View Article Online / Journal Homepage / Table of Contents for this issue

Sequence control in polymer synthesis

Nezha Badi and Jean-François Lutz*

Received 25th March 2009
First published as an Advance Article on the web 23rd July 2009
DOI: 10.1039/b806413j

Published on 23 July 2009. Downloaded by Portland State University on 07/06/2014 09:31:55.

The control over comonomer sequences is barely studied in macromolecular science nowadays. This is an astonishing situation, taking into account that sequence-defined polymers such as nucleic acids and proteins are key components of the living world. In fact, fascinating biological machines such as enzymes, transport proteins, cytochromes or sensory receptors would certainly not exist if evolution had not favored chemical pathways for controlling chirality and sequences. Thus, it seems obvious that synthetic polymers with controlled monomer sequences have an enormous role to play in the materials science of the next centuries. The goal of this *tutorial review* is to shed light on this highly important but embryonic field of research. Both biological and synthetic mechanisms for controlling sequences in polymerization processes are critically discussed herein. This state-of-the-art overview may serve as a source of inspiration for the development of new generations of synthetic macromolecules.

Introduction

The decoding of gene and protein sequences is certainly one of the biggest scientific achievements of the 20th century.[1–3] Indeed, the ordered monomer sequences contained in biopolymers such as nucleic acids and polypeptides are in large part responsible for the diversity, complexity and adaptability of living organisms. Thus, macromolecules such as DNA or RNA are not only crucial scientific subjects but also modern cultural icons. For instance, nucleotide sequences have escaped biology laboratories and invaded our daily lives through popular magazines, art galleries, paternity lawsuits, and forensic TV shows.

Research Group Nanotechnology for Life Science, Fraunhofer Institute for Applied Polymer Research, Geiselbergstrasse 69, 14476 Potsdam-Golm, Germany. E-mail: lutz@iap.fhg.de; Fax: +49 331 568 3000; Tel: +49 331 568 1127

In such a context, one could logically expect that the control over monomer sequences in polymerization processes is an established research field. Yet, even if molecular biologists and biochemists have understood for decades the importance of ordered monomer sequences, this topic is far from being central in other chemical disciplines. In particular, and rather surprisingly, this crucial subject is largely ignored in contemporary polymer chemistry. In fact, the most relevant discovery to date in the field of sequence-controlled macromolecules is due to a biochemist. In early 1963, Robert Bruce Merrifield introduced the solid-phase synthesis of peptides, which relies on the step-by-step attachment of protected monomers.[4] This method remains so far the most reliable pathway for synthesizing oligomers with tailored monomer sequences.

Nevertheless, it would be rather unfair to pretend that polymer chemists never showed interest in understanding macromolecular sequences. In fact, in the late 1960's and early

Nezha Badi

Nezha Badi obtained her MSc degree from the Université Paris VI in 2003. In 2006, she received her PhD in chemistry and physico-chemistry of polymers from the University of Evry-Val-D'Essonne, under the guidance of Prof. Philippe Guégan. After working as a postdoctoral researcher in the group of Dr. Jean-Paul Behr at the Université Strasbourg I, she joined in September 2008 the research group of Dr. Jean-François Lutz at the Fraunhofer Institute for Applied Polymer Research.

Jean-François Lutz

Jean-François Lutz received his BSc and MSc degrees from the Université Paris VI and his doctoral degree from the Université Montpellier II. Afterwards, he joined the research group of Prof. Krzysztof Matyjaszewski at Carnegie-Mellon University, where he worked for two years as a postdoctoral fellow. In late 2003, he moved to Potsdam in Germany and became team leader at the Fraunhofer Institute for Applied Polymer research. His current research interests include the synthesis of sequence-ordered polymers, the design of smart biocompatible polymers and the self-organization of macromolecules in aqueous media.

This journal is © The Royal Society of Chemistry 2009

View Article Online

Published on 23 July 2009. Downloaded by Portland State University on 07/06/2014 09:31:55.

1970's, a large amount of research was conducted for characterizing the monomer sequences of synthetic polymers. In particular, the pioneer development of NMR and mass spectrometry of large macromolecules allowed a comprehensive examination of polymer microstructures.[5,6] However, these topics have been slowly abandoned, or, more correctly, have fallen out of fashion. For instance, the number of articles on polymer sequences published in high-impact journals during the last 20 years is probably less than a dozen.[7] Actually, this research aspect has probably been obscured by the recent exceptional progress in controlling polymer topology (*i.e.* the shape of synthetic macromolecules). Thus, current synthetic copolymers still exhibit, for the most part, random monomer sequence distributions.

Yet, synthetic macromolecules with ordered monomer sequences could potentially play an important role in materials science. For instance, within the last few years, it has been demonstrated that sequenced-defined bio-oligomers such as oligonucleotides or oligopeptides can be efficiently used for organizing synthetic materials (*i.e.* guided self-assembly).[8–11] These interesting strategies could certainly be extended to other types of oligomers or polymers with controlled microstructures. In addition, one could think further and speculate that synthetic polymers with ordered sequences would eventually give access to man-made materials as complex and functional as biological assemblies.

However, although plausible, such visions seem far from current reality. First of all, synthetic chemists should reconsider the question of polymer sequences. Important breakthrough in this domain will likely not happen without a robust scientific community. Hence, this research area should be revitalized and, hopefully, become a priority topic in chemistry within the next years. In this context, the objective of the present tutorial is (i) to review existing pathways (*i.e.* biological or synthetic) for controlling polymer sequences and (ii) to highlight future challenges and directions in this research field.

Learning from Nature: synthesis of biopolymers

Natural biopolymers are an inexhaustible source of inspiration for synthetic chemists. Indeed, through billions of years of molecular evolution, Nature selected efficient routes for synthesizing isomolecular polymers with controlled sequences and configurations. Such a high degree of perfection is obtained with the help of efficient biocatalytic reactions. For instance, polymerization processes such as DNA replication, transcription and translation (*i.e.* protein synthesis) are important examples of sequence-controlled polymerizations. Yet, the objective of the present paragraph is certainly not to describe in detail these highly complex biological mechanisms. Comprehensive overviews on these processes can be found in biochemistry textbooks and in a variety of reviews.[12] Thus, our motivation is simply to extract the mechanistic essence of biological polymerizations. For instance, Table 1 divides replication, transcription and translation into discrete polymerization steps. Although significantly simplified, this classification allows a rational comparison of these complex processes.

It is quite obvious that all three biological polymerizations exhibit important mechanistic similarities. First of all, the sequential incorporation of monomer units is, in all cases, regulated by a templated mechanism. This simply means that the chain-growth of a given polymer is assisted by an another macromolecule of either similar (*e.g.* replication), slightly different (*e.g.* transcription) or completely different (*e.g.* translation) chemical nature. Typically, a temporary complex, involving the macromolecular template, the growing polymer chain and the monomer, is formed. As a result, the molecular sequence of the macromolecular template is copied or translated to the newly synthesized chain. Thus, in all biological polymerizations, the sequence information is inherently present in the system. It is somewhat difficult to imagine an alternative mechanism, taking into account that these polymerizations indeed proceed without any guiding intelligence. In contrast, and as described in the next paragraphs, man-made approaches for controlling sequences can be fundamentally different and do not necessarily require a prewritten template.

Nonetheless, it is perhaps too reductive to claim that the mechanisms of biological polymerizations are just templated. In polymer science, a template mechanism simply denotes a polymerization performed in the presence of a complementary macromolecule.[13] In such processes, multiple monomer molecules can be simultaneously associated with the macromolecular template. Yet, in biological polymerizations, this aspect is considerably more controlled. Indeed, biological monomers are usually associated one by one to their template and therefore polymerized in an exact sequential order. This precise monomer selection is in part due to a chain-walking mechanism. In all biological polymerizations, complex biocatalysts progress along the templated chains and allow a sequential molecular recognition of monomer units.

For instance, DNA and RNA polymerases regulate nucleic acid polymerizations in eukaryotes and prokaryotes.[12] These enzymes provide an optimized environment for confining a DNA-template, a growing chain and a nucleoside triphosphate (Fig. 1). However, the synthesis of a new phosphodiester bond only occurs if the template-strand and the monomer form a molecular complex of the Watson–Crick-type. As a consequence of this precise monomer selection, chain-growth is sequence-specific in replication and transcription processes.[12] Moreover, after each monomer incorporation, polymerases move forward on their template strands. This efficient chain-walking process is fueled by the hydrolysis of the triphosphate moieties of the monomers.[14]

Nevertheless, among all sequence-specific biological polymerizations, ribosomal protein synthesis is probably the most spectacular example in the sense that a sequence-ordered polymer chain is constructed from a template of very different chemical nature (*i.e.* nucleotide sequences of a single-strand mRNA template are translated into amino acid sequences). This fascinating aspect is in part due to the use of a very particular type of monomers. Indeed, in translation processes, each polymerizable amino acid is covalently attached to a specific molecule of transfer RNA *via* a labile ester bound.[12] This transient association allows the specific binding of the monomers with the mRNA template. The sequential

 This journal is © The Royal Society of Chemistry 2009

View Article Online

Published on 23 July 2009. Downloaded by Portland State University on 07/06/2014 09:31:55.

Table 1 "Through polymer chemists' eyes": a simplified view of the biological synthesis of sequence-ordered nucleic acids and proteins[a]

Polymerization	Initiation	Monomer	Reactive center[b]	Growth	Catalyst	Mechanism	Termination
Replication DNA → DNA	Initiation factor[c] RNA primer	Nucleoside triphosphate	3′-OH	Phosphodiester bond	DNA polymerase[c]	Template + chain-walking	Ligation, tus protein[d]
Transcription DNA → mRNA	Promoter + RNA polymerase	Nucleoside triphosphate	3′-OH	Phosphodiester bond	RNA Polymerase[c]	Template + chain-walking	Rho factor,[c] hairpin loop
Translation mRNA → protein	Start codon[e] + initiator tRNA	Aminoacyl-tRNA	C-terminus	Amide bond	Ribosome[f] (peptidyl transferase)	Template + chain-walking + transient monomer	Stop codon[g] + release factor[c]

[a] Generalization: differences between eukaryotic and prokaryotic mechanisms are not highlighted. [b] Reactive site of the growing macromolecule. [c] Proteins. [d] Tus stands for terminus utilizing substance. [e] AUG in most cases. [f] Protein–RNA complex. [g] UAA, UAG, or UGA.

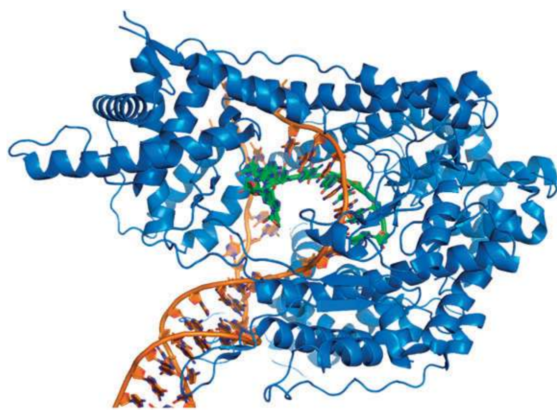

Fig. 1 Schematic representation of the transcription of a DNA-template (orange) into a sequence-defined mRNA chain (green), catalyzed by an RNA polymerase (blue). Artwork by T. Splettstoesser.

complexation of the aminoacyl-tRNA monomers with the template, their polymerization and the detachment of tRNA are regulated by ribosomes, which are large RNA–protein complexes.[15] Similarly to polymerases, ribosomes walk along the template strands during the polymerization process.

It is certainly premature to foresee fully-synthetic processes having the same degree of complexity as replication, transcription or translation. Current chemical tools do not allow the synthesis of catalysts, which could resemble polymerases or ribosomes. Nevertheless, the strategies (*e.g.* templated mechanisms) or the machineries (*e.g.* polymerases) of Nature can be simplified and exploited in man-made concepts. Several examples are described in the next paragraphs.

Synthetic pathways for preparing biopolymers

During the last 100 years, extensive research has been carried out for preparing natural biopolymers in artificial laboratory conditions. In particular, important progress has been made between the mid-1950's and the late 1970's for synthesizing sequence-defined oligopeptides and oligonucleotides. As mentioned in the introduction, the development of solid-phase synthesis was one of the most important breakthroughs in this area of research.[16] This technique was first optimized for the synthesis of peptides.[4] In this approach, the C-terminus of a N-protected amino acid is first linked to a crosslinked poly(styrene-*co*-divinyl benzene) bead *via* a labile covalent linkage. After attachment, the N-terminus of this first amino acid is deprotected and reacted with the C-terminus of a second N-protected amino acid, in the presence of a carbodiimide. Such deprotection/amidification cycles can be repeated several times to access oligopeptides with defined monomer sequences. The attachment of the peptide to a readily-filterable solid support allows straightforward and rapid purification steps. Nevertheless, this technique requires a highly optimized protection–deprotection chemistry. Besides the N-terminus of the amino acids, some reactive substituents (*e.g.* primary amines, guanidinium, carboxylic acids, alcohols, thiols) have to be protected. However, the side-chains protecting groups should be less labile than the one used at the N-terminus. All these aspects have been carefully optimized during recent decades.[17] Nowadays, synthetic protocols based on fluorenyl-methoxycarbonyl (Fmoc)-protected N-terminus are well established and widely used for synthesizing peptides.[11] Sequence-ordered oligomers with chain-lengths in the range of 2 to 50 amino acids can be routinely prepared with automated peptide synthesizers.

Merrifield's solid-phase strategy is indeed not restricted to peptides. This technique can be theoretically extended to any kind of step-growth polymerizations. For instance, solid-phase synthesis has been explored for preparing new types of sequence-defined oligomers based, for example, on ureas, carbamates or esters linkages. The synthesis of these non-biological structures is described in detail in the next paragraph. Another important application of solid-phase synthesis is certainly the preparation of sequence-ordered oligonucleotides.[18] In this approach, modified nucleotides are sequentially reacted on a solid support (typically a standard crosslinked polystyrene bead). Although various types of monomers have been investigated in the past decades, nucleoside phosphoramidites are nowadays primarily used in solid-phase oligonucleotide synthesis (Fig. 2). In these structures, the 3′-terminus is usually transformed into *N*,*N*-diisopropyl-*O*-cyanoethyl phosphoramidite, whereas the 5′-hydroxyl terminus is protected by di-*p*-methoxytrityl. In addition, some sites of the nucleobases (*i.e.* imide, lactam and exocyclic amine functions) and the 2′-hydroxy function in the case of ribonucleosides have to be protected as well. A comprehensive list of possible protecting groups can be found in a review of Beaucage and Iyer.[19]

Generally speaking, nucleoside phosphoramidites are relatively stable to hydrolysis or air oxidation and can therefore easily be

View Article Online

Published on 23 July 2009. Downloaded by Portland State University on 07/06/2014 09:31:55.

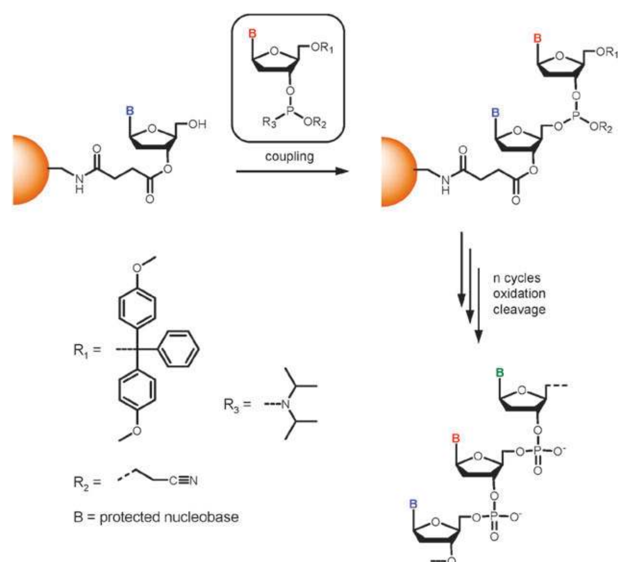

Fig. 2 Solid-phase synthesis of sequence-ordered oligonucleotides *via* nucleoside phosphoramidites oligomerization.[19]

stored as dried powders. However, in the presence of a weak acid (*e.g.* tetrazole), the 3′-terminus of these monomers is activated and reacts readily with free alcohol functions to form phosphotriester linkages, which can further be oxidized into stable phosphate triesters (*i.e.* protected phosphates).[18] Hence, this straightforward reaction can be conveniently used for growing oligonucleotides on a solid support. Typically, the 3′-terminus of the growing oligomer is attached to the support, whereas the 5′-terminus is involved in deprotection/reaction cycles (Fig. 2). After growth, the oligomer is cleaved from the support and deprotected into a natural nucleic acid structure.[18] Similarly to solid-phase peptide synthesis, this technique is limited to short oligomer chain-lengths. Yet, oligonucleotides with sequences of 100 to 200 nucleotides can nowadays be prepared on a solid support.

Much longer nucleic acids with defined monomer sequences can be synthesized in laboratory conditions using the polymerase chain reaction (PCR).[20] This process is not a chemical approach but a simplified *in vitro* version of DNA replication (see previous paragraph). In PCR, nucleic acid sequences are not newly designed as in solid-phase synthesis, but copied from existing templates. In fact, in terms of polymerization mechanism, PCR and replication are almost similar. In both cases, nucleoside triphosphates are enzymatically polymerized in the presence of a single-strand DNA-template. However, PCR requires experimental protocols, which significantly differ from physiological conditions. For instance, high temperatures (*i.e.* in the range 55–95 °C) are needed in this process for (i) dissociating hybridized DNA into single-strand templates and (ii) associating selectively short polymerization primers to the templates. Thus, thermally-stable polymerases are mandatory in PCR.[20] For instance, enzymes isolated from extremophile organisms can be used at very high temperatures. Several types of heat-tolerant polymerases have been identified and applied during the last twenty years. Nowadays, the PCR process is highly optimized and routinely used in molecular biology laboratories. Very recently, PCR procedures have been imported in synthetic polymer science and explored for synthesizing DNA-based block-copolymers.[21]

Nevertheless, PCR is not the only example of modified biological process used in biopolymer synthesis. Another important example is indeed genetic engineering.[9] This method relies on the cellular expression (*e.g.* in bacteria or yeasts) of artificial genes. For instance, the protein synthesis machinery of *Escherichia coli* can be exploited for preparing mutant proteins with defined sequences of either natural or non-canonical amino acids.[22] Typically, high molecular weight monodisperse polypeptides can be prepared in these conditions.[23] Thus, genetic engineering can be used for developing new varieties of protein-based materials.[9]

Solid-phase synthesis of non-biological polymers

As described in the preceding section, solid-phase synthesis has been primarily developed for the synthesis of biopolymers such as oligopeptides and oligonucleotides. However, peptide and phosphate linkages are definitely not the only types of polymer bonds, which can be formed on a solid support. In fact, virtually any monomer of the AB type (*i.e.* A and B are reactive termini, which can react with each other) can be sequentially polymerized on a solid support if one of the two reactive functions is temporarily protected. Thus, several types of nonnatural sequence-ordered oligomers have been synthesized on solid supports. Some selected examples are discussed in the present paragraph.

A wide variety of sequence-defined oligoamides have been reported in recent years. In particular, several examples of pseudo-peptides have been constructed by solid-phase chemistry. For example, β-peptides,[24] γ-peptides,[25] peptide nucleic acids (PNA)[26] and unnatural glycopeptides[27] can be easily synthesized using standard Boc- or Fmoc-based procedures (see previous paragraph for additional details). In all these cases, libraries of functional monomers have been developed and exploited for creating controlled pseudo-peptide sequences. Peptoids constitute another interesting class of pseudo-peptides, which can be prepared by solid-phase synthesis.[28] However, in these structures, the monomer units are not linked by conventional peptide bonds but *via* *N*,*N*-disubstituted amide linkages. Thus, these structures are usually constructed *via* stepwise oligomerization of AB monomers containing a –COOH terminus and a Fmoc-protected secondary amine function. Besides pseudo-peptides, diverse examples of unnatural polyamide backbones have been synthesized by solid-phase chemistry.[29–31] For example, Rose and Vizzavona prepared sequence-defined polyamides by successive reaction of diacids and diamines (*i.e.* building blocks of the AA and BB type). In this interesting sequential approach, the reactive amine- and acid-functions do not require a protecting group.[32]

Oligoureas and oligocarbamates with defined monomer sequences have been lately explored as novel peptidomimetics and foldamers. These unnatural biopolymers are in general prepared by solid-phase synthesis. For instance, Burgess and co-workers synthesized oligoureas by sequential reaction of phthalimide protected isocyanates.[34] Shortly after, Schultz and co-workers reported an original alternative method based on azido-4-nitrophenyl carbamate monomers.[35] More

View Article Online

Published on 23 July 2009. Downloaded by Portland State University on 07/06/2014 09:31:55.

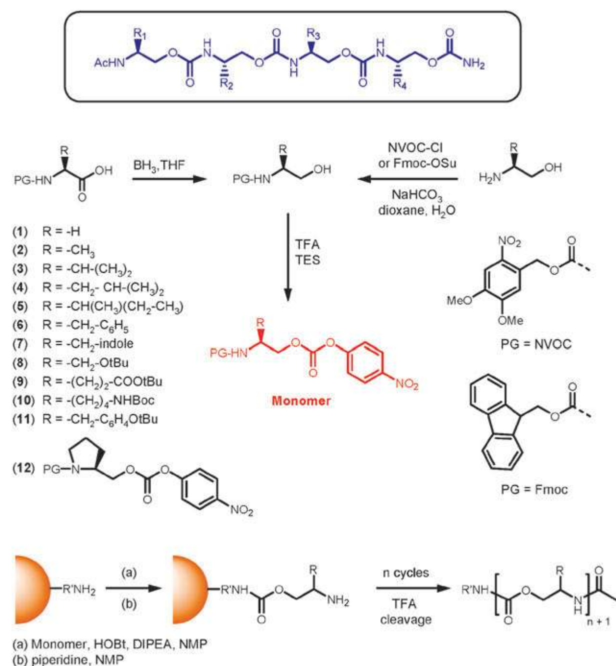

Fig. 3 Solid-phase synthesis of sequence-defined oligocarbamates: (top) general structure of the oligomers; (middle) synthesis of activated N-protected *p*-nitrophenyl carbonate monomers; (bottom) solid-phase protocols. R′ = H (Rink resin) or amino acid (Wang resin). Adapted from ref. 33. Copyright 1993 AAAS.

recently, a more conventional approach based on Boc-protected monomers have been described by the research group of Liskamp.[36] In comparison to oligoureas, very few studies have been reported in the literature about the solid-phase synthesis of oligocarbamates. So far, the most reliable approach was described by Schultz and co-workers and relies on the use of N-protected *p*-nitrophenyl carbonate monomers (Fig. 3).[33]

Numerous examples of solid-phase synthesis of sequence-defined polyamines have been reported in the literature. Detailed information on this topic can be found in the comprehensive reviews of Kong Thoo Lin *et al.*[37] and Papaioannou and Karigiannis.[38] Typically, linear polyamines can be synthesized by alkylation approaches or by reduction of oligoamide precursors. Very recently, Börner and co-workers reported fully automated protocols for the synthesis of monodisperse poly(amidoamine)s (PAA).[39] These segments were constructed by successive condensation of linear diamines with succinic anhydride (*i.e.* AA + BB strategy). This elegant approach was used to build up libraries of peptidomimetic PAA-*b*-poly(ethylene oxide) and PAA-*b*-oligopeptide conjugates where the cationic nature of the poly(amidoamine) block was varied. These defined oligocations were used as model molecules for understanding the mechanisms of polyelectrolyte–DNA complexation (*i.e.* polyplex formation for gene delivery).[39]

Several other types of defined unnatural oligomers such as oligoesters,[40] oligo(phenyl acetylene)s[41] oligo(phenylene ethynylene)s,[42] oligothiophenes[43] and oligosaccharides have been prepared on solid supports. The latter case constitutes a very established area of research, which is probably too broad to be covered in the present tutorial review. For additional information, the comprehensive reviews of Wong[44] and Seeberger[45] are highly recommended.

Sequence-control in liquid-phase polymerizations

Most of the current polymerizations methods, either step-growth or chain-growth, are performed in a batch mode. Reaction components are usually solubilized in the monomer (*i.e.* bulk polymerization) or in a monomer–cosolvent mixture (*i.e.* solution, emulsion, or dispersion polymerizations). These experimental procedures are far from the biological conditions described in the first paragraph of this review. Indeed, in Nature, polymer chains are synthesized one by one in confined environments, whereas, in man-made processes, a huge number of growing chains coexist in large reaction volumes. These conditions are clearly unfavorable for controlling sequences. However, in some cases, sequences can be controlled by physical or chemical means. These approaches are briefly reviewed in the present paragraph.

Step-growth polymerizations can be easily transformed into sequence-specific oligomerizations if appropriate protection–deprotection cycles are performed. As described in the previous two paragraphs, reactive monomers of the AB type can be sequentially polymerized if one of two functions is momentarily deactivated. In this approach, an undefined batch polymerization is simply decomposed into discrete reaction–deprotection–isolation steps. Thus, this type of chemistry does not necessarily require a solid support and can be performed in the liquid phase as well (*i.e.* using either sequential or convergent strategies). However, support-free approaches (*i.e.* solution synthesis) are dramatically limited by the purification steps and cannot be reasonably used for synthesizing long oligomers. Thus, some simplified procedures have been reported for controlling sequences in solution step-growth polymerizations. For instance, short sequence-defined repeating motifs can be created by selecting monomers of controlled reactivity.[46] Yet, these approaches are rather limited in scope. Alternatively, sequence-ordered oligomers can be directly synthesized on a soluble polymer segment. Indeed, linear macromolecules can be easily isolated from low molecular weight mixtures (*e.g. via* selective precipitation) and therefore used as efficient supports for organic synthesis.[47] For instance, some examples of oligonucleotide and oligopeptide synthesis on soluble polymer supports have been described in recent years.

Controlling monomer sequences in a chain-growth polymerization (*i.e.* polymerizations consisting of chain-initiation and chain-propagation steps) is, theoretically speaking, much more challenging than in a step-growth process.[7] Indeed, propagation steps rely on highly reactive transient species (*e.g.* radicals or ions), which are difficult to tame. Thus, chain-growth copolymerizations are in general statistical processes leading to random microstructures.[48] However, in some rare cases, sequences can be controlled.[49–51] One interesting exception is, for example, the radical copolymerization of styrene with cyclic monomers such as maleic anhydride or *N*-substituted maleimides. In such copolymerizations, the cross-propagation (*i.e.* the reaction of one comonomer with the other) is exceptionally favored as compared to

View Article Online

Published on 23 July 2009. Downloaded by Portland State University on 07/06/2014 09:31:55.

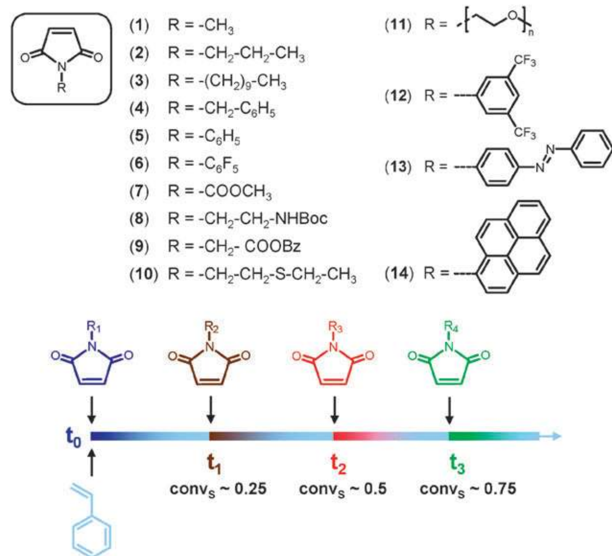

Fig. 4 Concept of the sequential atom transfer radical copolymerization of styrene and various *N*-substituted maleimides.[53,54]

homopolymerization.[50] Thus, conventional- or controlled-radical polymerizations (CRP) of these comonomer pairs typically lead to perfectly sequence-defined alternating copolymers. In fact, this tendency toward alternation is so pronounced that even for comonomer feeds containing a high excess of styrene, the cross-propagation still occurs in the early stages of the reaction, followed by the homopolymerization of the excess of styrene. For instance, Hawker, Russell *et al.* elegantly demonstrated that, if combined with a CRP process (*i.e.* a pseudo-living polymerization mechanism, in which all chains grow almost simultaneously), this kinetic behavior could result in the formation of well-defined block copolymers composed of short copolymer sequences connected to long polystyrene segments.[52]

We recently pushed this concept further and reported a novel sequential copolymerization strategy for preparing macromolecules with programmed sequences of functional comonomers.[53,54] This concept relies on the atom transfer radical copolymerization (ATRP) of functional *N*-substituted maleimides with styrene (Fig. 4). This copolymerization is a CRP process, which combines two unique kinetic features: (i) all the polymers chains are growing simultaneously and (ii) as aforementioned, the cross-propagation of the comonomers is highly favored. Thus, discrete amounts of *N*-substituted maleimides (*e.g.* 1 eq. as compared to initiator) are consumed extremely fast in the copolymerization process and are therefore locally incorporated in narrow regions of the growing polystyrene chains. MALDI-TOF analysis of model copolymers indicated that this kinetic concept is efficient.[54] Although a sequence-distribution is observed, well-defined polymer chains having only 1 or 2 functional maleimide units per chain were found to be the most abundant species. Furthermore, the position of the functional groups in the polystyrene chains can be kinetically-controlled by adding the *N*-substituted maleimides at desired times during the course of the polymerization. This method is very versatile and can be applied to a wide variety of *N*-substituted maleimides. For instance, a library of 20 different maleimides bearing various functional groups (*e.g.* aromatic moieties, fluorinated groups, hydroxy functions, protected esters, protected amines, light-responsive moieties, fluorophores and biorelevant functions such as short poly(ethylene glycol) segments or biotin moieties) was investigated. In most cases, the functional *N*-substituted maleimides could be efficiently incorporated in the polystyrene chains.[54]

Moreover, this concept could be extended for preparing macromolecules with programmed sequences of functional comonomers (Fig. 4).[53] Four different *N*-substituted maleimides (*N*-propyl maleimide, *N*-benzyl maleimide, *N*-methyl maleimide and *N*-[3,5-bis(trifluoromethyl)phenyl] maleimide) were consecutively added during the atom transfer radical polymerization of styrene performed at 110 °C in the presence of a catalyst composed of copper(I) bromide and 4,4′-dinonyl-2,2′-bipyridine. Still, the copolymer formed is not strictly sequence-defined at the molecular level. However, it possesses a pre-programmed distribution of functional side-groups. This proof of concept suggests that the present sequential method is a promising route for preparing copolymers with tailor-made microstructures.

In addition, other pathways have been studied for controlling sequences in chain-growth polymerizations. For example, Higashimura and co-workers investigated a concept for synthesizing sequence-ordered oligomers by living cationic polymerization.[55] In this strategy, a combination of HI and ZnI_2 was used to control the sequential oligomerization of vinyl ethers and styrene derivatives. For instance, HI reacts with a first monomer to form a stable iodo-adduct. The latter is a dormant species, which is only activated by ZnI_2. Hence, if this dormant adduct is activated in the presence of an equimolar equivalent of a second monomer, an iodo-terminated diadduct should be primarily formed. Theoretically, this strategy could be repeated to achieve sequence-defined oligomers. However, such an approach can only work if the activation/addition steps are kinetically favored as compared to homopolymerization. For instance, relatively monodisperse tetramers could be synthesized by reacting consecutively four monomers of decreasing reactivity.[55]

Polymers containing short sequence-defined repeating motifs can also be synthesized by ring opening polymerization of prestructured cyclic monomers. These interesting pathways have been recently reviewed by Cho.[56] Furthermore, some post-modification approaches have been described for controlling sequences. For example, Nishikubo and co-workers took advantage of the regioselective insertion of thiiranes in poly(S-aryl thioester) for synthesizing sequence-ordered motifs.[57] However, this attractive approach remains limited to very special classes of monomers and polymers.

Templated polymerizations

As discussed in the first paragraph of this review, sequence-specific biological polymerizations such as replication, transcription and translation rely in part on a template mechanism. Hence, during the last few decades, some research groups have investigated nucleic acid templates for the design of synthetic polymers with controlled monomer sequences.[58]

 This journal is © The Royal Society of Chemistry 2009

View Article Online

Published on 23 July 2009. Downloaded by Portland State University on 07/06/2014 09:31:55.

These strategies are, in most cases, an *in vitro* simplification of biological templating. Pioneer works in this area have been developed by Orgel, who reported interesting examples of nonenzymatic templated oligomerization.[59] In such approaches, activated phosphate units are preorganized on a template (*e.g.* DNA, RNA, PNA, hexitol nucleic acids (HNA) or altritol nucleic acids (ANA)) and subsequently polymerized. This strategy was proven to be efficient for transferring a sequence information from one type of nucleic acid to another (*e.g.* from DNA to RNA or from PNA to RNA). However, the sequence-specificity of these oligomerizations remains limited. The overall yields of formed polymers with correct sequence information rarely exceed a few percent. Richert and co-workers reported the use of 5′-acylamido-substituted DNA-templates to increase the selectivity and accelerate the incorporation of deoxyadenosine, deoxyguanosine and thymidine residues in single nucleotides extensions.[60] They demonstrated that cholic acid-modified templates made the reaction more specific and allowed a decrease in the error rates. Nevertheless, this approach remains limited to short chain-extensions.

The research group of Liu described promising examples of DNA-templated synthesis of peptide nucleic acids.[58] Interestingly, their approach does not rely on discrete monomer units but on the oligomerization of short sequence-defined tetrameric building blocks (Fig. 5). For instance, they first tried to combine five PNA tetramers using amine acylation as coupling reaction in the presence of DNA-templates and found out that this system led to the formation of 20% of full-lengths PNAs together with other mixture of products.[61] These moderate yields were attributed to the distance independence of the DNA-templated amine acylation reactions (*i.e.* the rate of formation of the products is not dependent on the distance between the annealed reagent and the template). In another attempt, basing their strategy on the work of Lynn and co-workers,[62] they used reductive amination (a distance-dependent reaction) templated by 5′-amino-terminated hairpin DNA oligonucleotides and obtained efficient, sequence-specific and regioselective oligomerization of PNAs (Fig. 5).[61] The efficiency is also maintained in the presence of multiple PNA building blocks bearing closely related sequences. On the basis of these results, Liu and co-workers recently expanded the scope of their studies to the synthesis of side-chain functionalized PNAs.[63]

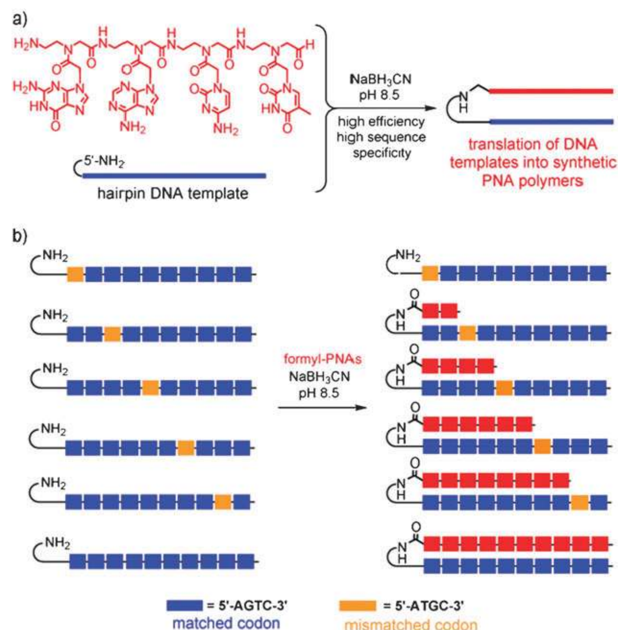

Fig. 5 DNA-templated formyl-PNA polymerization: (a) a 5′-amino-terminated DNA-template (blue) directs the oligomerization of a PNA tetramer (red); (b) mismatched codons (orange) in the template prevent the oligomerization of the tetramers. Reprinted with permission from ref. 58. Copyright 2004 Wiley-VCH.

Another appealing DNA-templated oligomerization has been described by Schuster and co-workers.[64,65] They synthesized oligoanilines (PANI) or oligo(4-aminobiphenyl) (PAB) by covalently linking the repeating units to the nucleobases of a DNA duplex. For instance, they functionalized a 22-mer strand DNA with aniline-modified cytosines.[64] Molecular modeling indicated that the aniline groups are involved in the major groove of the DNA double helix. Subsequently, they polymerized the aniline by adding horseradish peroxidase and hydrogen peroxide. This strategy allowed synthesis of high-conducting materials. Indeed, in comparison to the direct DNA-templated polymerization of aniline, the covalent linkage of aniline to the DNA nucleobase generated a head-to-tail *para* linkage, which led to a better structural orientation of the aniline chains and therefore increased their conductivity.[65] This interesting proof of concept indicates that DNA-templated oligomerizations can be exploited for preparing fully-synthetic structures. Hence, this model strategy could be certainly broadened to a wider palette of synthetic polymers.

Conclusions

An impressive research effort has been undertaken, during the last five decades, for characterizing, understanding, and controlling macromolecular sequences. Thus, important breakthroughs have been reported, for example, for the synthesis of oligopeptides and oligonucleotides. However, this research field still faces some limitations and challenges. For instance, sequence-defined biopolymers have been principally developed so far, while synthetic structures are still scarcely studied. Moreover, current technologies allow predominantly synthesis of short oligomers, whereas larger polymers (*e.g.* molecular weights comparable to those of structural proteins) with controlled monomer sequences are still out of reach. In addition, existing sequential processes, although automatized, remain, on the whole, complicated and pricy. Thus, the development of straightforward ‘‘one pot’’ polymerization procedures for controlling monomer sequence is certainly needed. In this regard, PCR is a particularly influential example.

However, to reach these goals, the research on sequence-specific polymerizations should be clearly intensified. In fact, many interesting strategies still have to be explored. As described in the last paragraph of this review, macromolecular templating is a promising pathway for tuning sequences. Nevertheless, several other ‘‘sleeping beauties’’ should be considered. For example, (i) controlled activation/deactivation

View Article Online

Published on 23 July 2009. Downloaded by Portland State University on 07/06/2014 09:31:55.

of growing species, (ii) kinetic control (*e.g.* starved conditions), (iii) chemical induction, (iv) selected reactivities, (v) transient complexation and (vi) confinement are potential strategies, which could be investigated for controlling polymer sequences. However, as learned from Nature, efficient sequence-specific pathways will probably not rely on a simple approach but almost certainly on a complex combination of physical- and chemical-conditions. Thus, an interdisciplinary research effort is most likely needed to design tomorrow's methods. In this regard, one can learn from the field of oligonucleotide synthesis. The current standard phosphoramidite approach is the product of a huge collective work.[19]

Nevertheless, the content of this review clearly shows that the control over macromolecular sequences is definitely not out of reach. Of course, it will certainly require time and efforts until new sequence-controlled polymerizations are developed. However, this research phase will be most probably shorter than the time used by Nature for developing replication. Realistically speaking, novel sequence-specific methodologies can be expected before the end of this century.

Notes and references

1 F. Sanger, *Nature*, 1948, **162**, 491–492.
2 M. F. Perutz, *Nature*, 1962, **194**, 914–917.
3 J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans and R. A. Holt, *et al.*, *Science*, 2001, **291**, 1304–1351.
4 R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.
5 F. A. Bovey and P. A. Mirau, *NMR of polymers*, Academic Press, Inc., San Diego, 1996.
6 *Mass Spectrometry of Polymers*, ed. G. Montaudo and R. P. Lattimer, CRC press LLC, Boca Raton, 2001.
7 J.-F. Lutz, T. Pakula and K. Matyjaszewski, *ACS Symp. Ser.*, 2003, **854**, 268–282.
8 J. J. Storhoff and C. A. Mirkin, *Chem. Rev.*, 1999, **99**, 1849–1862.
9 J. C. M. van Hest and D. A. Tirrell, *Chem. Commun.*, 2001, 1897–1904.
10 H. G. Börner and H. Schlaad, *Soft Matter*, 2007, **3**, 394–408.
11 J.-F. Lutz and H. G. Börner, *Prog. Polym. Sci.*, 2008, **33**, 1–39.
12 J. M. Berg, J. L. Tymoczko and L. Stryer, *Biochemistry*, 6th edn, W. H. Freeman, New York, 2006.
13 S. Polowinski, *Prog. Polym. Sci.*, 2002, **27**, 537–577.
14 J. Gelles and R. Landick, *Cell (Cambridge, Mass.)*, 1998, **93**, 13–16.
15 D. N. Wilson and K. H. Nierhaus, *Angew. Chem., Int. Ed.*, 2003, **42**, 3464–3486.
16 R. B. Merrifield, *Angew. Chem., Int. Ed. Engl.*, 1985, **24**, 799–810.
17 M. Amblard, J.-A. Fehrentz, J. Martinez and G. Subra, *Mol. Biotechnol.*, 2006, **33**, 239–254.
18 R. Eritja, *Int. J. Pept. Res. Ther.*, 2007, **13**, 53–68.
19 S. L. Beaucage and R. P. Iyer, *Tetrahedron*, 1992, **48**, 2223–2311.
20 R. Saiki, D. Gelfand, S. Stoffel, S. Scharf, R. Higuchi, G. Horn, K. Mullis and H. Erlich, *Science*, 1988, **239**, 487–491.
21 M. Safak, F. E. Alemdaroglu, Y. Li, E. Ergen and A. Herrmann, *Adv. Mater.*, 2007, **19**, 1499–1505.
22 R. E. Connor and D. A. Tirrell, *Polym. Rev.*, 2007, **47**, 9–28.
23 K. P. McGrath, M. J. Fournier, T. L. Mason and D. A. Tirrell, *J. Am. Chem. Soc.*, 1992, **114**, 727–733.
24 R. P. Cheng, S. H. Gellman and W. F. DeGrado, *Chem. Rev.*, 2001, **101**, 3219–3232.
25 T. Hintermann, K. Gademann, B. Jaun and D. Seebach, *Helv. Chim. Acta*, 1998, **81**, 983–1002.
26 E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chem., Int. Ed.*, 1998, **37**, 2796–2823.
27 H. Herzner, T. Reipen, M. Schultz and H. Kunz, *Chem. Rev.*, 2000, **100**, 4495–4538.
28 R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg and C. K. Marlowe, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 9367–9371.
29 N. R. Wurtz, J. M. Turner, E. E. Baird and P. B. Dervan, *Org. Lett.*, 2001, **3**, 1201–1203.
30 H. M. König, T. Gorelik, U. Kolb and A. F. M. Kilbinger, *J. Am. Chem. Soc.*, 2007, **129**, 704–708.
31 J. Farrera-Sinfreu, A. Aviñó, I. Navarro, J. Aymamí, N. G. Beteta, S. Varón, R. Pérez-Tomás, W. Castillo-Avila, R. Eritja, F. Albericio and M. Royo, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 2440–2444.
32 K. Rose and J. Vizzavona, *J. Am. Chem. Soc.*, 1999, **121**, 7034–7038.
33 C. Cho, E. Moran, S. Cherry, J. Stephans, S. Fodor, C. Adams, A. Sundaram, J. Jacobs and P. Schultz, *Science*, 1993, **261**, 1303–1305.
34 K. Burgess, H. Shin and D. S. Linthicum, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 907–909.
35 J.-M. Kim, Y. Bi, S. J. Paikoff and P. G. Schultz, *Tetrahedron Lett.*, 1996, **37**, 5305–5308.
36 A. Boeijen and R. M. J. Liskamp, *Eur. J. Org. Chem.*, 1999, 2127–2135.
37 V. Kuksa, R. Buchan and P. Kong Thoo Lin, *Synthesis*, 2000, 1189–1207.
38 G. Karigiannis and D. Papaioannou, *Eur. J. Org. Chem.*, 2000, 1841–1863.
39 L. Hartmann, S. Häfele, R. Peschka-Süss, M. Antonietti and H. G. Börner, *Chem.–Eur. J.*, 2008, **14**, 2025–2033.
40 T. M. Fyles, C. W. Hu and H. Luong, *J. Org. Chem.*, 2006, **71**, 8545–8551.
41 J. K. Young, J. C. Nelson and J. S. Moore, *J. Am. Chem. Soc.*, 1994, **116**, 10841–10842.
42 E. L. Elliott, C. R. Ray, S. Kraft, J. R. Atkins and J. S. Moore, *J. Org. Chem.*, 2006, **71**, 5282–5290.
43 C. A. Briehn, T. Kirschbaum and P. Bauerle, *J. Org. Chem.*, 2000, **65**, 352–359.
44 P. Sears and C.-H. Wong, *Science*, 2001, **291**, 2344–2350.
45 P. H. Seeberger, *Chem. Soc. Rev.*, 2008, **1**, 19–28.
46 M. Ueda, *Prog. Polym. Sci.*, 1999, **24**, 699–730.
47 D. J. Gravert and K. D. Janda, *Chem. Rev.*, 1997, **97**, 489–510; S. Pfeifer, Z. Zarafshani, N. Badi and J.-F. Lutz, *J. Am. Chem. Soc.*, 2009, **131**, 9195–9197.
48 G. Odian, *Principles of Polymerization*, 4th edn, John Wiley & Sons, Hoboken, 2004.
49 H. L. Hsieh, *US Pat.*, 3 661 865, 1969.
50 J. M. G. Cowie, *Alternating Copolymers*, Plenum Press, New York, 1985.
51 B. Kirci, J.-F. Lutz and K. Matyjaszewski, *Macromolecules*, 2002, **35**, 2448–2451.
52 D. Benoit, C. J. Hawker, E. E. Huang, Z. Lin and T. P. Russell, *Macromolecules*, 2000, **33**, 1505–1507.
53 S. Pfeifer and J.-F. Lutz, *J. Am. Chem. Soc.*, 2007, **129**, 9542–9543.
54 S. Pfeifer and J.-F. Lutz, *Chem.–Eur. J.*, 2008, **14**, 10949–10957.
55 M. Minoda, M. Sawamoto and T. Higashimura, *Macromolecules*, 1990, **23**, 4889–4895.
56 I. Cho, *Prog. Polym. Sci.*, 2000, **25**, 1043–1087.
57 A. Kameyama, Y. Murakami and T. Nishikubo, *Macromolecules*, 1996, **29**, 6676–6678.
58 X. Li and D. R. Liu, *Angew. Chem., Int. Ed.*, 2004, **43**, 4848–4870.
59 L. E. Orgel, *Acc. Chem. Res.*, 1995, **28**, 109–118.
60 J. A. Rojas Stütz and C. Richert, *J. Am. Chem. Soc.*, 2001, **123**, 12718–12719.
61 D. M. Rosenbaum and D. R. Liu, *J. Am. Chem. Soc.*, 2003, **125**, 13924–13925.
62 X. Li, Z.-Y. J. Zhan, R. Knipe and D. G. Lynn, *J. Am. Chem. Soc.*, 2002, **124**, 746–747.
63 R. E. Kleiner, Y. Brudno, M. E. Birnbaum and D. R. Liu, *J. Am. Chem. Soc.*, 2008, **130**, 4646–4659.
64 B. Datta, G. B. Schuster, A. McCook, S. C. Harvey and K. Zakrzewska, *J. Am. Chem. Soc.*, 2006, **128**, 14428–14429.
65 B. Datta and G. B. Schuster, *J. Am. Chem. Soc.*, 2008, **130**, 2965–2973.

 This journal is © The Royal Society of Chemistry 2009