

Editor's Summary

Practical Applications of Peptide Production

Biochemists, of course, cannot compete with ribosomes when it comes to speed. Whereas a bacterial (*Escherichia coli*) cell needs only a second or so to link 40 amino acids into a polypeptide chain, an automated peptide synthesizer requires a day or more to complete this task. The slower speed of the synthetic approach, however, is offset by several key advantages: the ability to include non-natural amino acids, if desired; the capacity to incorporate modifications in the backbone to confer resistance to proteases; and the increased safety and purity of synthetic peptides relative to proteins that are isolated from a pathogen of interest or produced from an expression vector, for example. Such benefits mean that synthetic peptides offer potential clinical advantages in the development of vaccines and drugs. Improvements in peptide synthesis technology that allow the production of long synthetic peptides (over 100 amino acids in length), together with bioinformatic techniques in which sequenced genomes are analyzed to determine appropriate peptides to synthesize, should both aid the synthetic peptide approach to clinical problems. Corradin and colleagues briefly review the history of synthetic peptide synthesis and discuss recent efforts to produce clinical products—including vaccines to prevent malaria, cancer, and allergies—based on long synthetic peptides.

A complete electronic version of this article and other services, including high-resolution figures, can be found at:

<http://stm.sciencemag.org/content/2/50/50rv3.full.html>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

Long Synthetic Peptides for the Production of Vaccines and Drugs: A Technological Platform Coming of Age

Giampietro Corradin,^{1*} Andrey V. Kajava,² Antonio Verdini¹

(Published 22 September 2010; Volume 2 Issue 50 50rv3)

Long synthetic peptides (LSPs) have a variety of important clinical uses as synthetic vaccines and drugs. Techniques for peptide synthesis were revolutionized in the 1960s and 1980s, after which efficient techniques for purification and characterization of the product were developed. These improved techniques allowed the stepwise synthesis of increasingly longer products at a faster rate, greater purity, and lower cost for clinical use. A synthetic peptide approach, coupled with bioinformatics analysis of genomes, can tremendously expand the search for clinically relevant products. In this Review, we discuss efforts to develop a malaria vaccine from LSPs, among other clinically directed work.

TECHNOLOGY FOR PRODUCING SYNTHETIC PEPTIDES

The ability to synthesize peptides using chemical techniques has numerous important applications, including synthetic vaccine and drug development. This technology has been honed for over a hundred years. Procedures of peptide bond formation were pioneered by Curtius and Fischer at the turn of the 19th century. A new era in peptide chemistry began in 1932 with the introduction of benzyloxycarbonyl for the temporary protection of the α -amino group (1), which culminated in 1954 with the synthesis of oxytocin, the first biologically active synthetic peptide (2). Perhaps the most spectacular application of the arsenal of new protecting groups and coupling methods incessantly developed during the 1960s was the synthesis, in 1963, of 39-residue porcine adrenocorticotrophic hormone (3).

Classically, peptides were synthesized with liquid-phase techniques. A major breakthrough in peptide chemistry came in 1963 with the description by Merrifield of the principles of solid-phase peptide chemistry (SPPS) and their application to the synthesis of a tetrapeptide (4). In this approach, peptide chains are synthesized while they are linked to resin. Synthesis occurs through repeated rounds of deprotection and coupling, proceeding from the C terminus to the N terminus; the retention of the growing chain on the resin allows reagents and by-products to be washed away. After synthesis, the peptide is released from the resin with a strong acid. Later, Sheppard very successfully modified the Merrifield's basic concept with alternative chemistry (5). A schematic outline of both methods—referred to as the butoxycarbonyl (Boc)-benzyl and 9-fluorenyl methoxycarbonyl (Fmoc)-*t*-butyl methods to indicate the protection groups used for α -amino and side-chain functional groups, respectively—is shown in Fig. 1.

The original Boc-benzyl method has been refined over the years to such an extent that HIV protease (99-amino acid residues) was assembled on a peptide synthesizer in 33 hours with a rate of synthesis of 75 residues per day (6). In previous peptide syntheses, rates of synthesis of about 5 to 10 residues per day have been reported depending on peptide sequences and synthetic protocols chosen. But the

hazards and difficulty in handling liquid hydrogen fluoride (HF; a volatile and highly toxic acid) during peptide-resin cleavage have hampered its widespread use in biomedical laboratories.

Fmoc-*t*-butyl chemistry is simpler and easier to use. It is also more flexible, requires fewer reagents, and involves the application of trifluoroacetic acid (TFA) and scavengers for peptide-resin cleavage and the removal of all side chain-protecting groups. An important feature of the method is that the removal of the Fmoc group by piperidine after each coupling step gives a piperidine-dibenzofulvene chromophoric adduct that allows monitoring of the individual coupling reactions by ultraviolet spectroscopy. This ability provides the operator with some assurance that the synthesis is progressing satisfactorily and with indications of further optimization in successive syntheses.

The spectacular achievements of SPPS led to a steady decline of the "classical" liquid-phase methods, whereas the solid-phase procedures were universally accepted for the assembly of peptides consisting of no more than 40 residues, a length considered until recently as an upper limit of this technology. However, in the 1990s, the substantial gains made in advancing the synthetic protocols as well as the methods of peptide purification and analysis increased this limit further (7). Indeed, long synthetic peptides (LSPs), well over 100-amino acid residues in length, have acquired increasing importance in biology and immunology. The main advantages of the solid-phase methods are their speed, the facile production of 10 to 100 mg of thoroughly purified products, and the possibility of including non-natural amino acids and backbone modifications to confer protease resistance and control of the enzymatic processing of the LSPs.

In the stepwise assembly of LSPs and small proteins (for example, chemokines), "difficult sequences" are frequently encountered. Such sequences almost invariably present problems in coupling and Fmoc removal steps, particularly at a position 5 to 15 residues from the C terminus, the result of aggregation of the growing peptide chains induced by the formation of β sheet structure. This problem has been satisfactorily solved by bringing about a structural change in the peptide backbone at one or more serine or threonine residues through the use of dimethylloxazolidinones or isoacyldipeptides and at glycine residues by the reversible protection of their amide NH groups in the peptide bond (7). Aspartimide formation at vulnerable Asp-Gly positions as a result of repeated exposure to piperidine can be avoided

¹Biochemistry Department, University of Lausanne, 1066 Epalinges, Switzerland. ²Centre de Recherche de Biochimie Macromoléculaire, CNRS, University of Montpellier, 34293 Montpellier, France.

*To whom correspondence should be addressed. E-mail: giampietro.corradin@unil.ch

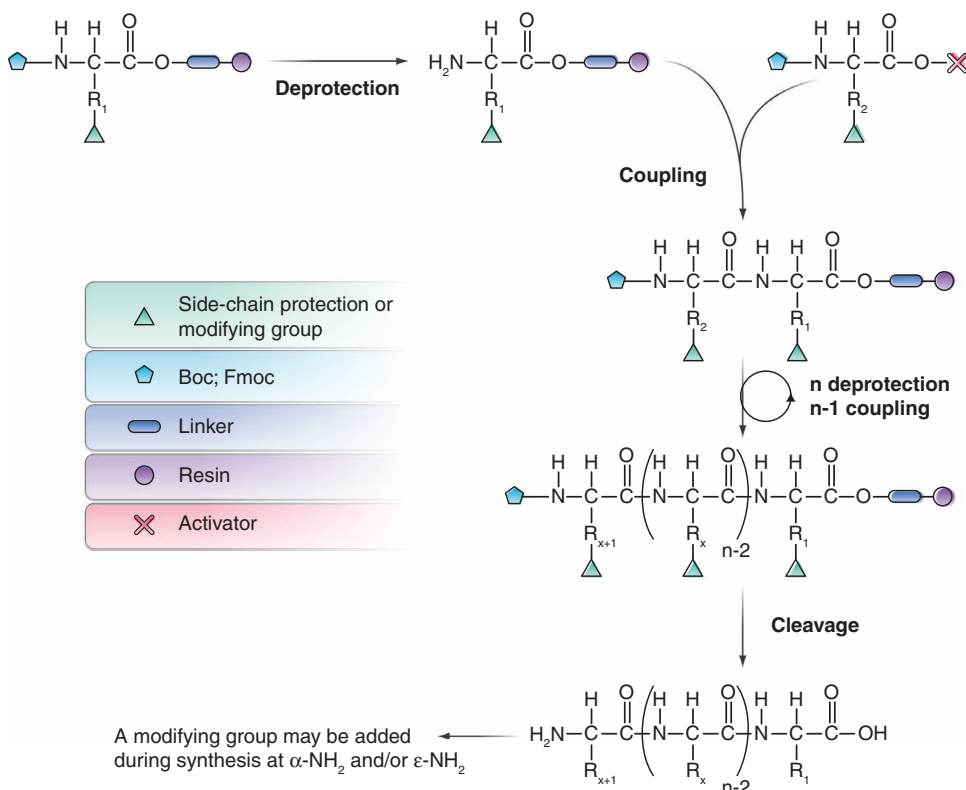


Fig. 1. General scheme of the Boc-benzyl and Fmoc-*t*-butyl methods of stepwise solid-phase synthesis of linear peptides. Deprotection of the α -amino-protecting group of the resin-linked growing peptide chain and coupling of the next protected amino acid constitute the repetitive cycle of the synthesis. X indicates the coupling reagent for amide bond formation. Both the α -amino group and the ϵ -amino group on lysine side chains can be modified for various purposes. The use of an insoluble support (resin) in a single reaction vessel allows for automation of the process. The crude peptide is obtained by resin-peptide acidolysis (liquid HF for Boc-benzyl and TFA for Fmoc-*t*-butyl methods) in the presence of scavengers. X: for example, diisopropylcarbodiimide-*N*-hydroxybenzotriazole; preformed symmetrical anhydrides; active esters; aminium salts such as HBTU [2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate], TBTU [2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate], BOP [benzotriazole-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate], PyBOP (benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate), or HCTU [2-(6-chloro-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate] in the presence of a tertiary base. [Credit: C. Bickel/*Science Translational Medicine*]

with Fmoc-Asp(OtBu)-(Hmb/Dmb)Gly derivatives (7). Poorly soluble LSPs have also been prepared by linking a polylysine-solubilizing tag to the insoluble peptide via a base-labile handle (8). Lipid, carbohydrate, phosphate, and chromophoric groups can be introduced at selected amino acid side chains of LSPs by following the Fmoc-*t*-butyl strategy.

PURIFICATION, ANALYSIS, AND FOLDING

The combined use of high-performance liquid chromatography (HPLC) and mass spectrometry [matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) and ion spray mass spectrometry] has revolutionized the purification and analysis of LSPs. Reversed-phase HPLC (RP-HPLC) is regarded as the most effective method for eliminating peptide impurities, particularly in combination with either ion exchange or

gel filtration chromatography. When the physicochemical properties of the impurities are very similar to those of the target product, a removable affinity tag introduced at the N terminus of the target peptide permits it to be isolated from the crude sample via a reaction (covalent capture) or the formation of a noncovalent bond with the stationary phase (9–13). Excellent resolution between biotin-labeled long peptides and truncated/alkylated impurities has been demonstrated by affinity chromatography on agarose coupled to the biotin-binding protein avidin (10). Poly-His tags, which bind to nickel (II), have also been successfully used in the purification of LSPs (11). The α -ketoacyl method [which is based on the selective binding of a keto-peptide to an aminoxy-functionalized resin followed by the transamination by glyoxylic acid and nickel (II) and the selective cleavage of the α -ketoacyl group with *o*-phenylenediamine], combined with RP-HPLC, has been shown to be a powerful purification strategy for the isolation of scarcely soluble peptides (13).

Finally, a new and simple yet very practical method has addressed a problem that occurs with LSPs containing multiple cysteine residues: the formation of non-native intramolecular and intermolecular disulfide bonds through oxidation of wrong pairs of thiol groups during the final oxidative folding. This method, based on polypeptide precursors S-protected as *t*-butylthio derivatives, eliminates the problems observed in the classical oxidative folding of polythiol precursors (14). Besides excellent yields of correctly folded peptides, a distinct practical advantage of the method is the

possibility of precisely evaluating the extent of folding and disulfide bond formation by MALDI-TOF mass spectrometry.

LSPS IN THE CLINIC: VACCINES AND DRUGS AGAINST INFECTIOUS AGENTS AND CANCER

To generate an immune response against disease-causing pathogens, vaccines generally contain killed or weakened forms of the microorganism, inactivated forms of toxins produced by the microorganism, or antigens produced by recombinant DNA technology. An alternative approach is to synthesize LSPs that resemble key regions of proteins on the pathogen's surface, for example. This synthetic approach is versatile, given that a wide variety of LSPs can be produced; it also offers speed of manufacturing, purity, and safety because the pathogen or impurities from the expression vector itself are not part of the vaccine. In addition, LSPs contain, in general, epitopes that are recognized

by helper (CD4⁺) and cytotoxic (CD8⁺) T cells and, in contrast to short peptides, may need to be processed and presented to T cells by professional antigen-presenting cells.

Malaria vaccines

We aimed to develop synthetic malaria vaccine candidates in the 1990s, an approach that was dictated by the advantages described above. It was also clear that a mixture of short peptides was not a sensible approach for a malaria vaccine because the possibility of mimicking conformational epitopes would be lost. In addition, small antigens (insulin, cytochrome, lysozyme, etc.) had been shown to be adequate to elicit an immune response in a genetically complex population.

Vaccines can be divided in two classes depending on the protective mechanism one wishes to obtain: either a T cell- or a B cell-based response. T cells minimally recognize 8- to 15-amino acid-long peptides derived from antigen degradation by antigen-presenting cells or serum proteases, followed by binding to major histocompatibility class I and/or II cell surface molecules and presentation to specific T cells, enabling antigen recognition by the T cells. Thus, no particular structural requirements of the antigen are required, facilitating vaccine manufacturing. On the other hand, in the case of vaccines that are designed to induce antibody secretion by B cells, the antigen must correctly mimic the three-dimensional structure of the parent protein to be effective. Unfortunately, the antigen structure is often not known.

Development of a malaria vaccine is complicated by the complex life cycle of the parasitic protists of the genus *Plasmodium* that cause this disease. *Plasmodium* sporozoites—the cell form that infects new hosts—are transmitted to hosts during the bite of an infected mosquito. The sporozoites then migrate to the liver where they multiply inside hepatocytes, eventually developing into merozoites, the form that infects erythrocytes. For vaccines against malaria, T cells and/or antibodies are the mediators of protection depending on the stage of the parasite to be targeted: the pre-erythrocytic stage (which includes the sporozoite) or erythrocytic stage. Sporozoite-specific antibodies can inhibit liver invasion by the free-living microorganism, whereas destruction of the parasite within hepatocytes is mediated by CD4⁺ and CD8⁺ T cells through secretion of lymphokines (that is, interferon- γ or tumor necrosis factor- α) upon recognition of infected hepatocytes. Thus, liver-stage antigens do not need to be conformationally similar to the native proteins. On the contrary, for the sporozoite and the erythrocytic stage, antibodies exert the protective function, which requires that synthetic antigens used for vaccination be conformationally similar to the native protein.

Following these general principles, the first antigen we selected was the circumsporozoite (CS) protein—a cell surface protein present in sporozoites and early liver stages of *Plasmodium falciparum*, the most deadly form of human malaria parasite. The gene encoding this protein was the first to be sequenced from *P. falciparum* (15) and has three distinct domains: the N- and C-terminal regions, which are both about 100 amino acids long, and a central repeat region consisting of about 40 copies of a tandemly repeated tetramer (Asn-Ala-Asn-Pro). We first concentrated on the C-terminal fragment because preclinical data suggested that it was important for protecting against liver infection through activation of CS-specific CD4⁺ and CD8⁺ T cells (16). Thus, the folded fragment, which contains four disulfide-bonded cysteine residues, was tested in two independent phase I clinical trials with different adjuvants (17, 18). Good immunological results were obtained:

Inoculation of the peptide together with an adjuvant, either Montanide ISA 720 (Seppic) or AS02A (GlaxoSmithKline), both oil-in-water emulsions that contain immunostimulants, elicited CD4⁺ and CD8⁺ T cell as well as antibody responses. In this instance, AS02A proved to be superior to Montanide ISA 720 in terms of antibody levels and production of interferon- γ in most volunteers (18). In view of these results, a phase IIa clinical trial was performed. Volunteers immunized with the antigen formulated with Montanide ISA 720 only, regrettably because of the unavailability of AS02A, were challenged by *P. falciparum* parasites through the bite of infected mosquitoes. No protection was obtained; this failure may be a result of the reduced response achieved with the adjuvant used (19).

Other clinical trials were then undertaken using LSPs with sequences derived from proteins of *P. falciparum* and *P. vivax*, another species that causes malaria in humans. These proteins included *P. vivax* CS protein (three individual or combined N, R, and C domains of 76, 48, and 71 amino acids, respectively), *P. falciparum* glutamic-rich protein (GLURP; 127 amino acids), which is expressed in several stages of the parasite life cycle, and merozoite surface protein 3 (MSP3; 96 amino acids) (7, 20). In particular, the MSP3-derived LSP underwent additional testing in phase Ib and IIa trials (21, 22) and then in a IIb clinical trial. Efficacy results from the latter trial will be available in the near future. GLURP has been further developed as a hybrid recombinant product in combination with MSP3 (23) and will enter phase Ib and II clinical trials in the near future.

Cancer vaccines

LSPs are also being used in the development of vaccines to induce T cell-based responses to prevent gynecological cancers. Several approaches involve vaccines against the human papillomavirus (HPV), because HPV infection can cause a number of cancers, including cervical cancer. In one case, longer than usual CD8⁺ T cell epitopes (30 to 35 amino acids) covering the entire HPV tumor antigen were synthesized to incorporate potential CD4⁺ T cell epitopes. Indeed, in a study of peptides representing *Plasmodium* sequences, the presence of helper T cell epitopes enhanced the CD8⁺ T cell response in an animal model (24). Two clinical trials involving LSPs representing HPV-16 E6 and E7 oncoproteins (which inactivate host tumor suppressor proteins) in one case and p53 protein (which is aberrantly expressed in many cancers) in the other were performed in vulvar intraepithelial neoplasia and ovarian cancer patients, respectively (25, 26). Results obtained were considered very promising (25, 27). Improvement of these formulations in terms of cost and, possibly, efficacy may be achieved by producing overlapping peptide segments up to 80 to 100 amino acids long.

Allergy vaccines

During an allergic reaction, the body interprets some substances—such as pollen or bee venom—as a disease agent and produces a type of antibody [immunoglobulin E (IgE)] that recognizes that substance. The IgE binds to basophils and mast cells; when the allergen is next encountered, these cells cause an inflammatory response that produces allergy symptoms. Allergen immunotherapy generally involves vaccinating the patient with increasing amounts of the allergen to induce tolerance to it. In the case of bee venom or pollen allergy treatment, it was hypothesized that injection of soluble denatured antigen would tolerate or anergize (make unresponsive) allergen-specific T cells without the antigen's concomitant recognition by IgE and subsequent activation of basophil

mast cells. This procedure was likely to avoid adverse events during the desensitization treatment and would take less time (weeks instead of years) than conventional approaches involving the intact allergen. Thus, three

long, carefully selected overlapping fragments of bee venom phospholipase A were synthesized. The LSPs were tested in vitro to determine any residual binding to IgE obtained from bee venom-allergic patients and

then injected in patients, where they were found to reduce the antigen-specific T cell response and IgE serum concentration (28). A similar approach was recently applied to prevent allergic reaction to the pollen Btv1 protein (29).

Chemokines

LSPs are also extensively used in research on chemokines, as well as in their clinical applications, and a variety of synthetic chemokines, either in native or in modified form, are commercially available. Chemokines are secreted proteins of 70 to 100 amino acids that contain four or six disulfide-bonded cysteines. By binding to chemokine receptors on target cells, they guide cell migration; for example, some chemokines attract inflammatory immune cells. One chemokine receptor, C-C chemokine receptor type 5 (CCR5), is used by HIV as a co-receptor to enter target cells. RANTES, the natural chemokine ligand of CCR5, has potential for suppressing HIV infection. A synthetic version of RANTES that is chemically modified at its N terminus has higher affinity for CCR5 than native RANTES; it also acts as a strong agonist of the CCR5 receptor, inducing receptor signaling. Based on these findings, recombinant analogs devoid of the agonistic activity were sought, with the goal of using them as topical agents to prevent HIV infection (30). A phase I clinical trial of such an analog is scheduled to start next year (31).

PRESENT AND FUTURE APPLICATIONS

Scaffold engineering

With the advances in peptide synthesis technology, the “engineering” approach to producing new peptides, in which small-size natural domains (less than 100 amino acids) are used as a scaffold for structural epitopes of nonhomologous large domains or proteins, became possible (32, 33). Successful epitope designs have been reported in a number of cases; the designed mini-proteins and their larger natural prototypes both bind the same surfaces of the relevant receptors (34). In another study,

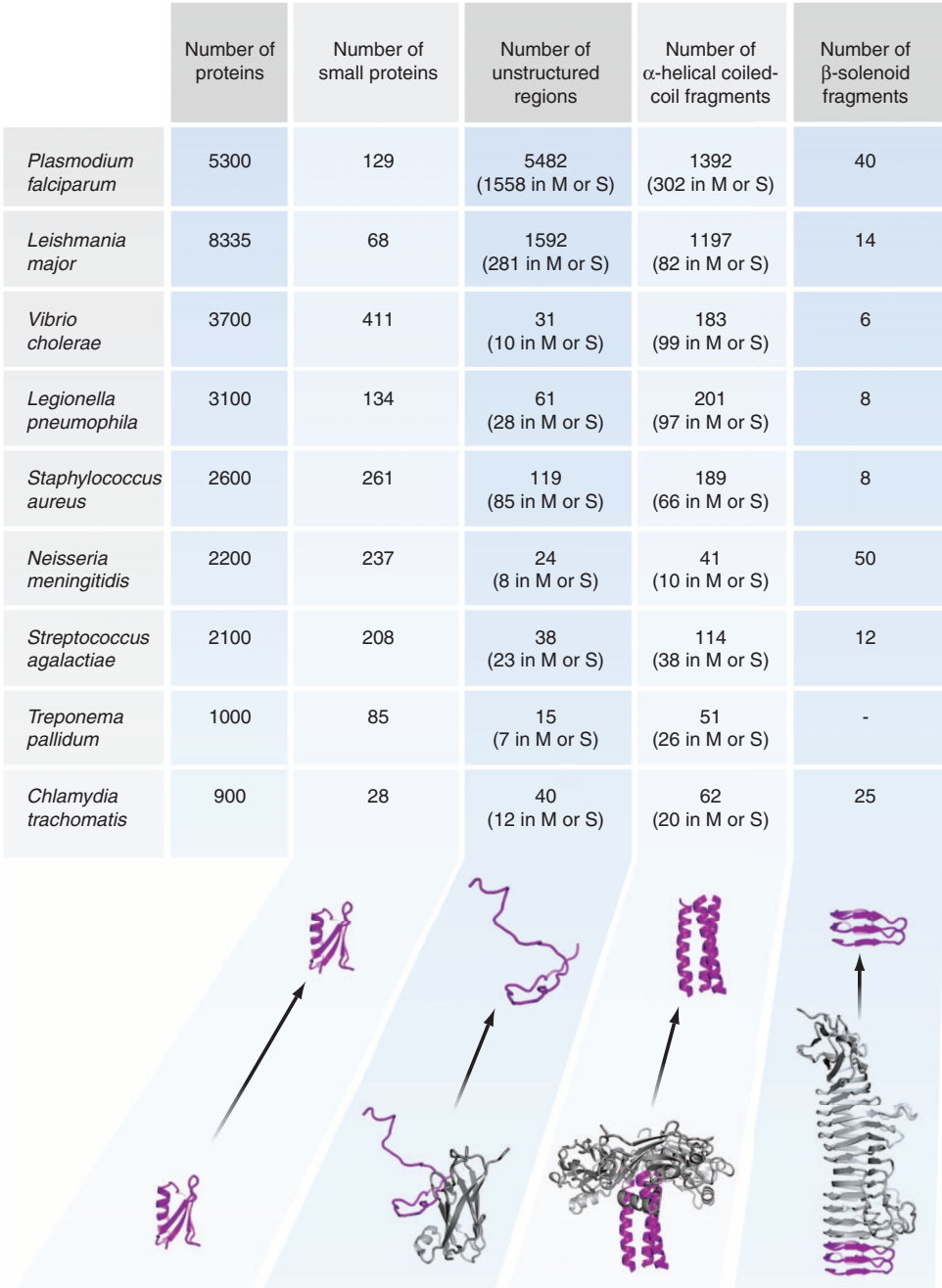


Fig. 2. Estimation of number of vaccine candidates that can be produced with peptide synthesis. The peptide numbers vary from hundreds to thousands depending on the pathogen. These numbers can be even higher if small structural domains of large proteins, such as zinc fingers or EGF modules, are taken into consideration. Among the candidates are (i) small proteins (less than 80 amino acids); (ii) nonoverlapping unstructured regions of more than 30 amino acids found in the whole proteome and in secreted (S) or membrane (M) proteins; (iii) nonoverlapping α -helical coiled-coil regions of more than 30 amino acids found in the whole proteome and in secreted or membrane proteins; and (iv) nonoverlapping β -solenoid fragments of 60 residues found in the outer membrane proteins (48). [Credit: C. Bickel/Science Translational Medicine]

a de novo–designed α -helical coiled-coil domain was used to generate an immune response against α -helical viral proteins or to produce a specific drug such as Fuzeon T-20, which blocks HIV from fusing to the target cell (35–37). Even though these engineering methodologies are poorly adapted to high-throughput screening, undoubtedly they will continue to be attractive with the improvements in peptide synthesis technology.

Reverse vaccinology

Today, the most impressive progress in vaccine development is linked to the increasing number of pathogen genomes that have been sequenced. The genomic data allow in silico selection of promising vaccine candidates regardless of their abundance, their expression conditions, or a need to grow the pathogen in vitro. The approach was named “reverse vaccinology” because in this case vaccine discovery starts from genetic information rather than from the pathogen itself (38). Reverse vaccinology has been successfully used to produce vaccines against some bacterial pathogens (39). At the same time, molecular biology and protein expression technology are still labor-intensive, making this approach fastidious and costly for high-throughput screening. Proper protein folding and solubility remain a limitation in numerous cases.

Structure-based bioinformatics coupled with peptide synthesis

These bottlenecks can be bypassed with an approach that uses the concept of structural epitopes and combines bioinformatics and peptide synthesis. Indeed, peptide synthesis technology is fast and cost-effective. Peptides of 40 to 60 amino acids, which are appropriate for high-throughput chemical synthesis, can form stable structural domains on their own. These peptides can mimic native structural epitopes when isolated from the context of the whole protein. Some of them can be identified in the proteomes of relevant organisms with bioinformatics approaches.

We successfully applied this structure-based bioinformatics and peptide synthesis approach to the identification of new malaria vaccine candidates (40). To select protein fragments that would retain their native conformation adopted in the parent protein, we used the α -helical coiled-coil motif. α -Helical coiled coils are abundant in proteomes; they are characterized by a seven-residue repeat (abcdefg_n), in which hydrophobic residues are located at positions **a** and **d** and, in general, polar residues are located elsewhere. Peptides containing three or more such heptad repeats with bulky apolar residues at the **a** and **d** positions are able to fold into a stable native structure (41). The sequence features of α -helical coiled-coil domains are easily identified by analysis of genomes. Their frequent occurrence in pathogen surface proteins makes them one of the most attractive templates for mimicking structural epitopes. Hence, bioinformatics screening of the *P. falciparum* genome (40) identified several hundred proteins containing putative α -helical coiled-coil motifs. Because we were searching for molecules that could be targets of a protective antibody-based response, we concentrated on those predicted to be expressed in the parasite's erythrocytic stage. The analysis identified over 200 segments that were associated with this stage and that displayed the putative α -helical coiled-coil motifs with a high probability score. Of these α -helical coiled-coil fragments, 170—derived from 140 putative proteins—were synthesized and purified. Ninety-five of them were extensively studied and processed through a rigorous selection procedure to result in 10 promising vaccine candidates (40, 42). The remaining 85 peptides are under study. To design immunogens able to provide immunity to the

general population, we colinearly linked these fragments of 30 to 40 residues via diethylene glycol derivatives, which are a nonimmunogenic handle used to prevent formation of neoantigens. This strategy is universal and can be easily applied to other pathogens and extended to other small structural domains [for example, zinc fingers or epidermal growth factor (EGF) modules] that are able to mimic structural epitopes and that are usually contained in larger proteins (Fig. 2).

β -Solenoid structures

In addition, over the past few years, a new set of proteins has been identified whose folds are based on solenoidal windings of the polypeptide chain in α -helical and (or) β -structural conformations (43). The high incidence of the β -solenoid structures among virulence factors of pathogenic bacteria and viruses (44) suggests that this fold may be an attractive target for new vaccines. Although it is not yet well established whether the β -solenoidal structures can be formed by the folding of shorter synthetic peptides, this possibility cannot be ruled out. Indeed, individual coils of solenoid consist of about 20 residues; therefore, it is possible to synthesize peptides containing three to five coils (Fig. 2). Testing for the presence of such folds in relatively short synthetic peptides by physicochemical or immunological techniques is an attractive task for future work.

Unstructured regions

Finally, the so-called natively unfolded or unstructured regions of proteins, in contrast to all the above-mentioned motifs, do not fold in any particular unique structure (45). Their peptide fragments are also unfolded and, therefore, mimic the native linear epitopes. These regions can be readily identified in protein sequences by bioinformatics analysis; the sequences that encode unstructured regions frequently occur in genomes (45). Selection of suitable candidates from predicted unstructured regions can proceed as described by Villard *et al.* (40). In malaria parasites, fragments of some of these “natively unstructured regions” are already being developed as malaria vaccine candidates (GLURP and MSP3; see above). By following these principles, we identified an unstructured region with very limited polymorphisms (an attractive feature for a vaccine candidate) in the newly identified protein PFF0165c (46). In a time frame of only 2 years, we synthesized, purified, and characterized its immunological and biological properties to lay the basis for a clinical trial.

CONCLUDING REMARKS

Thus, the wealth of proteome data, in combination with new bioinformatics tools developed for proteome analysis and progress in the synthesis of long peptides, opens up possibilities for new strategies of vaccine and drug development. Peptide synthesis methods will be constantly improved, facilitating the production of increasingly longer peptides. The native chemical ligation method already provides a powerful means to assemble unprotected synthetic peptides to proteins in aqueous solution (47). Therefore, larger structural motifs than those described in this Review may be considered. Thus, advances in bioinformatics tools for prediction of structural domains and peptide synthesis, coupled with sequencing of pathogen genomes, should ensure the increasing interest of researchers and developers in the concept of structural epitopes for the development of synthetic drugs and vaccines.

REFERENCES AND NOTES

1. M. Bergmann, L. Zervas, Über ein allgemeines verfahren der peptid-synthese. *Ber. Deutsch. Chem. Ges.* **65**, 1192–1201 (1932).
2. V. du Vigneaud, C. Ressler, J. M. Swan, C. W. Roberts, P. G. Katsoyannis, The synthesis of oxytocin. *J. Am. Chem. Soc.* **76**, 3115–3121 (1954).
3. R. Schwyzler, P. Sieber, Total synthesis of adrenocorticotrophic hormone. *Nature* **199**, 172–174 (1963).
4. R. B. Merrifield, Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
5. E. Atherton, R. C. Sheppard, *Solid Phase Peptide Synthesis: A Practical Approach* (IRL Press, Oxford, 1989).
6. S. B. H. Kent, Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **57**, 957–989 (1988).
7. S. Olugbile, C. Habel, C. Servis, F. Spertini, A. Verdini, G. Corradin, Malaria vaccines—the long synthetic peptide approach: Technical and conceptual advancements. *Curr. Opin. Mol. Ther.* **12**, 64–76 (2010).
8. M. A. Hossain, A. Belgi, F. Lin, S. Zhang, F. Shabanpoor, L. Chan, C. Belyea, H.-T. Truong, A. R. Blair, S. Andrikoupolos, G. W. Tregear, J. D. Wade, Use of a temporary “solubilizing” peptide tag for the Fmoc solid-phase synthesis of human insulin glargine via use of regioselective disulfide bond formation. *Bioconjug. Chem.* **20**, 1390–1396 (2009).
9. N. Robertson, R. Ramage, Total chemical synthesis of deglycosylated human erythropoietin. *J. Chem. Soc. Perkin Trans. 1*, 1015–1021 (1999).
10. W. M. Kazmierski, J. McDermed, Synthesis of the carbonic acid benzotriazol-1-yl-ester-(2-biotinylamino)-9H-fluoren-9-ylmethyl ester: A convenient transient-biotinylation reagent for use in affinity chromatography. *Tetrahedron Lett.* **36**, 9097–9100 (1995).
11. M. A. Roggero, C. Servis, G. Corradin, A simple and rapid procedure for the purification of synthetic polypeptides by a combination of affinity chromatography and methionine chemistry. *FEBS Lett.* **408**, 285–288 (1997).
12. J. Vizzavona, M. Villain, K. Rose, Covalent capture purification of polypeptides after SPPS via a linker removable under very mild conditions. *Tetrahedron Lett.* 8693–8696 (2002).
13. T. Hara, A. Tainosho, K. Nakamura, T. Sato, T. Kawakami, S. Aimoto, Peptide purification by affinity chromatography based on α -ketoacyl group chemistry. *J. Pept. Sci.* **15**, 369–376 (2009).
14. A. Verdini, S. Terenzi, V. Brossard, M. Roggero, G. Corradin, Oxidative folding of synthetic polypeptides S-protected as *tert*-butylthio derivatives. *J. Pept. Sci.* **14**, 1271–1282 (2008).
15. J. B. Dame, J. L. Williams, T. F. McCutchan, J. L. Weber, R. A. Wirtz, W. T. Hockmeyer, W. L. Maloy, J. D. Haynes, I. Schneider, D. Roberts, G. S. Sanders, E. P. Reddy, C. L. Diggs, L. H. Miller, Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. *Science* **225**, 593–599 (1984).
16. E. H. Nardin, R. S. Nussenzweig, T cell responses to pre-erythrocytic stages of malaria: Role in protection and vaccine development against pre-erythrocytic stages. *Annu. Rev. Immunol.* **11**, 687–727 (1993).
17. J. A. López, C. Weilenman, R. Audran, M. A. Roggero, A. Bonelo, J. M. Tiercy, F. Spertini, G. Corradin, A synthetic malaria vaccine elicits a potent CD8⁺ and CD4⁺ T lymphocyte immune response in humans. Implications for vaccination strategies. *Eur. J. Immunol.* **31**, 1989–1998 (2001).
18. R. Audran, F. Lurati-Ruiz, B. Genton, H. E. Blythman, O. Ofori-Anyinam, O. C. Reymond, G. Corradin, F. Spertini, The synthetic *Plasmodium falciparum* circumsporozoite peptide PfCS102 as a malaria vaccine candidate: A randomized controlled phase I trial. *PLoS One* **4**, e7304 (2009).
19. B. Genton, V. D’Acremont, F. Lurati-Ruiz, D. Verhage, R. Audran, C. Hermesen, L. Wolters, C. Reymond, F. Spertini, R. Sauerwein, Randomized double-blind controlled Phase I/IIa trial to assess the efficacy of malaria vaccine PfCS102 to protect against challenge with *P. falciparum*. *Vaccine* **28**, 6573–6580 (2010).
20. S. Herrera, O. L. Fernández, O. Vera, W. Cárdenas, O. Ramírez, R. Palacios, M. Chen-Mok, G. Corradin, M. Arévalo-Herrera, Phase I safety and immunogenicity trial of *Plasmodium vivax* CS derived long synthetic peptides adjuvanted with Montanide ISA 720 or Montanide ISA 51. *Am. J. Trop. Med. Hyg.*, in press.
21. S. B. Sirima, A. B. Tiono, A. Ouédraogo, A. Diarra, A. L. Ouédraogo, J. B. Yaro, E. Ouédraogo, A. Gansané, E. C. Bougouma, A. T. Konaté, Y. Kaboré, A. Traoré, R. Chilengi, I. Soulama, A. J. Luty, P. Druilhe, S. Cousens, I. Nébédjé, Safety and immunogenicity of the malaria vaccine candidate MSP3 long synthetic peptide in 12–24 months-old Burkinabe children. *PLoS One* **4**, e7549 (2009).
22. J. P. Lusingu, S. Gesase, S. Msham, F. Francis, M. Lemnge, M. Seth, S. Semuche, A. Rutta, D. Minja, M. D. Segeja, S. Bosompah, S. Cousens, R. Noor, R. Chilengi, P. Druilhe, Satisfactory safety and immunogenicity of MSP3 malaria vaccine candidate in Tanzanian children aged 12–24 months. *Malar. J.* **8**, 163 (2009).
23. M. Esen, P. G. Kremsner, R. Schleucher, M. Gässler, E. B. Imoukhuede, N. Imbault, O. Leroy, S. Jepsen, B. W. Knudsen, M. Schumm, J. Knobloch, M. Theisen, B. Mordmüller, Safety and immunogenicity of GMZ2—a MSP3–GLURP fusion protein malaria vaccine candidate. *Vaccine* **27**, 6862–6868 (2009).
24. C. Widmann, P. Romero, J. L. Maryanski, G. Corradin, D. Valmori, T helper epitopes enhance the cytotoxic response of mice immunized with MHC class I-restricted malaria peptides. *J. Immunol. Methods* **155**, 95–99 (1992).
25. G. G. Kenter, M. J. Welters, A. R. Valentijn, M. J. Lowik, D. M. Berends-van der Meer, A. P. Vloon, F. Essahsah, L. M. Fathers, R. Offringa, J. W. Drijfhout, A. R. Wafelman, J. Oostendorp, G. J. Fleuren, S. H. van der Burg, C. J. Melief, Vaccination against HPV-16 oncoproteins for vulvar intraepithelial neoplasia. *N. Engl. J. Med.* **361**, 1838–1847 (2009).
26. N. Leffers, A. J. Lambeck, M. J. Gooden, B. N. Hoogbeem, R. Wolf, I. E. Hamming, B. G. Hepkema, P. H. Willemse, B. H. Molmans, H. Hollema, J. W. Drijfhout, W. J. Sluiter, A. R. Valentijn, L. M. Fathers, J. Oostendorp, A. G. van der Zee, C. J. Melief, S. H. van der Burg, T. Daemen, H. W. Nijman, Immunization with a P53 synthetic long peptide vaccine induces P53-specific immune responses in ovarian cancer patients, a phase II trial. *Int. J. Cancer* **125**, 2104–2113 (2009).
27. C. J. Melief, S. H. van der Burg, Immunotherapy of established (pre)malignant disease by synthetic long peptide vaccines. *Nat. Rev. Cancer* **8**, 351–360 (2008).
28. C. von Garnier, M. Astori, A. Kettner, N. Dufour, C. Heusser, G. Corradin, F. Spertini, Allergen-derived long peptide immunotherapy down-regulates specific IgE response and protects from anaphylaxis. *Eur. J. Immunol.* **30**, 1638–1645 (2000).
29. F. Spertini, personal communication.
30. H. Gaertner, F. Cerini, J. M. Escola, G. Kuenzi, A. Melotti, R. Offord, I. Rossitto-Borlat, R. Nedellec, J. Salkowitz, G. Gorochov, D. Mosier, O. Hartley, Highly potent, fully recombinant anti-HIV chemokines: Reengineering a low-cost microbicide. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17706–17711 (2008).
31. O. Hartley, personal communication.
32. L. Martin, C. Vita, Engineering novel bioactive mini-proteins from small size natural and de novo designed scaffolds. *Curr. Protein Pept. Sci.* **1**, 403–430 (2000).
33. M. E. Houston Jr., A. Wallace, E. Bianchi, A. Pessi, R. S. Hodges, Use of a conformationally restricted secondary structural element to display peptide libraries: A two-stranded α -helical coiled-coil stabilized by lactam bridges. *J. Mol. Biol.* **262**, 270–282 (1996).
34. C. C. Huang, F. Stricher, L. Martin, J. M. Decker, S. Majeed, P. Barthe, W. A. Hendrickson, J. Robinson, C. Roumestand, J. Sodroski, R. Wyatt, G. M. Shaw, C. Vita, P. D. Kwong, Scorpion-toxin mimics of CD4 in complex with human immunodeficiency virus gp120 crystal structures, molecular mimicry, and neutralization breadth. *Structure* **13**, 755–768 (2005).
35. C. T. Wild, D. C. Shugars, T. K. Greenwell, C. B. McDanal, T. J. Matthews, Peptides corresponding to a predictive α -helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9770–9774 (1994).
36. R. Miceli, D. Myszk, J. Mao, G. Sathe, I. Chaiken, The coiled coil stem loop miniprotein as a presentation scaffold. *Drug Des. Discov.* **13**, 95–105 (1996).
37. S. M. Lu, R. S. Hodges, A de novo designed template for generating conformation-specific antibodies that recognize α -helices in proteins. *J. Biol. Chem.* **277**, 23515–23524 (2002).
38. R. Rappuoli, Reverse vaccinology. *Curr. Opin. Microbiol.* **3**, 445–450 (2000).
39. B. Capecchi, D. Serruto, J. Adu-Bobie, R. Rappuoli, M. Pizzi, The genome revolution in vaccine research. *Curr. Issues Mol. Biol.* **6**, 17–27 (2004).
40. V. Villard, G. W. Agak, G. Frank, A. Jafarshad, C. Servis, I. Nébédjé, S. B. Sirima, I. Felger, M. Arevalo-Herrera, S. Herrera, F. Heitz, V. Bäcker, P. Druilhe, A. V. Kajava, G. Corradin, Rapid identification of malaria vaccine candidates based on α -helical coiled coil protein motif. *PLoS One* **2**, e645 (2007).
41. P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**, 1401–1407 (1993).
42. G. W. Agak, P. Bejon, G. Fegan, N. Gicheru, V. Villard, A. V. Kajava, K. Marsh, G. Corradin, Longitudinal analyses of immune responses to *Plasmodium falciparum* derived peptides corresponding to novel blood stage antigens in coastal Kenya. *Vaccine* **26**, 1963–1971 (2008).
43. B. Kobe, A. V. Kajava, When protein folding is simplified to protein coiling: The continuum of solenoid protein structures. *Trends Biochem. Sci.* **25**, 509–515 (2000).
44. A. V. Kajava, A. C. Steven, The turn of the screw: Variations of the abundant β -solenoid motif in passenger domains of type V secretory proteins. *J. Struct. Biol.* **155**, 306–315 (2006).
45. P. E. Wright, J. H. Dyson, Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331 (1999).
46. S. Olugbile, C. Kulangara, G. Bang, S. Bertholet, E. Suzarte, V. Villard, G. Frank, R. Audran, A. Razaname, I. Nebie, O. Awobusuyi, F. Spertini, A. V. Kajava, I. Felger, P. Druilhe, G. Corradin, Vaccine potentials of an intrinsically unstructured fragment derived from the blood stage-associated *Plasmodium falciparum* protein PFF0165C. *Infect. Immun.* **77**, 5701–5709 (2009).
47. P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).
48. The *P. falciparum*, *L. major*, and bacterial genomes were downloaded from <http://plasmodb.org>, http://ftp.sanger.ac.uk/pub/databases/Lmajor_sequences/DATASETS/, and <http://ftp.ncbi.nih.gov/genomes/>, correspondingly. The α -helical coiled-coil regions were identified by using

generalized sequence profiles as described (40). β -Solenoid fragments were identified by using the CRBM library of profiles (44) (<http://bioinfo.montp.cnrs.fr/?p=profiles&r=profiles>). Membrane (M) and (or) secreted (S) proteins were selected by using generalized sequence profiles of typical transmembrane regions and the N-terminal signal peptides (49). Unstructured regions were identified by calculating TopIDP index in the 30-residue sliding window (50).

49. P. Bucher, K. Karplus, N. Moeri, K. Hofmann, A flexible motif search technique based on generalized profiles. *Comput. Chem.* **20**, 3–23 (1996).
50. A. Campen, R. M. Williams, C. J. Brown, J. Meng, V. N. Uversky, A. K. Dunker, TOP-IDP-scale: A new amino acid scale measuring propensity for intrinsic disorder. *Protein Pept. Lett.* **15**, 956–963 (2008).
51. **Acknowledgments:** We thank J. Maüel, M. Roggero, and C. Servis for discussion and F. Spertini, O. Hartley, and J.-P. Salles for sharing relevant information. For more information on clinical trials

based on LSPs, please consult <http://clinicaltrials.gov/ct2/results?term=long+synthetic+peptides> and <http://www.emvi.org>. **Competing interests:** G.C. and A.V.K. hold patents related to malaria vaccine candidates. G.C. is Vaxéal-Group vice president of technical operations and malaria vaccine development.

Submitted 16 June 2010

Accepted 3 September 2010

Published 22 September 2010

10.1126/scitranslmed.3001387

Citation: G. Corradin, A. V. Kajava, A. Verdini, Long synthetic peptides for the production of vaccines and drugs: A technological platform coming of age. *Sci. Transl. Med.* **2**, 50rv3 (2010).