From Sequence-Defined Macromolecules to Macromolecular Pin Codes

Joshua O. Holloway, Filip Van Lijsebetten, Nezha Badi, Hannes A. Houck, and Filip E. Du Prez*

1. Introduction

In recent years, polymer chemists worldwide have been striving to prepare monodisperse synthetic macromolecules that possess properties mimicking the ones found in nature. This has resulted in the emergence of the field of sequence-defined polymers.[1–9] While until recently, research has been mostly devoted toward the development of synthesis routes to design sequence-defined macromolecules,[1–3] attention is now shifting toward the potential use of these structures and their wider impact on our daily lives.[10,11] Hence, biomedical and biomimicking applications,[12,13] (bio)catalysis,[14] long-term data storage,[6,11,15,16] and product labeling for anti-counterfeiting purposes[17] are now all emerging. Information storage, in particular, is an area of increasing scientific interest as we move toward a more digital society, generating an ever-increasing amount of data that will soon exceed the chip-grade silicon resources required to store this.[18,19] As a result, polymer chemists are turning their interests toward a more compact and durable chemical solution.[11,20] While nonbiological information storage has already been successfully demonstrated using DNA,[21,22] along with its read-out using DNA sequencing,[23] there are drawbacks when considering this as a feasible data storage medium. These include, reliability as a result of possible degradation, a restricted number of building blocks (i.e., the monomer alphabet), limited rewritability, and cost.[15,20,24,25] In order to solve some of the drawbacks related to DNA, the use of discrete synthetic macromolecules has emerged as a viable alternative for data encoding at the molecular level and has the potential to greatly enhance data storage capacity.[10,11,15–17,20,26]

Although many protocols have been developed over recent years to store information on non-natural sequence-defined macromolecules,[15–17] data encryption still remains an important aspect in future molecular data management, as individuals and enterprises seek more sophisticated and fraud-resistant technologies.[10,17] To date, only a limited number of approaches using sequence-defined macromolecules have been developed for the specific purpose of data encryption.[16,27,28] For instance, Lutz and co-workers reported the use of sequence-defined poly(alkoxyamine amide)s for the encoding and decoding of information, which can be read-out via tandem mass spectrometric (MS) analysis.[27] Their study was completed by the total and uncontrolled erasing of the information using thermally induced degradation of the polymeric structure, which they claimed to be potentially easier to sequence than using biopolymers. Since then, the same group has expanded their research toward the elegant use of sequence-defined (oligo)urethanes as readable, molecular barcodes to label polymer-based materials for anti-counterfeiting purposes.[17] More recently, the use of photochemistry, as opposed to heat, has been used to introduce...
mutations in a sequence-defined code, altering the information carried within.\cite{28}

In a similar context, Meier and co-workers reported the use of multicomponent reactions for the preparation of synthetic macromolecules for means of secret communication.\cite{16} To achieve this, molecular messages were deposited onto supports such as envelopes, from which they could be extracted and later deciphered using MS/MS. The fragmentation patterns were used to read out the information that could be translated into an alphanumeric code.\cite{16} Unlike alternative strategies, this was not limited to a simple binary code, but utilizing 130 commercially available components lead to up to 500 000 possible combinations. Finally, our own research group synthesized thiolactone-based sequence-defined oligomers, using an acrylate-based alphabet to synthesize 71 different sequences in an automated fashion, which could be encoded and decoded to result in the unique storage of a QR code in a chemical form.\cite{15}

When developing data management technologies, the reading part of the process is arguably more important than the writing and must be unambiguously demonstrated, which poses challenges to both data storage and encryption in order to retrieve molecularly encoded information in a fast, straightforward, and non-destructive manner using readily available analytical techniques.\cite{10} Indeed, the current decryption of information-containing macromolecules is mainly carried out using MS/MS,\cite{6,10,29} a technique that requires a high investment cost and relies on specific and readily fragmentable bonds in order to give distinct fragmentation patterns that can be identified and rationalized.\cite{10} For more complex encoded data, however, MS/MS sequencing rapidly requires additional—and often in-house developed—computer algorithms.\cite{31}

Motivated by the need for both a more advanced encryption and a simplified read-out strategy, we herein introduce cleavable covalent linkages into the backbone of sequence-defined oligomers that exhibit a reliable and clean dynamic behavior at elevated temperatures. Specifically, we aimed to synthesize macromolecules bearing a numeric code by using the thermoreversible reaction of 1,2,4-triazoline-3,5-dione (TAD) coupling reagents with indole substrates.\cite{32–34} This was used in combination with the Passerini three-component reaction (P-3CR), which is the one-pot reaction between a carboxylic acid, aldehyde or ketone, and isocyanide (Figure 1).\cite{35} In contrast to the thermal unzipping of sequences that requires high degradation temperatures,\cite{23} the bond scission profile of TAD-indole linkages is accessible at a much milder temperature range (i.e., 60–150 °C). This guarantees cleavage to occur in a reliable manner at specific sites along the backbone, thereby resulting in simplified structural fragments that can be decrypted by 1D MS techniques. Moreover, the dynamic nature also offers the potential to scramble the TAD-indole combinations, allowing the encryption of the molecularly encoded information. While the word “encryption” generally means “to convert information (…) into a secret code that people cannot understand or use on normal equipment,”\cite{36} we are more specifically referring to encryption as the concealment of the numeric code in the context of this study.

Following our interests in the synthesis of precise sequence-defined macromolecules, the aim of this research was, to demonstrate our encryption/decryption protocol through the design of macromolecular pin codes whereby chemical digits (e.g., 1–4, Figure 1) were written in a defined order using commercially available P-3CR building blocks. Importantly, tailored indole

---

**Figure 1.** Schematic overview of the encryption and decryption of the macromolecular pin codes. a) The reversible TAD-indole covalent bonds imbedded in the backbone of the pin codes enable cleavage of the digits upon heating. b) Upon cooling, the digits reconnect with one another in a random manner, thereby leading to a statistical mixture of unreadable oligomers. c) The encrypted pin code can only be decrypted upon subsequent heating in the presence of a trapping agent, which allows the monomer digits to separate from one another. d) The resulting mixture of monomers can eventually be read out via ESI-MS analysis and their original order can be deciphered from the attached chemical fingerprints, hence revealing the initially written pin code.
compounds, bearing a distinct isotopic pattern, were simultaneously introduced during the writing process in order to mark the digit’s position within the sequence (i.e., A to D) and enable easier read-out. As the digits are connected through the thermally dynamic TAD-indole linkages, the written pin codes can readily be dissociated upon heating (Figure 1a) and subsequently transformed into a statistical distribution of oligomers upon cooling to room temperature (Figure 1b), resulting in unreadable information. Additionally, the individual monomer digits could later be isolated by heating in the presence of a kinetic trap (Figure 1c) and the information could eventually be retrieved by means of electrospray ionization-MS (ESI-MS) analysis, based on their difference in mass (Figure 1d). The original position of the digit in the macromolecular pin code could hence be deciphered through the attached “chemical fingerprints” of the indole compounds. In other words, we demonstrate that only those who know the initial order of the indole marker compounds (here ABCD) are able to decipher the encoded information and thus recover the original pin code (thereby for instance differentiating between 1234 and 4321).

2. Results and Discussion

2.1. Writing Macromolecular Pin Codes

Prior to evaluating our above outlined encryption/decryption strategy (see Figure 1; Movie S1, Supporting Information), we first identified a chemical procedure to write TAD-indole-based molecular pin codes that could undergo the desired reversible transformations. For this, a recently established iterative synthesis protocol that combines the reactivity of TADs with thermoreversible TAD-indole chemistry. For this, kinetic reversibility experiments of PC1 (0.02 M in DMSO-d6) were first conducted at three distinct temperatures (i.e., 120, 135, and 150 °C) in the presence of a slight excess of a low molecular weight conjugated diene (i.e., (E,E)-2,4-hexadien-1-ol). The latter served as an irreversible and kinetically favored trap, forming an irreversible TAD–Diels–Alder adduct, effectively quenching any released TAD as the pin code was subjected to elevated temperatures. Thus, the fraction of TAD-indole dissociation could be closely monitored over time through offline 1H NMR measurements (Figure S3, Supporting Information). Several key proton resonances indeed provided the confirmation of the successful TAD-indole dissociation, thus demonstrating the successful deconstruction of the sequence-defined tetramer.

As expected, the temperature had a significant effect on the rate of oligomer dissociation, with complete conversion to the diene-capped monomers—and hence full dissociation of the pin code—reached within 15 min at 150 °C, 45 min at 135 °C, and 180 min at 120 °C (Figure S4, Supporting Information). Moreover, an observed activation energy $E_{a,obs} = 116.1 \pm 8.5$ kJ mol$^{-1}$ could be derived from the kinetic study (Table S1, Supporting Information), which is in accordance with the values of previous studies on simple low molecular weight TAD-indole systems (see Supporting Information for details).

While it was demonstrated that the reversibility is relatively fast at 150 °C, this high temperature is often regarded as the upper limit at which reliable TAD-indole exchange reactions can occur without significant TAD-degradation. In order to anticipate this potential issue, a second macromolecular pin code (PC2) was synthesized using the alternative aromatic TAD-COOH derivative L1b, as this was previously reported to significantly reduce the temperature of reversibility of the corresponding indole adduct compared to its aliphatic counterpart. Similar to the first pin code, PC2 was again characterized using SEC (Figure 3b), LCMS, and NMR (see Supporting Information) during the pin code writing process and complemented by MALDI-MS/MS to retrieve the encoded information (see Figure S2, Supporting Information). The thermal reversibility kinetics was also investigated by offline 1H NMR (Figure S5, Supporting Information). As anticipated, complete dissociation of PC2 was now observed to be completed in a faster timescale, that is, after 15 min at 120 °C (Figure S6, ...
Supporting Information), with an observed activation energy $E_{a,obs} = 95.7 \pm 5.3$ kJ mol$^{-1}$ (Table S1, Supporting Information). Besides a significant acceleration of the encryption process, these results highlight the potential to further tune the thermoreversible properties by changing the structure of the TAD derivative.

2.3. Encryption of Macromolecular Pin Codes

Following the kinetic studies designed to gain basic insights into the thermal deconstruction process of the pin codes, the encryption process—whereby the written information is (temporarily) secured—was evaluated. In this context, both PC1 and PC2 were heated for 15 min at 150 and 120 °C, respectively, this time in the absence of any co-reactant. During this heating step, the retro-reaction of the TAD-indole covalent bonds was initiated, which led to the complete dissociation of the monomer units from the oligomer (see Figure 1a). Since no trapping agent was present during this encryption stage, the in situ liberated TAD moieties were no longer scavenged from the reaction mixture and consequently allowed the TAD-indole systems to exchange with one another. Subsequent cooling to room temperature eventually resulted in a recombination of the covalent bonds, albeit in a completely random order (see Figure 1b). Hence, a statistical mixture of disperse oligomers was formed, as evidenced by the SEC elugrams (Figure 3c,d), confirming the successful encryption of the macromolecular pin codes. Moreover, MALDI-MS/MS analysis of the encrypted pin codes showed that the original order of the sequence could no longer be determined, as different dimer combinations were observed, which resulted in ambiguity during read-out (Figures S7 and S8, Supporting Information). Additionally, the mixture proved to be unreadable by means of ESI-MS (Figures S9 and S10, Supporting Information), although the Diels–Alder adduct formed as a result of starting with $S_1$, thus
encoding A1, could be identified from the ESI-MS spectrum. The sacrificial identification of the first digit during encryption could be avoided, however, by either using an indole instead of the conjugated diene to initiate the sequence growth, or by first incorporating a non-information containing dummy digit.

2.4. Decryption and Read-Out of Macromolecular Pin Codes

In a final stage, we aimed to retrieve the encrypted information embedded in the macromolecular pin codes. As with the encryption studies, a heat treatment was carried out on PC1 and PC2 (at 120 °C for 5 and 2 h, respectively), but this time in the presence of the conjugated diene trapping agent S1, which was also used as the anchor from which the iterative sequence-defined growth was initiated. As the TAD-indole bonds of the unreadable oligomeric mixture opened at elevated temperature, the sequence could now be effectively dissociated into its separate monomer units (see Figure 1c). When cooling the mixture in the presence of S1, the Diels–Alder reaction between the newly liberated TAD moieties and conjugated diene was kinetically favored over the recombination reaction with the indole compounds, which is known from our previous research.[33] In other words, S1 acted as a kinetic trap to isolate the monomer units.

In contrast to the previously obtained disperse statistical distribution of oligomers, the resulting mixture of defined monomer fragments could be readily identified by means of LCMS and ESI-MS (Figures S12 and S13, Supporting Information). The initial position of the different monomer digits could be deciphered using the chemical fingerprint associated with their corresponding substituted indole markers (Figure S14, Supporting Information). Indeed, by making use of the unique isotopic patterns, the related m/z values could be linked to the correct chemical structure of the digits based on their difference in mass. As a result, the different positions and contained information of the digits could be acquired and the molecular pin codes deciphered, thereby retrieving the initial sequencing order (Figure 3e,f). Note, the ESI-MS can be complemented with LCMS when poor ionization of the cleaved monomer fragments is observed. It should be also noted that if (part of) the pin code is revealed, the synthesis of a new macromolecular pin code would be required.

2.5. Anti-Counterfeiting Demonstration

Among the many emerging applications of sequence-defined macromolecules is their use as chemical labels for anti-counterfeiting purposes.[17] To demonstrate the use of our conceptual encryption/decryption approach in this context, the encrypted oligomers were deposited onto biaxially oriented polypropylene banknotes, to add a concealed chemical label as an additional layer of security. In general, such polymer-based banknotes contain multiple layers of printing and security features, which are ever-changing and evolving to protect the currency from counterfeiting.
Thus, as a proof-of-concept study, a £5 Bank of England and a $5 Reserve Bank of Australia polymer banknote (Figure 4a), which had been in circulation and hence subjected to the usual mechanical wear of daily use, were labeled, respectively, with the encrypted PC1 and PC2. Prior to this, the banknotes were suspended overnight in ethanol to verify that they remained unaltered in the solvent (Figure S15, Supporting Information). Furthermore, the products extracted from the banknote during suspension were analyzed by ESI-MS in order to record a background reference (see Figure S16, Supporting Information). Ethanol was chosen here as non-toxic solvent, which could be used to both deposit and extract the chemical information. Approximately 10 mg of encrypted PC1 and PC2 was dissolved in a minimal amount of ethanol and deposited on the surface of the respective banknotes. Mild heating (<50 °C) was applied to evaporate the solvent and the notes were then left to dry overnight. To the casual eye, the encoded banknotes had the same appearance as the pristine ones (Figure 4a; Figure S15, Supporting Information). They were placed again in ethanol for 5 h to re-extract the encrypted pin codes from the surface (Figure 4b,c). Solvent removal in vacuo allowed for the recovery of the encrypted codes, which were finally decrypted by heating in butyl acetate (4 h at 120 °C) in the presence of the trapping agent S1. Eventually, the initially written pin codes were successfully deciphered using ESI-MS, following the above-outlined strategy (see Figures S17 and S18, Supporting Information).

3. Conclusion

In this work, the successful combination of the P-3CR and TAD-indole chemistry was demonstrated for the synthesis of dynamic sequence-defined macromolecules. The thermoreversible TAD-indole reaction was utilized to construct, scramble, and deconstruct well-defined macromolecular structures, carrying a molecular pin code. The introduction of a range of indole-aldehyde analogues meant that the original order of the sequence could be readily deciphered and read-out through distinct isotopic patterning (or “chemical fingerprinting”) using ESI-MS, rather than the currently employed sophisticated tandem mass analyses. Such tandem mass analysis was used in this study for the sole purpose of unambiguously proving the necessity of the thermal encryption for the complete concealment of the original sequence order following scrambling. By altering the structure of the TAD-COOH compound, we were able to tune the temperature of reversibility to conduct both the encryption and decryption process. Finally, we visually demonstrated how this conceptual approach could potentially be used to further advance the use of precision chemistry for anti-counterfeiting applications. Currently, the use of four different isocyanides allows for 44 (256) possible combinations, which translates to a small storage capacity of 8 bits. Our strategy can, however, be expanded to a wide range of other isocyanide compounds, thereby significantly extending the amount of codable pin code digits and hence increase the amount of possible permutations to $n^4$. The implementation of a reliable, dynamic covalent chemistry within sequence-defined macromolecules could be considered as an important advancement in the development of the next generation of encryption and decryption technologies.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

J.O.H. would like to acknowledge funding of his PhD scholarship from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 642083 and from the Ghent University Special Research Fund-ITN (BOF-ITN). H.A.H. thanks the Research Foundation-Flanders (FWO) for the funding of his PhD fellowship. The authors would like to thank Chiel Mertens...
for his assistance in MALDI-MS/MS measurements. Additionally, this project received funding and support from the FWO under EOS-project 30650939. The authors would like to thank Bernhard De Meyer, Jan Goeman, and Timothee Courtin for technical assistance.

Conflict of Interest
The authors declare no conflict of interest.

Keywords
1,2,4-triazoline-3,5-dione, chemical encryption, multicomponent reactions, sequence-defined macromolecules

Received: December 19, 2019
Revised: February 10, 2020
Published online:

[1] S. Martens, J. O. Holloway, F. E. Du Prez, Macromol. Rapid Commun. 2017, 38, 1700469.
[2] S. C. Solleder, R. V. Schneider, K. S. Wetzel, A. C. Boukis, M. A. R. Meier, Macromol. Rapid Commun. 2017, 38, 1600711.
[3] J. A. De Neve, J. J. Haven, L. Maes, T. Junkers, Polym. Chem. 2018, 9, 4692.
[4] N. Badi, J.-F. Lutz, Chem. Soc. Rev. 2009, 38, 3383.
[5] S. C. Solleder, D. Zengel, K. S. Wetzel, M. A. R. Meier, Angew. Chem., Int. Ed. 2016, 55, 1204.
[6] N. Zydziak, W. Konrad, F. Feist, S. Afonin, S. Weidner, C. Barner-Kowollik, Nat. Commun. 2016, 7, 14166.
[7] F. A. Leibfarth, J. A. Johnson, T. F. Jamison, Proc. Natl. Acad. Sci. USA 2015, 112, 10617.
[8] J. O. Holloway, C. Mertens, F. E. Du Prez, N. Badi, Macromol. Rapid Commun. 2019, 40, 1800685.
[9] J. O. Holloway, S. Aksakal, F. E. Du Prez, C. Remzi Becer, Macromol. Rapid Commun. 2017, 38, 1700500.
[10] M. A. R. Meier, C. Barner-Kowollik, Adv. Mater. 2019, 31, 1806027.
[11] J.-F. Lutz, Macromolecules 2015, 48, 4759.
[12] L. Hartmann, Macromol. Chem. Phys. 2011, 212, 8.
[13] M. Porel, D. N. Thornlow, C. M. Artim, C. A. Alabi, ACS Chem. Biol. 2017, 12, 715.
[14] P. Chandra, A. M. Jonas, A. E. Fernandes, J. Am. Chem. Soc. 2018, 140, 5179.
[15] S. Martens, A. Landuyt, P. Espeel, B. Devreese, P. Dawyndt, F. E. Du Prez, Nat. Commun. 2018, 9, 4451.
[16] A. C. Boukis, K. Reiter, M. Frölich, D. Hofheinz, M. A. R. Meier, Nat. Commun. 2018, 9, 1439.
[17] U. S. Gunay, B. E. Petit, D. Karamessini, A. Al Ouahabi, J. A. Amalian, C. Chendo, M. Bouquey, D. Gigmes, L. Charles, J.-F. Lutz, Chem 2016, 1, 114.
[18] M. Hilbert, P. López, Science 2011, 332, 60.
[19] D. Panda, K. A. Molla, M. J. Baig, A. Swain, D. Behera, M. Dash, 3 Biotech 2018, 8, 239.
[20] J. Steinkönig, R. Aksakal, F. E. Du Prez, Eur. Polym. J. 2019, 120, 109260.
[21] C. Bancroft, T. Bowler, B. Bloom, C. T. Clelland, Science 2001, 293, 1763c.
[22] G. M. Church, Y. Gao, S. Kosuri, Science 2012, 337, 1628.
[23] H. Mutlu, J.-F. Lutz, Angew. Chem., Int. Ed. 2014, 53, 13101.
[24] J. Bonnet, M. Colotte, D. Couidy, V. Couallier, J. Portier, B. Morin, S. Tuffet, Nucleic Acids Res. 2010, 38, 1531.
[25] T. Lindahl, Nature 1993, 362, 709.
[26] A. C. Boukis, M. A. R. Meier, Eur. Polym. J. 2018, 104, 32.
[27] R. K. Roy, A. Meszynska, C. Laure, L. Charles, C. Verchin, J.-F. Lutz, Nat. Commun. 2015, 6, 7237.
[28] N. F. König, A. Al Ouahabi, L. Oswald, R. Szewda, L. Charles, J.-F. Lutz, Nat. Commun. 2019, 10, 3774.
[29] J.-A. Amalian, G. Cavallo, A. Al Ouahabi, J.-F. Lutz, L. Charles, Anal. Chem. 2019, 91, 7266.
[30] G. Cavallo, S. Poyer, J.-A. Amalian, F. Dufour, A. Burel, C. Carapito, L. Charles, J.-F. Lutz, Angew. Chem. 2018, 130, 6374.
[31] A. Burel, C. Carapito, J.-F. Lutz, L. Charles, Macromolecules 2017, 50, 8290.
[32] H. A. Houck, K. De Bruycker, C. Barner-Kowollik, J. M. Winne, F. E. Du Prez, Macromolecules 2018, 51, 3156.
[33] H. A. Houck, K. De Bruycker, S. Billiet, B. Dhanis, H. Goossens, S. Catak, V. Van Speybroeck, J. M. Winne, F. E. Du Prez, Chem. Sci. 2017, 8, 3098.
[34] S. Billiet, K. De Bruycker, F. Driessen, H. Goossens, V. Van Speybroeck, J. M. Winne, F. E. Du Prez, Nat. Chem. 2014, 6, 815.
[35] M. Passerini, Gazz. Chim. Ital. 1921, 51, 126.
[36] Cambridge University Press, Encryp can be found under www.dictionary.cambridge.org/dictionary/english/encrypt (accessed: January 2020).
[37] J. O. Holloway, K. S. Wetzel, S. Martens, F. E. Du Prez, M. A. R. Meier, Polym. Chem. 2019, 10, 3859.
[38] H. Wamhoff, K. Wald, Chem. Ber. 1977, 110, 1699.
[39] The authors would like to declare that permanently defacing or destroying banknotes of the Bank of England or the Reserve Bank of Australia was neither intended nor carried out. As can be seen in Figure 4c, the banknotes were cleaned following the extraction of the macromolecular pin codes and were thus returned to their original condition.