CHEMISTRY

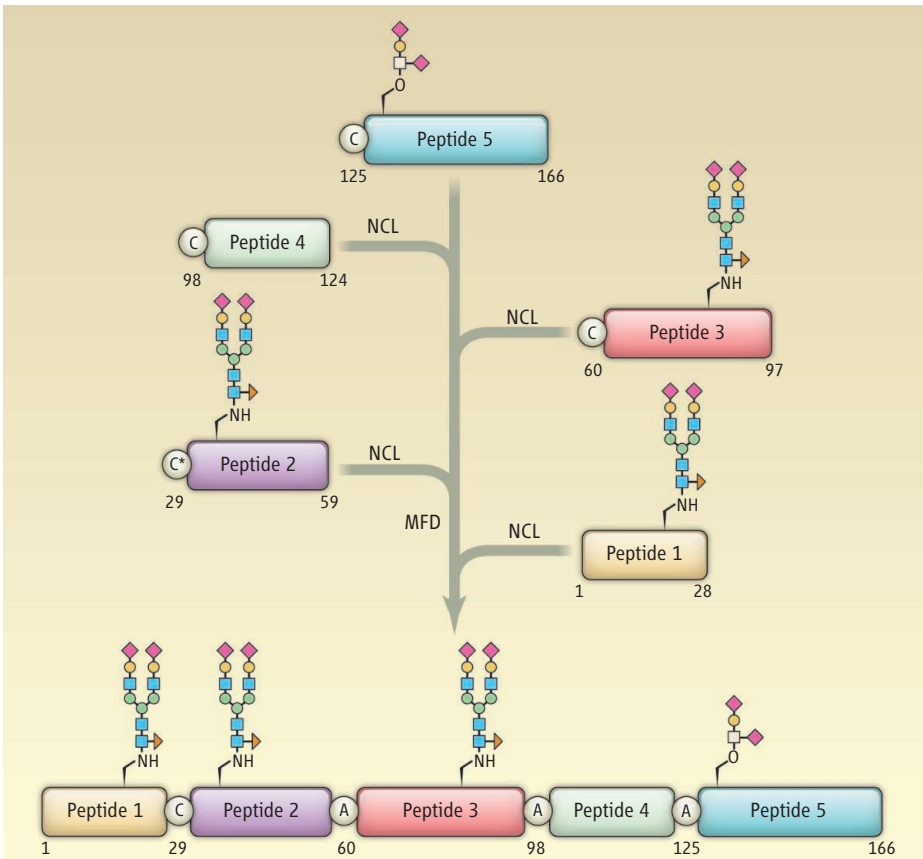
Improving Biologic Drugs via Total Chemical Synthesis

The chemical synthesis of a single glycosylated form of erythropoietin may help elucidate the functions of its sugars and tailor its properties.

Linda C. Hsieh-Wilson and Matthew E. Griffin

Most biologic therapeutics are large, complex molecules or heterogeneous mixtures of molecules that are manufactured in a living system (e.g., microorganism, plant cell, or animal cell) through recombinant DNA technology. Biologics are now being used to treat a wide range of diseases, including cancer, autoimmune disorders, and diabetes. It is estimated that half of the top 100 best-selling medications will soon be biologics, with Roche's anticancer biologic Avastin and AbbVie's anti-inflammatory antibody Humira rivaling the success of Pfizer's small-molecule drug Lipitor (*1*). Biologics differ fundamentally from small-molecule drugs in terms of purity, composition, and production. Small-molecule drugs typically have homogeneous, well-defined structures that have been finely tuned with atomic-level precision via chemical synthesis. On page 1357 of this issue, Wang *et al*. (*2*) bridge the gap between biologics and small-molecule drugs by accomplishing the total chemical synthesis of the biologic erythropoietin (EPO) in a single, pure form.

Erythropoietin is a 166–amino acid glycoprotein hormone that regulates the production of red blood cells (*3*). A major challenge in obtaining pure, homogeneous EPO is that cells produce the protein as a mixture of glycosylated forms (glycoforms). The protein undergoes extensive and variable glycosylation (the attachment of complex sugars) at four distinct sites (serine 126 and asparagine 24, 38, and 83). The commercial analogs of EPO, Procrit (Janssen), Epogen (Amgen), and Aranesp (Amgen), are sold as a mixture of glycoforms and are therapeutically used to treat anemia resulting from cancer, chemotherapy, or kidney failure. Studies suggest that the specific sugar modifications affect the stability and erythropoietic activity of EPO (*4*–*6*). Thus, systematic investigations into the structure-activity relations of the sugars and targeted production of the most active glycoform could improve the efficacy of EPO, as well as many other biologics. However, the individual therapeutic contributions of each glycoform are presently unknown because of an inability to precisely alter the sugar structures and isolate pure EPO glycoforms.



A homogeneous glycosylated erythropoietin. Wang *et al*. used native chemical ligation (NCL) to connect five separate peptides and form a single glycoform of the biologic drug EPO. An improved method for metal-free desulfurization (MFD) of unprotected cysteine residues (C) allows for ligation at more common alanine (A) junction sites between peptides. The application of convergent aspartylation to append complex *N*-glycans to these peptides before NCL permits the substitution of specific glycan structures to generate homogeneous glycoforms and study their individual activities.

Wang *et al*. tackle this problem by constructing a single, complex glycoform of EPO via chemical synthesis (see the figure). This tour-de-force effort was made possible by several recent advances in chemical protein synthesis. Early work by Kent and colleagues paved the way with the development of native chemical ligation (NCL), a mild, selective reaction that enables two unprotected polypeptide chains to be joined (*7*). Long polypeptides beyond the reach of solid-phase peptide synthesis (which is generally limited to ~50 residues) can be synthesized by NCL, but the process traditionally requires a cysteine residue at the ligation site between the peptides, and like many proteins, EPO lacks suitable cysteine residues at the desired junctions. To circumvent this problem, Wang *et al*. expanded on desulfurization methods to convert cysteine residues to alanine (*8*) and developed a mild, radical-based desulfurization procedure (*9*). This advance greatly increased the number of potential ligation sites, as alanine is far more prevalent than cysteine in proteins.

Still, these improvements only addressed half of the problem—the construction and attachment of glycans also posed a formidable challenge. The synthesis of complex

Division of Chemistry and Chemical Engineering and Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125, USA. E-mail: lhw@caltech.edu

Published by AAAS

glycans is notoriously difficult because of the presence of multiple hydroxyl groups of comparable chemical reactivity and the stereospecific connections between the sugar units. Unlike the synthesis of peptides and oligonucleotides, there are no universal building blocks or methods for the synthesis of all glycans; thus, each synthesis requires meticulous design and experimentation to produce the exact desired structure. *N*-Glycans like those found in EPO may contain multiple branches of up to 20 sugar units and are some of the most complex oligosaccharide structures known (*10*).

Once synthesized, the glycan must then be attached to the peptide, which is again no small feat. However, a technique known as convergent aspartylation elegantly accomplishes this task (*11*, *12*). This reaction combines the carboxylic acid of an aspartate residue and an amine-terminated oligosaccharide to attach the glycan to the peptide and simultaneously convert the aspartate to the native asparagine. To generate a single EPO glycoform of physiologically relevant complexity, Wang *et al*. synthesized and appended a "consensus" dodecasaccharide for the three *N*-glycans and a glycophorin tetrasaccharide for the single *O*-glycan found on EPO.

However, their synthetic route was not without pitfalls. When the researchers attempted to join peptides containing the full consensus *N*-glycans using a ligation strategy previously shown to work with smaller glycans, one of the ligation steps was not successful, presumably because of the greater steric bulk of the larger *N*-glycan. Fortunately, reconfiguration of the synthesis with a different junction site distant from the problematic *N*-glycan allowed for smooth ligation, and assembly of the remaining peptide fragments produced the desired EPO structure. The fully synthetic EPO glycoform exhibited enhanced stability relative to its counterpart with smaller glycans and, upon folding, displayed in vivo erythropoietic activity comparable to that of the clinical drug Procrit.

Further advances are still required for chemical synthesis to be able to compete with biologic methods in terms of speed, scalability, and production. Nonetheless, the chemical synthesis of such a complex molecule represents an impressive culmination of more than a decade of synthetic efforts by the Danishefsky laboratory. Their studies set the stage for generating "synthetic biologics" whose structures are controlled with the same atomic-level precision as small-molecule drugs. An understanding of how various, defined glycan structures affect the pharmacological properties of EPO and other biologics may ultimately contribute to the development of better and safer drugs.

References

1. B. Huggett, *Nat. Biotechnol.* **31**, 697 (2013).
2. P. Wang *et al*., *Science* **342**, 1357 (2013).
3. A. J. Sytkowski, *Erythropoietin*: *Blood, Brain and Beyond* (Wiley-VCH, Weinheim, Germany, 2004).
4. L. O. Narhi *et al*., *J. Biol. Chem.* **266**, 23022 (1991).
5. M. Mattio *et al*., *Biotechnol. Prog.* **27**, 1018 (2011).
6. M. Higuchi *et al*., *J. Biol. Chem.* **267**, 7703 (1992).
7. P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **266**, 776 (1994).
8. L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **123**, 526 (2001).
9. Q. Wan, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **46**, 9248 (2007).
10. P. H. Jensen, N. G. Karlsson, D. Kolarich, N. H. Packer, *Nat. Protoc.* **7**, 1299 (2012)
11. S. T. Cohen-Anisfeld, P. T. Lansbury Jr., *J. Am. Chem. Soc.* **115**, 10531 (1993).
12. P. Wang, B. Aussedat, Y. Vohra, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **51**, 11571 (2012).

10.1126/science.1247615

MOLECULAR BIOLOGY

Finding the Right Partner in a 3D Genome

The three-dimensional organization of the genome plays a role in controlling legitimate and illegitimate DNA recombination.

Pedro P. Rocha, Julie Chaumeil, Jane A. Skok

DNA integrity is frequently compromised as a result of exposure to cytotoxic agents, as well as the normal wear and tear of cellular processes like transcription and replication. DNA double-strand breaks (DSBs) are arguably the most dangerous type of DNA damage as they can lead to chromosomal translocations when their repair joins noncontiguous genomic regions together. Indeed, numerous malignancies have been associated with signature translocations in which an oncogene becomes deregulated through joining with another gene that exerts control over its expression.

By definition, illegitimate joining of broken DNA ends requires that the two partner genes are physically brought together, and in this context, frequent exchange partners have been found in closer spatial proximity in the nucleus than rare exchange partners (*1*), and to occupy the same polymerase II–enriched transcription factories (*2*). More recently, translocation capture data sets, combined with chromosome conformation capture (3C) (*3*), have further confirmed that nuclear organization has a major impact on the choice of translocation partners and that within a population of cells, most translocations occur between sites that are found most frequently in the same neighborhood (*4*). These conclusions are nonetheless limited to an end product that is the sum of the data from a population of formaldehyde-fixed cells.

Researchers are now starting to look at the architecture of the nucleus in real time. Using high-throughput time-lapse imaging in live cells containing DNA breaks in defined chromosomal locations marked by binding sites for fluorescent reporter proteins, Roukos *et al*. were able to track the formation of translocations after DSB induction (*5*). They found that translocations form within hours of a break after transitioning through three phases: DSB partner search, transient pairing, and persistent pairing. Breaks that result in a permanent fusion between two distant parts of the genome are more mobile than nontranslocating breaks. Curiously, the two ends of the same break move in unison during the break partner search and separate only after completion of a translocation. Ostensibly, this provides a mechanism to promote the correct rejoining of the two broken ends as opposed to illegitimate joining with a noncontiguous partner. The orchestrated movement of the two ends of a break also explains how reciprocal translocations can arise when unfaithful repair occurs between loci on different chromosomes. By tracking the cells over time, Roukos *et al*. could determine that most translocations arise from breaks in close proximity at the time of formation, but

Department of Pathology, New York University School of Medicine, 550 First Avenue, MSB 599, New York, NY 10016, USA. E-mail: jane.skok@nyumc.org

Improving Biologic Drugs via Total Chemical Synthesis

Linda C. Hsieh-Wilson, and Matthew E. Griffin

Science, 342 (6164), • DOI: 10.1126/science.1247615