Nature-Inspired Circular-Economy Recycling for Proteins: Proof of Concept

Simone Giaveri, Adeline Marie Schmitt, Laura Roset Julià, Vincenzo Scamarcio, Anna Murello, Shiyu Cheng, Laure Menin, Daniel Ortiz, Luc Patiny, Sreenath Bolisetty, Raffaele Mezzenga, Sebastian Josef Maerkl, and Francesco Stellacci**

The billion tons of synthetic-polymer-based materials (i.e. plastics) produced yearly are a great challenge for humanity. Nature produces even more natural polymers, yet they are sustainable. Proteins are sequence-defined natural polymers that are constantly recycled when living systems feed. Digestion is the protein depolymerization into amino acids (the monomers) followed by their re-assembly into new proteins of arbitrarily different sequence and function. This breaks a common recycling paradigm where a material is recycled into itself. Organisms feed off of random protein mixtures that are "recycled" into new proteins whose identity depends on the cell's specific needs. In this study, mixtures of several peptides and/or proteins are depolymerized into their amino acid constituents, and these amino acids are used to synthesize new fluorescent, and bioactive proteins extracellularly by using an amino-acid-free, cell-free transcription–translation (TX–TL) system. Specifically, three peptides (magainin II, glucagon, and somatostatin 28) are digested using thermolysin first and then using leucine aminopeptidase. The amino acids so produced are added to a commercial TX–TL system to produce fluorescent proteins. Furthermore, proteins with high relevance in materials engineering (β-lactoglobulin films, used for water filtration, or silk fibroin solutions) are successfully recycled into biotechnologically relevant proteins (fluorescent proteins, catechol 2,3-dioxygenase).

1. Introduction

The world's projected population will be 10 billion by 2050.[1] One of the most daunting sustainability challenges linked to such a large population size will be the handling of all plastic products,[2] that is, the production and recycling of polymers.[3] Not surprisingly, there are large world-wide efforts in research for polymer recycling. Mechanical recycling tends to lead to the original material but with lower quality.[4] A better possibility is chemical recycling,[5,6] that is, thermally,[7] chemically,[7] or biologically[8] catalyzed depolymerization of a polymer into its constituent monomers in order to re-polymerize them into either the same virgin quality material, or a new (co)polymer.[9,10] Another approach is repurposing of a polymer into a different value-added chemical (upcycling).[11–15] Both methods are a closed-loop, that is, compatible with circular-economy principles.[16]

S. Giaveri, A. M. Schmitt, L. Roset Julià, V. Scamarcio, A. Murello, F. Stellacci
Institute of Materials
École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne 1015, Switzerland
E-mail: francesco.stellacci@epfl.ch

S. Giaveri, S. Cheng, S. J. Maerkl, F. Stellacci
Institute of Bioengineering
École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne 1015, Switzerland
E-mail: sebastian.maerkl@epfl.ch

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adma.202104581.

DOI: 10.1002/adma.202104581

A. M. Schmitt
Faculty of Chemistry
Université de Strasbourg
Strasbourg 67000, France

L. Menin, D. Ortiz, L. Patiny
Institute of Chemical Sciences and Engineering
École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne 1015, Switzerland

S. Bolisetty, R. Mezzenga
Institute of Food, Nutrition and Health
ETH Zurich
Zürich 8092, Switzerland

R. Mezzenga
Department of Materials
ETH Zurich
Zürich 8093, Switzerland

It is fair to state that recycling a material into the same material is the current paradigm in recycling. To go beyond this paradigm, current trends in polymer recycling involve their degradation into small molecules, that are then re-used in further chemical processes. Alternative approaches include the use of biosourced/-degradable polymers, that is, materials that are derived from renewable sources and that can be degraded into environmentally benign substances.[17,18] This approach takes inspiration from the way nature handles some natural polymers such as, lignin and cellulose. Yet, these natural materials grow slowly, remain in use a long time, and biodegrade slowly. This balance is always present in nature's approach to recycling. Currently, man-made biosourced/-degradable polymers are produced for consumers products that often have very short lifetimes (days or less), but in the environment degrade over months or years. As a consequence, no matter how "green" such materials will appear to be, there will be significant environmental concerns due to their accumulation into the environment. By 2050, $\approx 10^{12}$ kg of plastics are projected to be produced yearly.[3] Were all polymers to be biosourced and biodegraded (i.e., the best-case scenario), the sustainability problem would remain. Sourcing will generate issues in deforestation and in competition for land with food production.[19] The issues in disposing will be the accumulation into the environment. A major concern will be the deleterious effects of the intermediate degradation components on soils.[4] Moreover, the final degradation products will have negative ecological effects, as such large quantities will inevitably shift the equilibrium of local ecosystems.[4]

It is clear that humans should move toward the use of models that rely on the principles of a circular economy[6,20] where materials, once produced, remain in use for the longest possible amount of time, taxing earth the minimum possible. The question is whether this is at all possible for polymers, that mostly lose quality upon recycling as opposed, for example, to metals. To address this question, one can be inspired by nature. It is undeniable that nature is sustainable: it takes most of its energy from the sun, food production is commensurate to population, and materials are used in a circular manner. While we have more than 1 billion tons of biological soft-matter produced on earth yearly, we do not have a sustainability concern with it. When pausing to observe nature's main polymers, for example, proteins, each one characterized by its own specific sequence of monomers, the 20 proteinogenic amino acids (AAs), it is possible to admire the circularity in their use. A vast over-simplification of protein metabolism shows that, proteins can be depolymerized into AAs that, in turn, can be reassembled into a new protein by the ribosomal machinery of the cell. The newly formed protein can have a sequence that differs from any of the sequences of the original proteins. It is fair to say that this approach breaks the recycling paradigm, that is, that materials are recycled into a lower version of themselves. In nature this is not the case, a protein can be of much higher complexity than its "parent" proteins with which it has only the individual AA building blocks in common. Nature can achieve this impressive result because proteins are sequence-defined polymers (SDPs), that is, their remarkable structural and property diversity derives from the sequence of the AAs that compose them, and not from their chemical diversity.[21] Furthermore, the backbone chemical bonds that link AAs are reversibly cleavable and there exists a machinery (the ribosome) capable of synthesizing proteins starting from a random mixture of AAs. Nature is able of recycling a mixture of n SDPs into an arbitrary (n+1)th SDP whose sequence (and consequently property) can be completely different from any of the sequences of the n parent polymers. Here we show that the described recycling approach can be implemented in the laboratory extracellularly for proteins, protein mixtures, and protein materials. We call this approach nature-inspired circular-economy recycling (NaCRe) (**Figure 1**).

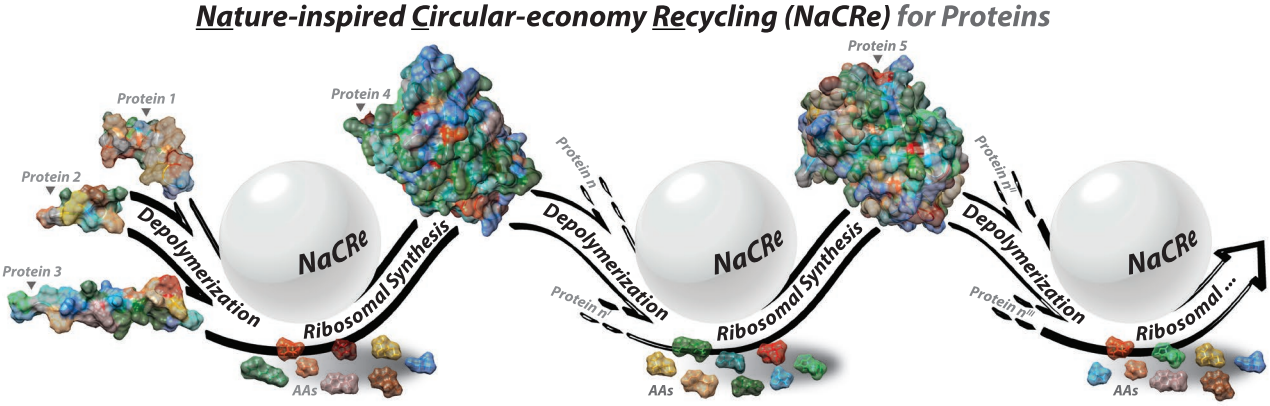

Figure 1. Schematic illustration for the main concept of NaCRe. Multiple possible NaCRe cycles are shown. The illustrated examples are close to what is shown herein. It should be clear that the overall concept of NaCRe goes beyond what is illustrated. The sketched process starts from three different short peptides (drawn as the ones used in the paper, magainin II, glucagon, and somatostatin), and produces GFP. In the second round of recycling, GFP, together with other arbitrary proteins, is used to produce red fluorescent protein (mScarlet-i). In the last recycling round mScarlet-i is recycled into something not specified, to stimulate the reader's imagination. Molecular graphics of the proteins 3D structures and of the AAs conformers are visualized from the experimentally determined 3D structures in the PDB databank (protein 1(2LSA), protein 2(2MI1), protein 3(1GCN), protein 4(5B61), and protein 5(5LK4)) and from the computed 3D structures in PubChem (https://pubchem.ncbi.nlm.nih.gov/compound/CID#section=3D-Conformer, where CID = 5950, 6322, 5960, 5961, 33032, 6274, 6306, 6106, 5962, 6137, 6140, 145742, 5951, 6305, 6057), respectively (e.g., https://pubchem.ncbi.nlm.nih.gov/compound/5950#section=3D-conformer for CID = 5950). All were edited in UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at UCSF, with support from NIH P41-GM103311.

Here, we aim to show that the current revolution in using more and more protein-based materials to realize advanced objects[22–25] has one more advantage: proteins are recyclable in a unique way. Arguably, it would be a breakthrough if, in the future, a large quantity of different objects all made of various protein-based materials could be NaCRe-recycled into the protein that a community needs in that specific moment. Clearly, this vision will take decades (if not centuries) to be implemented, as the technological challenges are significant. This paper is intended as a proof-of-concept where we present the feasibility of the overall process, showing that it is possible to recycle mixtures of peptides and engineering-relevant proteins into proteins with relevance in biotechnology outside living organisms.

2. Results and Discussion

The initial attempt to establish the feasibility of NaCRe was performed by enzymatically depolymerizing three peptides separately, and by recombining the AAs so achieved using the cell machinery to express a target protein. The latter task was achieved in a standard method. We purchased a commonly used cell-free transcription–translation (TX–TL) system (PURE, Protein synthesis Using Recombinant Elements, PURExfrex, Kaneka Eurogentec SA, Supporting Information, Section a) that is known to "transcribe" the information that we provided by feeding a specific DNA into a messenger RNA (mRNA), and then "translate" the mRNA code by "polymerizing" the target protein. The main issue with commercial TX–TL systems is that they contain free AAs. We chose PURExfrex because it is composed of multiple separate solutions, with only one of them that contains free AAs, and it is relatively simple to replace such solution with a home-made one that is AAs-free. The home-made solution lacking the AAs was produced by using a protocol adapted from the original reference from Ueda and co-workers.[26] It should be noted that the PURExfrex system contains a single AA (glutamic acid) as a component of one of the other solutions. Hereafter, we will refer to this home-made AAs-free form of PURExfrex simply as TX–TL system. To establish the absence of AAs in our TX–TL system, we performed control experiments that show the lack of any detectable protein expression (Figure S1, Supporting Information). In order to have a simple way to detect protein expression in the TX–TL system, we decided to focus all the work presented here on expressing fluorescent proteins. As a first choice, we focused on mScarlet-i, a fluorescent protein whose sequence contains 19 of the 20 proteinogenic AAs with cysteine missing. For later work, we expressed green fluorescent protein (GFP) as it is the most commonly expressed fluorescent protein and it contains all 20 proteinogenic AAs.

We felt that it would be simpler to develop a robust depolymerization method starting with shorter molecules, thus our initial attempts were based on short peptides. We selected magainin II, and glucagon by reading the whole PDB databank searching for peptides composed of a short number n of residues ($20 \leq n \leq 30$), with no cysteine, and no unnatural/modified residues (see Supporting Information, Section d). From the hits, we selected commercially available peptides, presenting different secondary structures, and different functions. Magainin II (Table S1, Supporting Information) is an antimicrobial peptide, and glucagon (Table S1, Supporting Information) is a peptide hormone. Somatostatin 28 (Table S1, Supporting Information), a peptide hormone, was selected a posteriori because it is rich in proline (missing in magainin II and glucagon), and structurally different from the other two peptides, that is, disulfide cyclized. The three peptides together contain all 20 proteinogenic AAs (see **Figure 2**a–c for AAs contained in each peptide).

We depolymerized magainin II, glucagon, and somatostatin 28 by means of two consecutive enzymatic reactions, following the approach developed by Teixeira et al.[27] We incubated the peptides first with thermolysin endoprotease (that cleaves at the N-terminus of Leu, Phe, Val, Ile, Ala, Met), then with leucine aminopeptidase (LAP), as described in Supporting Information, Sections e and f. Mass spectrometric (MS) analysis of the materials before (Figures S14–S16, Supporting Information), and after thermolysin treatment (Figures S18–S31, Supporting Information) shows extensive cleavage at the N-terminus of the hydrophobic amino acids (see Supporting Information, Section j). Cleaved fragments were incubated with LAP and depolymerized to their free AAs (Figure 2a–c). For each AA we defined a depolymerization yield as the ratio between the amount of AAs produced by the depolymerization divided by the total amount of AAs present in the starting material (green and gray bars in Figure 2, respectively). Quantification was performed using MS (Supporting Information, Section k).

We achieved an average depolymerization yield of ≈66% ± 19%. The large standard deviation ($1\sigma = 19\%$) is caused by the large variation between depolymerization yields of different AAs, with a maximum of ≈99% for aspartic acid (for glucagon) and a minimum of ≈17% for phenylalanine (averaged for all three peptides). We observed variations in yield also across peptides, for example alanine was efficiently recovered from the depolymerization of magainin II and glucagon, but not from somatostatin 28. We noticed that the aromatic AAs were consistently recovered in poor yields (for all three peptides), and that such yields were dependent on the number (type) of aromatic residues in the material to be depolymerized. Specifically, the recovery of the aromatics in glucagon was higher (≈73% for Trp, ≈52% for Tyr, and ≈36% for Phe) than in somatostatin 28 (≈41% for Trp, and ≈15% for Phe), that was in turn higher than in magainin II (≈null for Phe). The free AAs achieved by depolymerizing separately the three peptides were combined, and added into the TX–TL system supplemented with an mScarlet-i DNA template (Table S2, Supporting Information, Section h). As shown in Figure 2d we successfully expressed mScarlet-i. As a reference control and yield reference, we ran a TX–TL reaction with a solution containing the concentration of each AA that would have been achieved had the depolymerization yield been 100% for each peptide (that ideal result of a complete depolymerization, Supporting Information, Section h).

A first attempt to determine the efficiency of NaCRe was performed by comparing the fluorescence values of the expression plateau for the recycling curve with that for the reference control (the green and gray curves in Figure 2d, respectively),

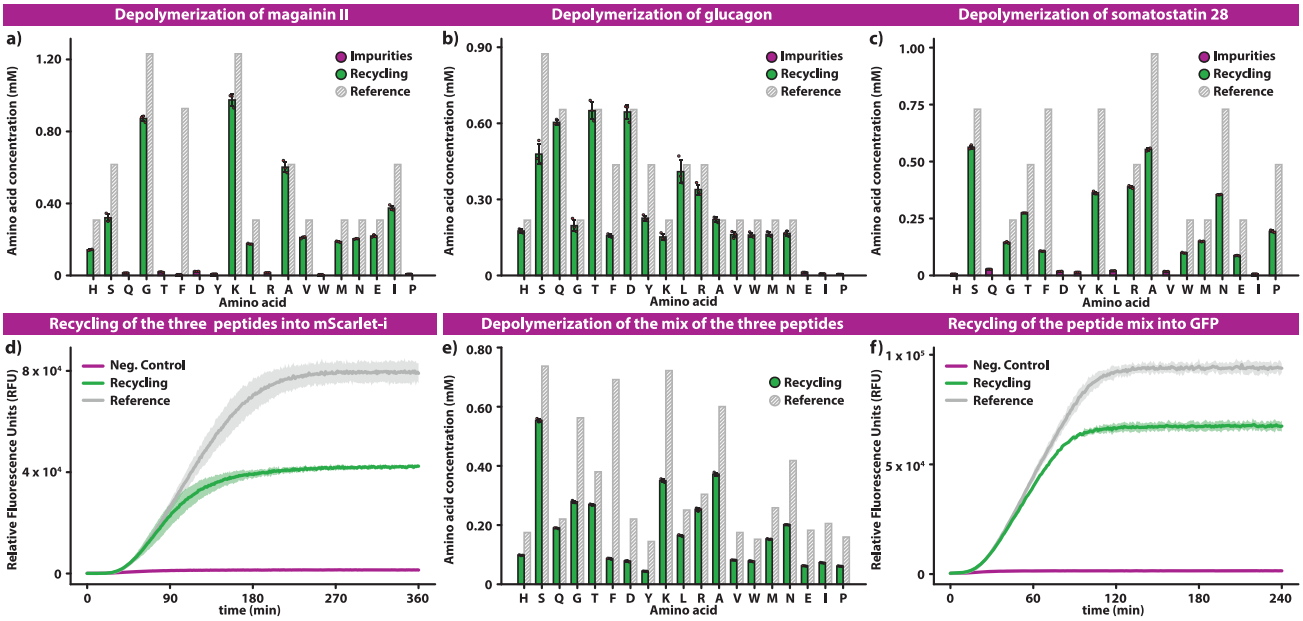

Figure 2. Recycling of magainin II, glucagon, and somatostatin 28 into mScarlet-i. a–c,e) Bar graphs showing the result of the amino acid analysis performed using mass spectrometry of the result of the depolymerization of magainin II (a), glucagon (b), and somatostatin 28 (c), and their mixture (e). The experimental results are represented with green bars to be compared with the gray bars that are the ideal reference concentrations of each AA calculated by assuming the complete conversion of the starting peptide into free AAs. The violet bars represent trace concentration of the AAs that theoretically should have not been observed, they are possibly the result of depolymerization of the digestion enzymes themselves. Such impurities are present for all the recovered AAs. The additive effect due to the impurities is by definition difficult to estimate, and probably contributes to slightly overestimate the green bars. This becomes more evident when the obtained depolymerization yield is close to 100%. (Note: cysteine is not detected by the amino acid analysis, hence the quantification of cysteine is n.a.). d,f) Plots of the fluorescence signal resulting from the expression of mScarlet-i (d) and GFP (f) in a TX–TL reaction. The green curves are data obtained preforming NaCRe, the gray curves are obtained as the results of expression experiments with the TX–TL reactions supplemented with concentrations for each AA matching the gray bars shown in a), b), c), and e). In the negative control expressions (violet curves), the TX–TL system was supplemented with the solution resulting from the same depolymerization process used for the individual peptides, without adding the peptides initially. Bar-plots of the statistical mean of the results of the repeated injections (triplicates) of each sample are shown; the error bars represent the standard deviation of the same data. The TX–TL reactions were all run in duplicates. The expression curves represent the statistical mean of the results at any acquisition time; the shadow represents the standard deviation of the same data.

leading to a yield of ≈50%. We also used NaCRe to express GFP (see Table S5, Supporting Information). In this case we spiked cysteine into the free AAs solution obtained from the depolymerization of magainin II, glucagon, and somatostatin 28, the resulting yield for GFP was ≈80% (Figure S2, Supporting Information). The results presented so far were achieved performing the depolymerization of each peptide separately, and by combining the obtained solutions at the end of the depolymerization process. In order to establish NaCRe as a recycling method that starts from mixtures of proteins and/or peptides, we also performed it starting with a mixture of the three peptides, depolymerizing them together, and expressing GFP. As shown in Figure 2e,f, the process was successful in depolymerization and expression, leading to a yield of ≈70% that is approximately the same yield we obtained when expressing GFP starting from the product of the separate depolymerization of the peptides.

It would be obvious at this point to wonder about the difference in observed yields for the expression of mScarlet-i and GFP. First, the yields mentioned so far are relative yields (RY), defined as:

$$\frac{P_1}{P_2} \times 100 \tag{1}$$

where P_1 and P_2 are the fluorescence intensity signal for the NaCRe (P_1) and the reference (P_2) expressed proteins, averaged over the last 30 min of the experiment.

The evaluation of a yield for NaCRe is rather complex because of the sequence-defined nature of the product. In fact, when expressing a protein from a mixture of free AAs there will always be a limiting reactant. This limiting AA will be the one that determines the amount of protein expressed in the reference control. By virtue of this definition, the limiting AA depends both on the proteins/peptides that were depolymerized as well as on the specific sequence of the protein to be expressed. As shown in **Table 1**, when recycling the three peptides, the limiting AAs for expressing the reference mScarlet-i is either proline, tyrosine, or valine, while for GFP it is valine. Note that the limiting AA does not necessarily need to be the AA with the lowest concentration in the reference reactant mixture, indeed in our case this was tyrosine. Also, the concentration of cysteine is irrelevant when expressing mScarlet-i because it lacks cysteine. Therefore, the yield of NaCRe can be tailored by enriching the mixture of proteins to be depolymerized with proteins/protein-based materials that contain the residues that are highly used in the sequence of the protein to be expressed. When determining the RY we make the implicit assumption that the limiting AA in NaCRe and in the reference control is the same. As shown in Table 1, this is not

Table 1. Overview of the depolymerization and expressions efficiencies for key experiments in this study; minima are formatted in bold (depolymerization) and bold italics (expressions).

Depolymerization (AAs)				Expression (protein chains)[a]				
[nmol] of AAs in 10 µL				[nmol] of protein chains in 25 µL TX–TL system supplemented with 10 µL of AAs from depolymerization				
AAs	Reference AAs	AAs	AAs	Reference (mScarlet-i)	NaCRe (mScarlet-i)	Reference (GFP)	NaCRe (GFP)	NaCRe (GFP)
		From separated peptides	From peptides mix	From separated peptides			From separated peptides	From peptides mix
(Type)	(Ideal)	(Experimental)	(Experimental)	(Ideal)	(Experimental)	(Ideal)	(Experimental)	(Experimental)
H	1.76	1.11 ± 0.04	1.00 ± 0.01	0.16	0.10 ± 0.01	0.12	***0.07 ± 0.01***	0.07 ± 0.01
S	7.41	4.57 ± 0.27	5.43 ± 0.06	0.44	0.27 ± 0.02	0.74	0.46 ± 0.03	0.54 ± 0.01
Q	2.19	2.17 ± 0.04	1.90 ± 0.01	0.27	0.27 ± 0.01	0.27	0.27 ± 0.01	0.24 ± 0.01
G	5.67	4.05 ± 0.18	2.80 ± 0.01	0.19	0.14 ± 0.01	0.26	0.18 ± 0.01	0.13 ± 0.01
T	3.81	3.14 ± 0.14	2.70 ± 0.01	0.22	0.19 ± 0.01	0.24	0.20 ± 0.01	0.17 ± 0.01
F	6.98	0.88 ± 0.02	0.87 ± 0.06	0.63	0.08 ± 0.01	0.58	***0.07 ± 0.01***	0.07 ± 0.01
D	2.19	2.29 ± 0.13	0.83 ± 0.06	0.14	0.14 ± 0.01	0.12	0.13 ± 0.01	0.05 ± 0.01
Y	**1.46**	0.83 ± 0.04	**0.40 ± 0.01**	***0.12***	0.07 ± 0.01	0.13	0.08 ± 0.01	***0.04 ± 0.01***
K	7.29	5.00 ± 0.21	3.50 ± 0.11	0.33	0.23 ± 0.01	0.36	0.25 ± 0.01	0.18 ± 0.01
L	2.49	2.02 ± 0.20	1.63 ± 0.06	0.18	0.14 ± 0.02	0.12	0.10 ± 0.01	0.08 ± 0.01
R	3.08	2.50 ± 0.11	2.53 ± 0.06	0.22	0.18 ± 0.01	0.51	0.42 ± 0.02	0.42 ± 0.01
A	6.03	4.61 ± 0.17	3.73 ± 0.06	0.43	0.33 ± 0.02	0.75	0.58 ± 0.03	0.47 ± 0.01
V	1.76	1.31 ± 0.07	0.80 ± 0.01	***0.12***	0.09 ± 0.01	***0.10***	***0.07 ± 0.01***	***0.04 ± 0.01***
W	1.54	0.87 ± 0.04	0.80 ± 0.01	0.51	0.29 ± 0.02	1.54	0.87 ± 0.04	0.80 ± 0.01
M	2.57	1.66 ± 0.06	1.50 ± 0.01	0.23	0.15 ± 0.01	0.43	0.28 ± 0.01	0.25 ± 0.01
N	4.19	2.41 ±0.08	2.00 ± 0.01	0.70	0.40 ± 0.02	0.32	0.19 ± 0.01	0.15 ± 0.01
E	1.84	1.07 ± 0.04	0.60 ± 0.01	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]
I	2.06	1.32 ± 0.06	0.70 ± 0.01	0.26	0.17 ± 0.01	0.17	0.11 ± 0.01	0.06 ± 0.01
P	1.62	**0.68 ± 0.02**	0.60 ± 0.01	***0.12***	***0.05 ± 0.01***	0.16	***0.07 ± 0.01***	0.06 ± 0.01

[a]Calculated as the ratio between the amount of each AA (nmol) and the number of its incorporations inside a single protein chain, see Table S5, Supporting Information; [b]n.a. not assessable because E is present in the TX–TL system as Potassium glutamate (buffer), see Supporting Information, Section h.

necessarily always the case. Therefore, even though the RY is a simple measure of the efficiency of our process, it depends critically on the starting and final proteins, hence it is a powerful tool solely to compare and optimize the yield of NaCRe when starting and ending from and into the same proteins.

The true efficiency of NaCRe should be its absolute yield (AY) defined as a mass-to-mass ratio of the output divided by the input. When the limiting AA is the same for the NaCRe and reference control, the AY can be written as:

$$RY \times Y \tag{2}$$

where Y is the yield of expression of the TX–TL system. AY (mass-to-mass ratio) in the case of the expression of mScarlet-i is ≈7% (see Supporting Information, Section n). The present results show a mass-to-mass yield for NaCRe for the limiting AA of proline in the expression of m-Scarlet-i of ≈15%. This is the most accurate measurement of the absolute yield of the process.

To go beyond peptides, we performed NaCRe starting from larger proteins with defined tertiary structures. We started by recycling β-lactoglobulin A (≈18 kDa, Table S1, Supporting Information), a protein that can be obtained in large quantities as a side product of bovine milk production. As shown in **Figure 3a**, β-lactoglobulin A was successfully depolymerized into its constitutive AAs with a yield comparable to the ones obtained for the peptides (see Supporting Information, Sections e and f). These AAs were used to express GFP (Figure S3, Supporting Information, Section h). The RY for β-lactoglobulin A recycled into GFP was ≈40%.

To better establish the potential of NaCRe we recycled technologically relevant materials. We first recycled a film composed of β-lactoglobulin amyloid fibrils, known to be able to adsorb a variety of different heavy metal ions with outstanding efficiency.[24] Such amyloids are assemblies of peptides obtained from the hydrolysis of β-lactoglobulin chains (A and B) and their re-assembly into filamentous proteins with a typical cross-β secondary structure. Because amyloids have been postulated to be the ground state in the protein folding landscape,[28] carrying out NaCRe starting from these systems ideally showcase the universality and the reach of the method. A solution of amyloid fibrils was dried on a cellulose membrane, as shown in Figure 3b (see Supporting Information, Section c). The dry film

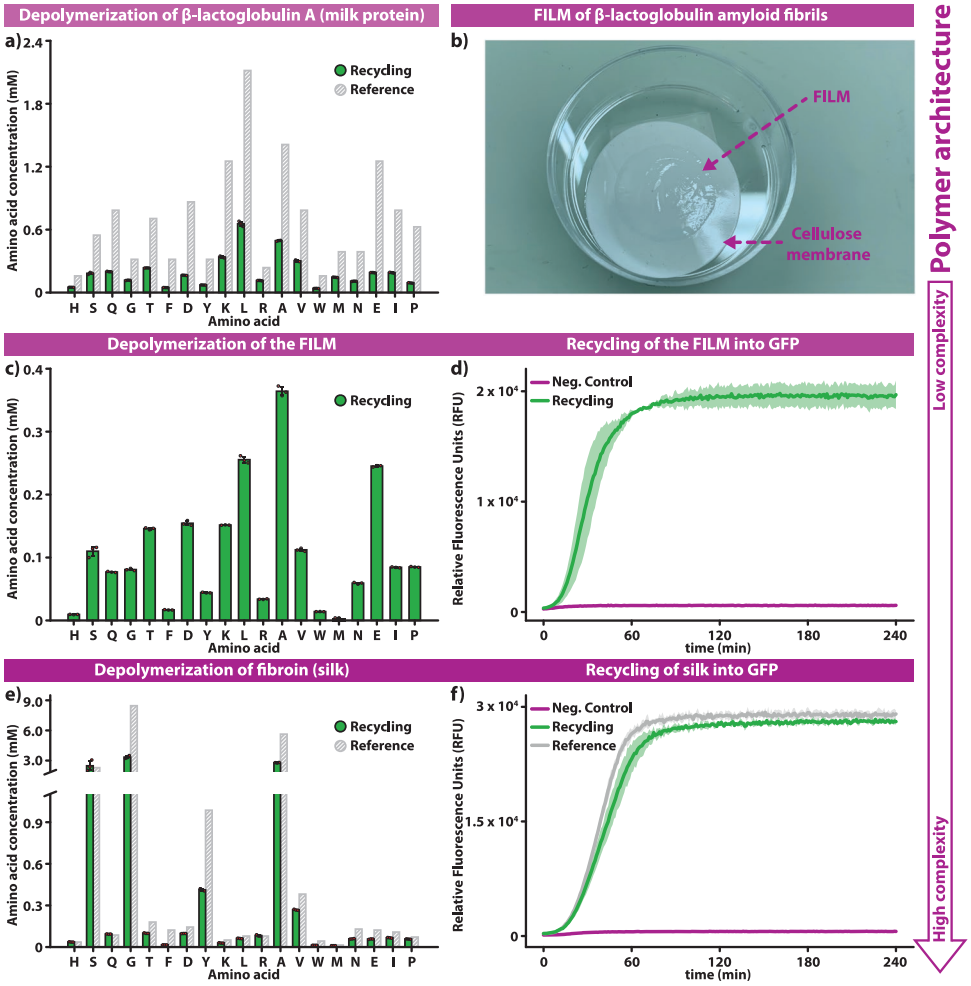

Figure 3. Recycling of technologically relevant materials. a,c,e) Bar graphs of the results of the amino acid analysis for the depolymerization of *β*-lactoglobulin A (a), *β*-lactoglobulin amyloid film (c), and silk fibroin solution (e). The color scheme (and its meaning) is identical to the one used in Figure 2, the ideal reference (gray bar) is missing from (c) because the exact composition of the amyloids composing the film is unknown. b) Photograph of a film composed of *β*-lactoglobulin amyloid fibrils. d,f) Real-time plots of the fluorescence signal resulting from the expression of GFP from the depolymerization of the amyloid film (d) and the silk fibroin solution (f); color scheme as in Figure 2.

was removed from the support, the film powder was weighed, and first incubated with pepsin endoprotease (that cleaves at the C-terminus of Leu, Phe, Tyr, Trp), then with LAP (Supporting Information, Sections e and f, respectively). In order to support the mass spectrometry evaluation of the depolymerization process, we performed atomic force microscopy analysis of the amyloid fibrils as prepared, and after full depolymerization. The images of the as prepared amyloids show an abundance of fibrils, that were absent after depolymerization (Figure S54, Supporting Information). The mass spectrometry result of the consecutive cleavage, and depolymerization is shown in Figure 3c. In this case we do not have a reference standard, as the exact amyloid composition is unknown due to the hydrolysis process. We note that methionine and histidine were obtained only at low concentrations. As shown in Figure 3d, the free AAs obtained from the *β*-lactoglobulin film were recycled into GFP, by spiking cysteine, methionine, and histidine (see Supporting Information, Section h). We then recycled a solution of silk fibroin (Table S1, Supporting Information), another technologically relevant protein used in many devices, ranging from biomedical[29] to electronics applications.[22] After incubating fibroin with thermolysin, and then LAP (see Supporting Information, Sections e and f, respectively), we successfully recovered fibroin's free AAs (Figure 3e), and used them to express GFP in our TX–TL system (Figure 3f) spiked with cysteine, and methionine (see Supporting Information, Section h). RY for silk fibroin recycling into GFP was ≈95%. Figure 3d,f demonstrate that NaCRe is capable of recycling high molecular weight polymeric structures, either composed of the supramolecular assembly of low molecular weight peptides or characterized by multiple high molecular weight chains.

As described above, we decided to spike cysteine every time we were expressing GFP because we could not detect cysteine, that is, quantify it, in the AAs solutions from the depolymerizations. We then tried to assess if cysteine could be part of NaCRe by recycling magainin II, glucagon, and somatostatin 28 into GFP, without adding any cysteine (Supporting Information, Section h). As shown in Figure S8, Supporting Infor-

mation, spiking cysteine was not necessary, since the two recycling curves reach basically the same plateau, this means that cysteine from the disulfide cyclization of somatostatin 28 is recycled into GFP. This result strengthens the visionary idea of NaCRe, where materials are recycled into completely different ones, without the need of any external monomer feed, that is fulfilling the principles of a circular-economy model for polymers. After proving that cysteine can be recycled by NaCRe (as well as the other AAs), we performed every experiment without the need of spiking any amino acid. We produced a mixture of low and high molecular weight proteins (glucagon, β-lactoglobulin A, and silk fibroin), and we successfully recycled it into GFP, as shown in Figure S9, Supporting Information. RY for recycling this mixture of proteins into GFP was ≈70%.

In order to show that NaCRe can undergo more than one complete cycle, we first scaled-up the NaCRe processes described just above to produce either GFP or mScarlet-i. We purified these proteins (see Figure S57, Supporting Information), and characterized them by proteomic analysis (see Supporting Information, Section m). For GFP we identified 24 exclusive unique peptides (55 exclusive unique spectra), with 87% sequence coverage. For mScarlet-i we identified 21 exclusive unique peptides (48 exclusive unique spectra), with 77% sequence coverage. We then performed a second NaCRe cycle on the purified GFP (≈0.1 mg) to produce mScarlet-i (Figure S10, Supporting Information), without the need of any spike AAs (see Supporting Information, Section i). After performing NaCRe starting from the mixture of low and high molecular weight proteins, we applied the same strategy to recycle a very complex mixture of proteins, that is our whole TX–TL system. As shown in Figure S11, Supporting Information, we successfully recycled into mScarlet-i the whole solution resulting from a first cycle of NaCRe in which glucagon, β-lactoglobulin A, and silk fibroin were recycled into GFP (see Supporting Information, Section i). This experiment demonstrates the robustness of NaCRe that can perform multiple cycles of recycling for truly complex protein mixtures, in the presence of other polymers such as nucleic acids.

Starting from the same mixture of glucagon, β-lactoglobulin A, and silk fibroin, we have also performed NaCRe to obtain catechol 2,3-dioxygenase (see Table S5, Supporting Information), an enzyme which converts catechol into 2-hydroxymuconate semialdehyde.[30] Figure S12, Supporting Information, shows that the product of NaCRe is indeed catalytically active.

After having shown that NaCRe is capable of recycling a variety of structurally different proteins, and protein-based materials, we demonstrated that NaCRe is not limited to the functionalities present in the 20 proteinogenic AAs. Thus, we recycled 2 unnatural amino acids (UAAs, L-norleucine, and L-canavanine) originating from a peptide containing several UAAs (see Table S1, Supporting Information), some present as DL-stereoisomers (3-fluoro-DL-valine and DL-3-hydroxynorvaline). The non-natural peptide was incubated first with thermolysin, then with LAP, as described in Supporting Information, Sections e and f. MS analysis before (Figure S17, Supporting Information), and after thermolysin incubation (Figures S32–S35, Supporting Information, Section j) shows extensive cleavage. After depolymerization with LAP, we identified all the residues composing the non-natural peptide (Figures S36–S41, Supporting Information, Section j). L-norleucine and L-canavanine were successfully recycled into GFP (Figures S42–S53, Supporting Information, Sections h and m), following the protocol developed in refs. [31,32]. The final product of this approach is a SDP composed of a set of monomers that goes beyond the 20 proteinogenic AAs. It should be noted that the GFP produced in this way is not fluorescent (Figure S13, Supporting Information). If one wanted to obtain from NaCRe proteins with their full set of biological properties then NaCRe should be based solely on the 20 proteinogenic AAs.

3. Conclusion

The results presented show that it is possible to envision a way of recycling protein-based materials, outside living organisms, where mixtures are transformed into a single targeted final protein. The recycling of the β-lactoglobulin film, and of silk fibroin into GFP can be used as a proof-of-concept for the generality of NaCRe, showing its potential to recycle materials composed of complex molecular architectures. Particularly noteworthy is that such a process can be carried out also from protein templates with extremely robust secondary structures, as in the case of amyloids. There are many challenges ahead for NaCRe, clearly the first one is the upscaling of the process in an economically viable way. This will require optimization of all the processes in terms of reactants and reaction conditions. An upscaled process will allow for the opportunity of identifying proteins that are particularly suitable for NaCRe and consequently for the focusing of NaCRe toward true technological opportunities.

The NaCRe is a complex yet powerful way to think about recycling, were mixtures of (soft) materials are transformed into new (soft) materials, not (necessarily) related to the parent ones. NaCRe has two key requirements, the materials must be sequence-defined macromolecules (not strictly necessary for the first NaCRe cycle) based on links that can be readily depolymerized, and the polymerization reaction must be based on an approach that can use random mixtures of monomers as starting materials (this is what the ribosome does exceptionally well). Many efforts have recently focused on developing increasingly complex synthetic SDPs[33,34] for applications ranging from data storage[35–37] to catalysis.[38,39] It is true that these SDPs are all synthesized with approaches that are incompatible with NaCRe as they are based on step-growth synthesis and require the separation of the starting monomers. It should be noted though that significant efforts exist in creating a synthetic equivalent of the ribosome,[40,41] that is a system capable of synthesizing SDPs starting from mixtures of monomers. With this work we hope to highlight an additional advantage of SDPs, their amazing ability to be recycled in ways that fulfill the vision of a circular economy.

4. Experimental Section

Comprehensive information about the materials and procedures is detailed in the Supporting Information.

Supporting Information

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

Acknowledgements

The research was supported by the Swiss National Science Foundation (SNF) NCCR Bio-Inspired Materials 51NF40-182881, the SNF Spark project CRSK-2_190167, and the ERC Advanced Grant (884114-NaCRe). A.M.S. received funding by the French National Research Agency (ANR) through the Programme d'Investissement d'Avenir under contract 17-EURE-0016. The authors would like to thank the Protein production and structure Core Facility (EPFL), and the Proteomics Core Facility (EPFL) for helping to express the calibrants, and to perform the proteomic analyses, respectively. The authors thank Prof. Correia, Dr. Laohakunakorn, Dr. Swank, Dr. Olson, Dr. Istomin, Dr. Michielin, and Matteo Gasbarri for useful discussions.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

F.S. conceived the work first. S.G., F.S., and S.J.M. designed the experiments and discussed the results. S.G., A.M.S., L.R.J., V.S., and S.C. performed the experiments under the supervision of F.S. and S.J.M. S.G., D.O., and L.M. performed the MS characterizations, analyzed the data, and discussed the results. A.M. performed the AFM characterization. L.P. adapted the software for analyzing the MS data. A.M. and S.G. wrote the script for selecting the peptides. S.B. and S.G. prepared the amyloid film under the supervision of R.M. All the authors contributed to the writing of the paper.

Data Availability Statement

The data supporting the findings of this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.5167184. The raw data for the mass spectrometry quantifications of amino acids are available from the corresponding author, upon reasonable request.

Keywords

protein-based materials, recycling, sequence-defined polymers, sustainability

Received: June 15, 2021
Revised: July 29, 2021
Published online:

[1] Our World in Data, https://ourworldindata.org/world-population-growth (accessed: September 2020).
[2] OECD Environment Policy Papers, https://read.oecd-ilibrary.org/environment/improving-plastics-management_c5f7c448-en (accessed: September 2020).
[3] The New Plastics Economy, https://www.ellenmacarthurfoundation.org/assets/downloads/The-New-Plastics-Economy-Rethinking-the-Future-of-Plastics.pdf (accessed: September 2020).
[4] M. Hong, E. Y.-X. Chen, *Green Chem.* **2017**, *19*, 3692.
[5] A. Rahimi, J. M. García, *Nat. Rev. Chem.* **2017**, *1*, 0046.
[6] G. W. Coates, Y. D. Y. L. Getzler, *Nat. Rev. Mater.* **2020**, *5*, 501.
[7] J.-B. Zhu, E. M. Watson, J. Tang, E. Y.-X. Chen, *Science* **2018**, *360*, 398.
[8] V. Tournier, C. M. Topham, A. Gilles, B. David, C. Folgoas, E. Moya-Leclair, E. Kamionka, M.-L. Desrousseaux, H. Texier, S. Gavalda, M. Cot, E. Guémard, M. Dalibey, J. Nomme, G. Cioci, S. Barbe, M. Chateau, I. André, S. Duquesne, A. Marty, *Nature* **2020**, *580*, 216.
[9] J.-B. Zhu, E. Y.-X. Chen, *Angew. Chem., Int. Ed.* **2018**, *57*, 12558.
[10] G. W. Fahnhorst, T. R. Hoye, *ACS Macro Lett.* **2018**, *7*, 143.
[11] X. Tang, E. Y.-X. Chen, *Chem* **2019**, *5*, 284.
[12] I. Vollmer, M. J. F. Jenks, M. C. P. Roelands, R. J. White, T. Harmelen, P. Wild, G. P. Laan, F. Meirer, J. T. F. Keurentjes, B. M. Weckhuysen, *Angew. Chem., Int. Ed.* **2020**, *59*, 15402.
[13] X. Jia, C. Qin, T. Friedberger, Z. Guan, Z. Huang, *Sci. Adv.* **2016**, *2*, e1501591.
[14] F. Zhang, M. Zeng, R. D. Yappert, J. Sun, Y.-H. Lee, A. M. LaPointe, B. Peters, M. M. Abu-Omar, S. L. Scott, *Science* **2020**, *370*, 437.
[15] G. O. Jones, A. Yuen, R. J. Wojtecki, J. L. Hedrick, J. M. García, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 7722.
[16] D. E. Fagnani, J. L. Tami, G. Copley, M. N. Clemons, Y. D. Y. L. Getzler, A. J. McNeil, *ACS Macro Lett.* **2021**, *10*, 41.
[17] R. A. Gross, B. Kalra, *Science* **2002**, *297*, 803.
[18] T. Iwata, *Angew. Chem., Int. Ed.* **2015**, *54*, 3210.
[19] M. Harvey, S. Pilgrim, *Food Policy* **2011**, *36*, S40.
[20] H. Sardon, A. P. Dove, *Science* **2018**, *360*, 380.
[21] J.-F. Lutz, M. Ouchi, D. R. Liu, M. Sawamoto, *Science* **2013**, *341*, 1238149.
[22] H. Tao, D. L. Kaplan, F. G. Omenetto, *Adv. Mater.* **2012**, *24*, 2824.
[23] T. P. J. Knowles, R. Mezzenga, *Adv. Mater.* **2016**, *28*, 6546.
[24] S. Bolisetty, R. Mezzenga, *Nat. Nanotechnol.* **2016**, *11*, 365.
[25] D. L Barreiro, Z. Martín-Moldes, J. Yeo, S. Shen, M. J. Hawker, F. J. Martin-Martinez, D. L. Kaplan, M. J. Buehler, *Adv. Mater.* **2019**, *31*, 1904720.
[26] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, *Nat. Biotechnol.* **2001**, *19*, 751.
[27] P. F. Teixeira, B. Kmiec, R. M. M. Branca, M. W. Murcha, A. Byzia, A. Ivanova, J. Whelan, M. Drag, J. Lehtiö, E. Glaser, *Nat. Chem. Biol.* **2016**, *13*, 15.
[28] P. C. Ke, R. Zhou, L. C. Serpell, R. Riek, T. P. J. Knowles, H. A. Lashuel, E. Gazit, I. W. Hamley, T. P. Davis, M. Fändrich, D. E. Otzen, M. R. Chapman, C. M. Dobson, D. S. Eisenberg, R. Mezzenga, *Chem. Soc. Rev.* **2020**, *49*, 5473.
[29] F. G. Omenetto, D. L. Kaplan, *Science* **2010**, *329*, 528.
[30] M. Verosloff, J. Chappell, K. L. Perry, J. R. Thompson, J. B. Lucks, *ACS Synth. Biol.* **2019**, *8*, 902.
[31] K. Josephson, M. C. T. Hartman, J. W. Szostak, *J. Am. Chem. Soc.* **2005**, *127*, 11727.
[32] M. C. T. Hartman, K. Josephson, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 4356.
[33] B. Lewandowski, G. De Bo, J. W. Ward, M. Papmeyer, S. Kuschel, M. J. Aldegunde, P. M. E. Gramlich, D. Heckmann, S. M. Goldup, D. M. D'Souza, A. E. Fernandes, D. A. Leigh, *Science* **2013**, *339*, 189.
[34] J. Niu, R. Hili, D. R. Liu, *Nat. Chem.* **2013**, *5*, 282.
[35] R. K. Roy, A. Meszynska, C. Laure, L. Charles, C. Verchin, J.-F. Lutz, *Nat. Commun.* **2015**, *6*, 7237.
[36] R. Dong, R. Liu, P. R. J. Gaffney, M. Schaepertoens, P. Marchetti, C. M. Williams, R. Chen, A. G. Livingston, *Nat. Chem.* **2019**, *11*, 136.
[37] N. Zydziak, W. Konrad, F. Feist, S. Afonin, S. Weidner, C. Barner-Kowollik, *Nat. Commun.* **2016**, *7*, 13672.
[38] M. A. R. Meier, C. Barner-Kowollik, *Adv. Mater.* **2019**, *31*, 1806027.
[39] J. C. Barnes, D. J. C. Ehrlich, A. X. Gao, F. A. Leibfarth, Y. Jiang, E. Zhou, T. F. Jamison, J. A. Johnson, *Nat. Chem.* **2015**, *7*, 810.
[40] K. Wang, H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nat. Biotechnol.* **2007**, *25*, 770.
[41] J. W. Chin, *Nature* **2017**, *550*, 53.