypeptides. Whether such a copolymerization strategy might have been adopted by nature for the prebiotic origin of polypeptides might deserve further investigation.[30] Although the present study has been confined to the use of stable imines, their successful copolymerization with CO has indicated that this transformation can be both thermodynamically and kinetically viable. Further expandsion of the scope of this reaction and full elucidation of the reaction mechanism are the subject of current research.

Experimental Section

Imine (about 1 mL) was added to a Par pressure reactor containing a suitable amount of catalyst **1** in dioxane (50 mL), which had been pressurized under 800 psi of CO overnight. After the addition, the pressure of CO was returned to 800 psi, and the reactor was heated in an oil bath at 50 °C while the reaction mixture was magnetically stirred for the period of time specified in Table 1. After cooling to room temperature, the pressure was released and the reactor opened. The resulting solution was transferred to a flask, and the solvent was removed under vacuum to afford the crude product polypeptide.

Received: February 12, 2007
Revised: April 16, 2007
Published online: July 3, 2007

Keywords: carbon monoxide · copolymerization · homogeneous catalysis · imines · polypeptides

[1] N. Sewald, H.-D. Jakubke, *Peptides: Chemistry and Biology*, Wiley-VCH, Weinheim, **2002**.
[2] M. Bodanszky, A. Bodanszky, *Principles of Peptide Synthesis, 2nd ed.*, Springer, New York, **1994**.
[3] T. J. Deming, *Adv. Mater.* **1997**, *9*, 299 – 311.
[4] *Applied Homogeneous Catalysis with Organometallic Compounds, 2nd ed.* (Eds.: B. Cornils, W. A. Herrmann), Wiley-VCH, Weinheim, **2002**.
[5] S. Kacker, J. S. Kim, A. Sen, *Angew. Chem.* **1998**, *110*, 1335 – 1337; *Angew. Chem. Int. Ed.* **1998**, *37*, 1251 – 1253.
[6] R. D. Dghaym, K. J. Yaccato, B. A. Arndtsen, *Organometallics* **1998**, *17*, 4 – 6.
[7] R. D. Dghaym, R. Dhawan, B. A. Arndtsen, *Angew. Chem.* **2001**, *113*, 3328 – 3330; *Angew. Chem. Int. Ed.* **2001**, *40*, 3228 – 3230.
[8] J. K. Funk, H. Yennawar, A. Sen, *Helv. Chim. Acta* **2006**, *89*, 1687 – 1695.
[9] R. Dhawan, R. D. Dghaym, B. A. Arndtsen, *J. Am. Chem. Soc.* **2003**, *125*, 1474 – 1475.
[10] J. L. Davis, B. A. Arndtsen, *Organometallics* **2000**, *19*, 4657 – 4659.
[11] D. Lafrance, J. L. Davis, R. Dhawan, B. A. Arndtsen, *Organometallics* **2001**, *20*, 1128 – 1136.
[12] L. Jia, H. Sun, J. T. Shay, A. M. Allgeier, S. D. Hanton, *J. Am. Chem. Soc.* **2002**, *124*, 7282 – 7283.
[13] V. Galamb, G. Palyi, F. Ungvary, L. Marko, R. Boese, G. Schmid, *J. Am. Chem. Soc.* **1986**, *108*, 3344 – 3351.
[14] M-X. Wang, J. Liu, D-X. Wang, Q-Y. Zheng, *Tetrahedron: Asymmetry* **2005**, *16*, 2409 – 2416.
[15] H. O. Bayer, R. Huisgen, R. Knorr, F. C. Schaefer, *Chem. Ber.* **1970**, *103*, 2581 – 2597.
[16] See the Supporting Information.
[17] W. N. Speckamp, M. J. Moolenaar, *Tetrahedron* **2000**, *56*, 3817 – 3856.
[18] L. Cavallo, *J. Am. Chem. Soc.* **1999**, *121*, 4238 – 4241.
[19] F. Ungvary, *J. Organomet. Chem.* **1986**, *303*, 251 – 255.
[20] G. W. Coates, P. D. Hustad, S. Reinartz, *Angew. Chem.* **2002**, *114*, 2340 – 2361; *Angew. Chem. Int. Ed.* **2002**, *41*, 2236 – 2257.
[21] R. R. Schrock, *Acc. Chem. Res.* **1990**, *23*, 158 – 165.
[22] H. R. Kricheldorf, *α-Aminoacid-N-Carboxyanhydrides and Related Materials*, Springer, New York, **1987**.
[23] A. Cosani, M. Palumbo, M. Terbojevich, E. Peggion, *Macromolecules* **1978**, *11*, 1041 – 1045.
[24] D. G. H. Ballard, C. H. Bamford, *J. Chem. Soc.* **1958**, 355 – 360.
[25] T. J. Deming, *Nature* **1997**, *390*, 386 – 389.
[26] J. Urban, T. Vaisar, R. Shen, M. S. Lee, *Int. J. Pept. Protein Res.* **1996**, *47*, 182 – 189.
[27] M. Goodman, C. J. Creighton, US patent 6,417,397, **2002**.
[28] M. Teixido, F. Albericio, E. Giralt, *J. Pept. Res.* **2005**, *65*, 153 – 166.
[29] T. M. Trnka, R. H. Grubbs, *Acc. Chem. Res.* **2001**, *34*, 18 – 29.
[30] L. Leman, L. Orgel, M. R. Ghadiri, *Science* **2004**, *306*, 283 – 286.

