Synthetic Biology

Synthetic Biology

What is synthetic biology?

Synthetic biology refers to both:

→ the design and fabrication of biological components and systems that do not already exist in the natural world

→ the re-design and fabrication of existing biological systems.

http://syntheticbiology.org/FAQ.html

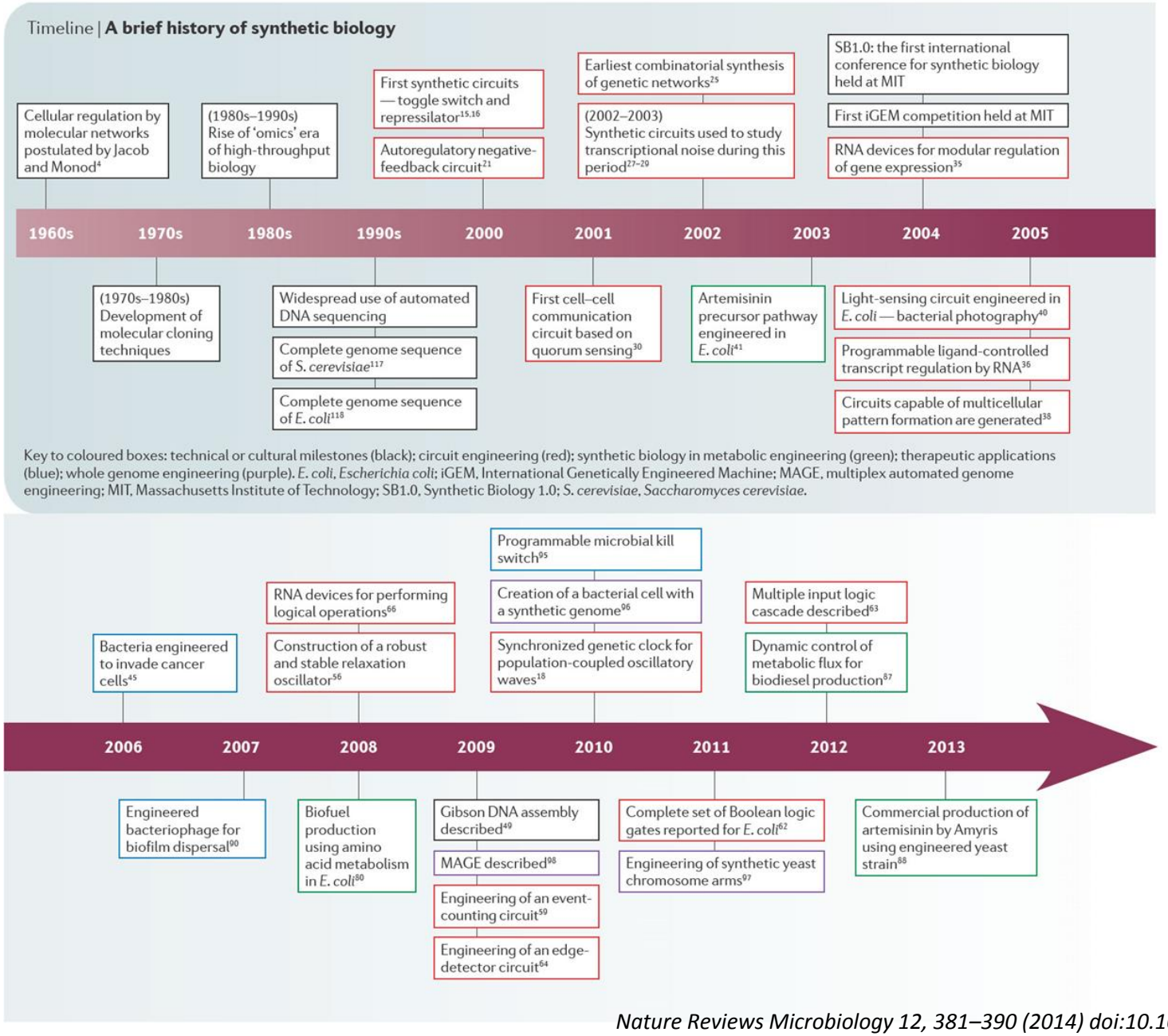
What is the difference between synthetic biology and systems biology?

Systems biology

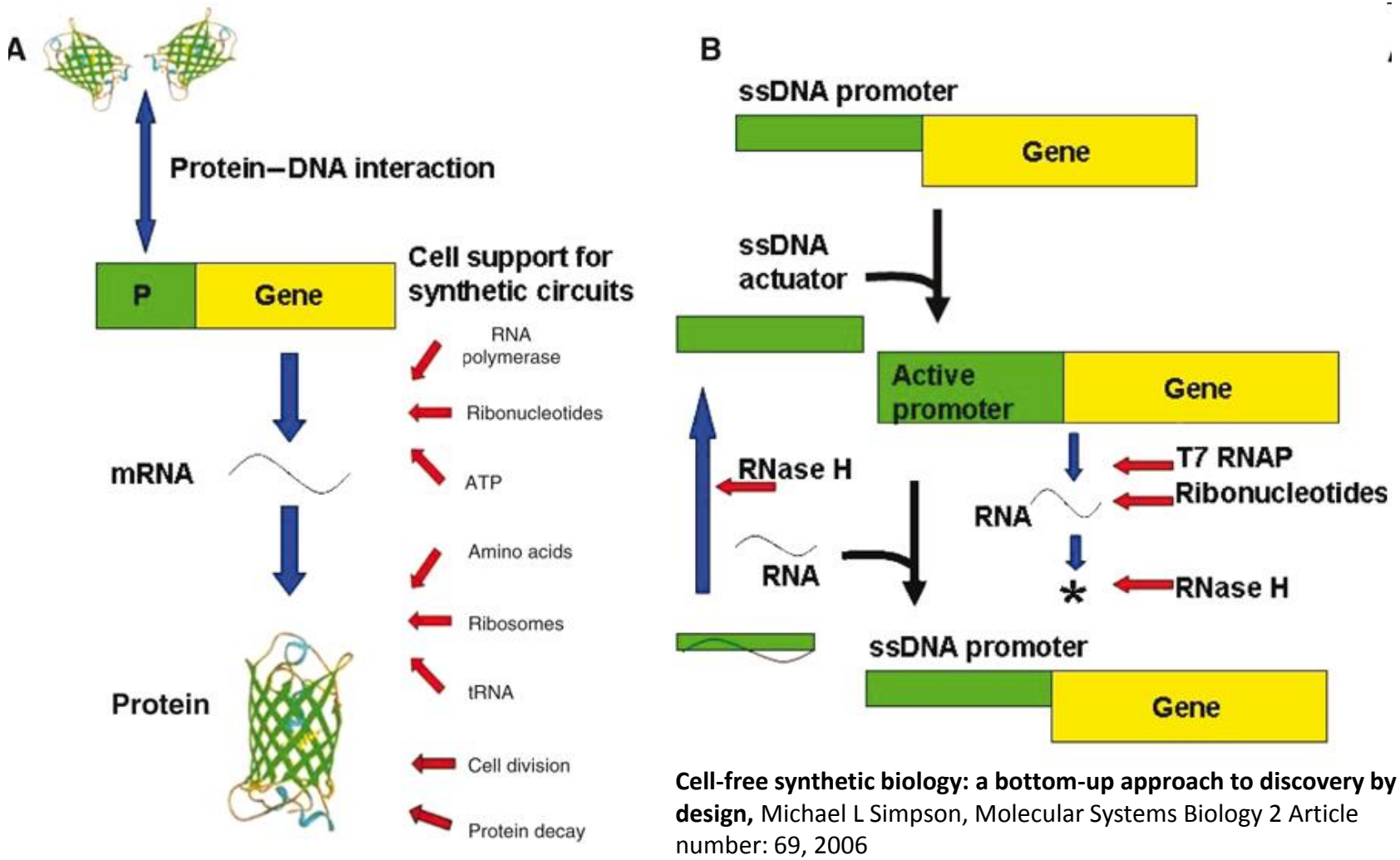
→ studies complex biological systems as integrated wholes, using tools of modeling, simulation, and comparison to experiment. The focus tends to be on natural systems, often with some (at least long term) medical significance.

Synthetic biology

→ studies how to build artificial biological systems for engineering applications, using many of the same tools and experimental techniques.
But the work is **fundamentally an engineering application** of biological science, rather than an attempt to do more science. The focus is often on ways of taking parts of natural biological systems, characterizing and simplifying them, and using them as a component of a highly unnatural, engineered, biological system.



Nature Reviews Microbiology 12, 381–390 (2014) doi:10.1038/nrmicro3239



Cell-free synthetic biology: a bottom-up approach to discovery by design, Michael L Simpson, Molecular Systems Biology 2 Article number: 69, 2006

Comparison of whole-cell and cell-free synthetic biology. (**A**) Global cellular resources support the operation of synthetic circuits in whole-cell synthetic biology, but also provide subtle pathways for interactions between otherwise unconnected gene circuits. Much of transcriptional regulation is achieved through protein–DNA interactions. (**B**) The cell-free synthetic biology approach reported by Kim *et al* removes global resources to support gene circuit operation, translation, and protein transcription factors. Transcriptional regulation is achieved by the competing actions of single-stranded DNA activators and complementary strands of RNA.

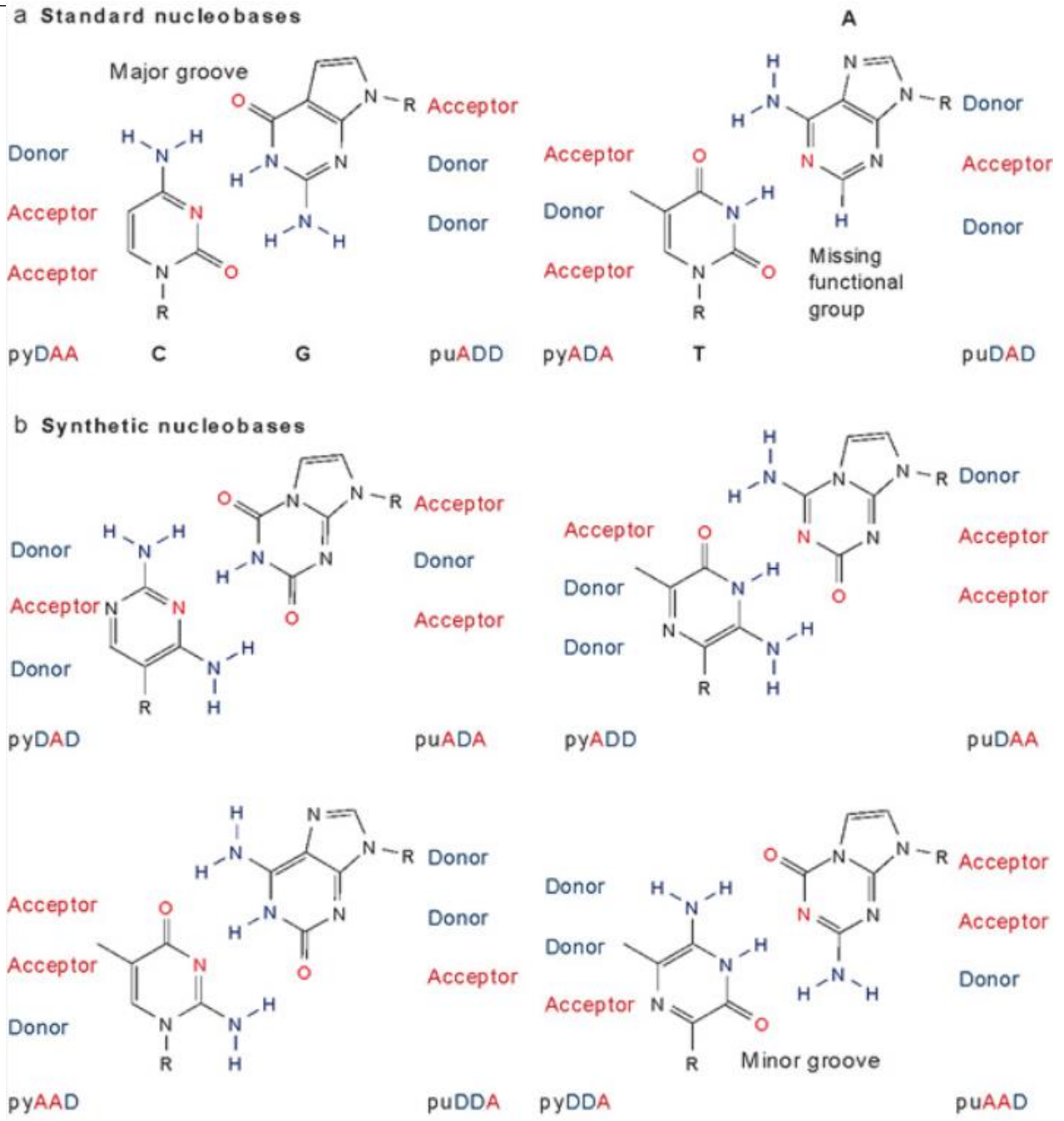


Figure 1 | **Examples of alternative nucleobases.** Parts of the nucleobases of DNA can be used as interchangeable building modules. The blue units are the hydrogen bonding donor (D) collections of atoms. The red units are the hydrogen bonding acceptor (A) collections of atoms. a | The four standard nucleobases are shown. b | Shuffling the hydrogen bond donor and acceptor modules generates eight additional nucleotides, which constitute a synthetic genetic system. These synthetic bases have been used in an artificial genetic system that can support Darwinian evolution. A, adenine; C, cytosine; G, guanine; Pu, purine; Py, pyrimidine; T, thymine.

Nature Reviews Genetics 6, 533-543 (July 2005)

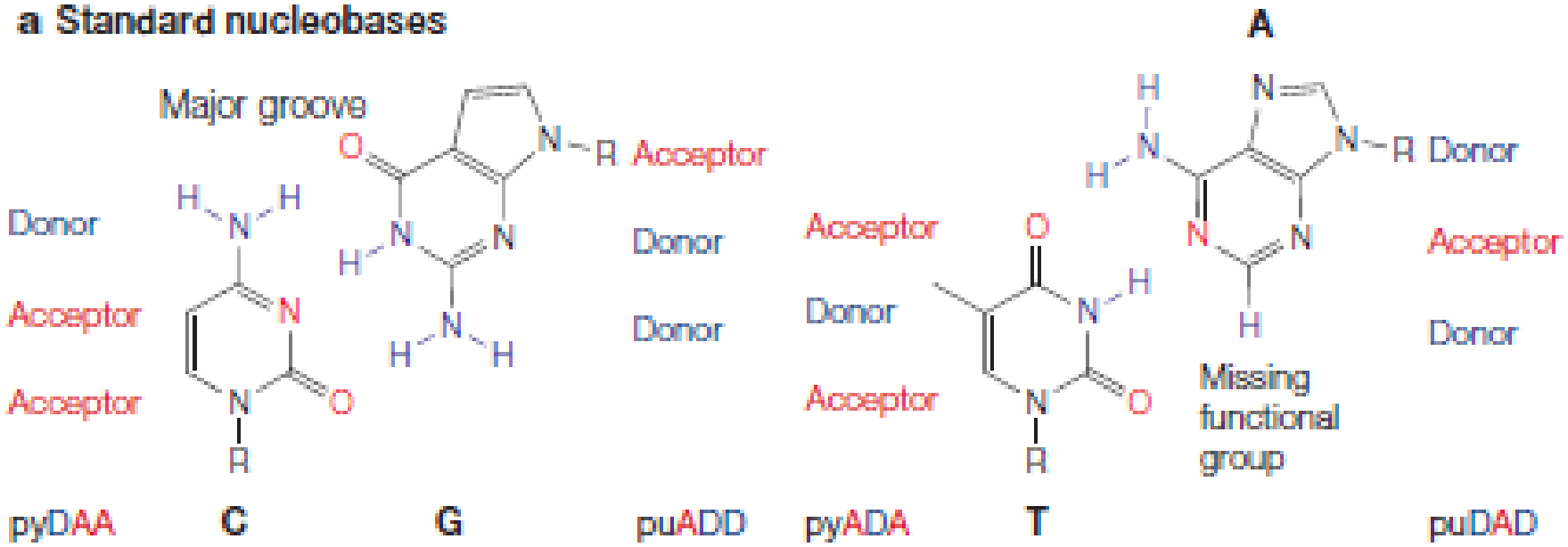


Figure 1 | **Examples of alternative nucleobases.** Parts of the nucleobases of DNA can be used as interchangeable building modules. The blue units are the hydrogen bonding donor (D) collections of atoms. The red units are the hydrogen bonding acceptor (A) collections of atoms.

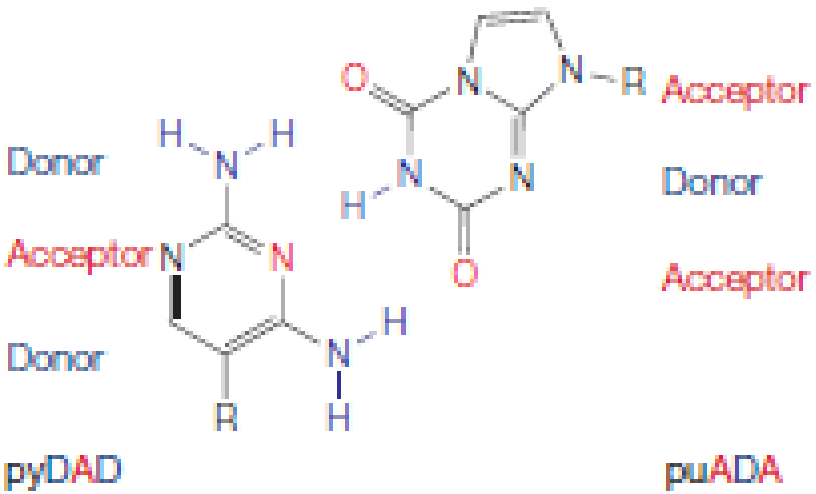
a | The four standard nucleobases are shown.

b | Shuffling the hydrogen bond donor and acceptor modules generates eight additional nucleotides, which constitute a synthetic genetic system. These synthetic bases have been used in an artificial genetic system that can support Darwinian evolution.

A, adenine; C, cytosine; G, guanine; Pu, purine; Py, pyrimidine; T, thymine.

Nature Reviews Genetics 6, 533-543 (July 2005)

b Synthetic nucleobases



Nature Reviews Genetics 6, 533-543 (July 2005)

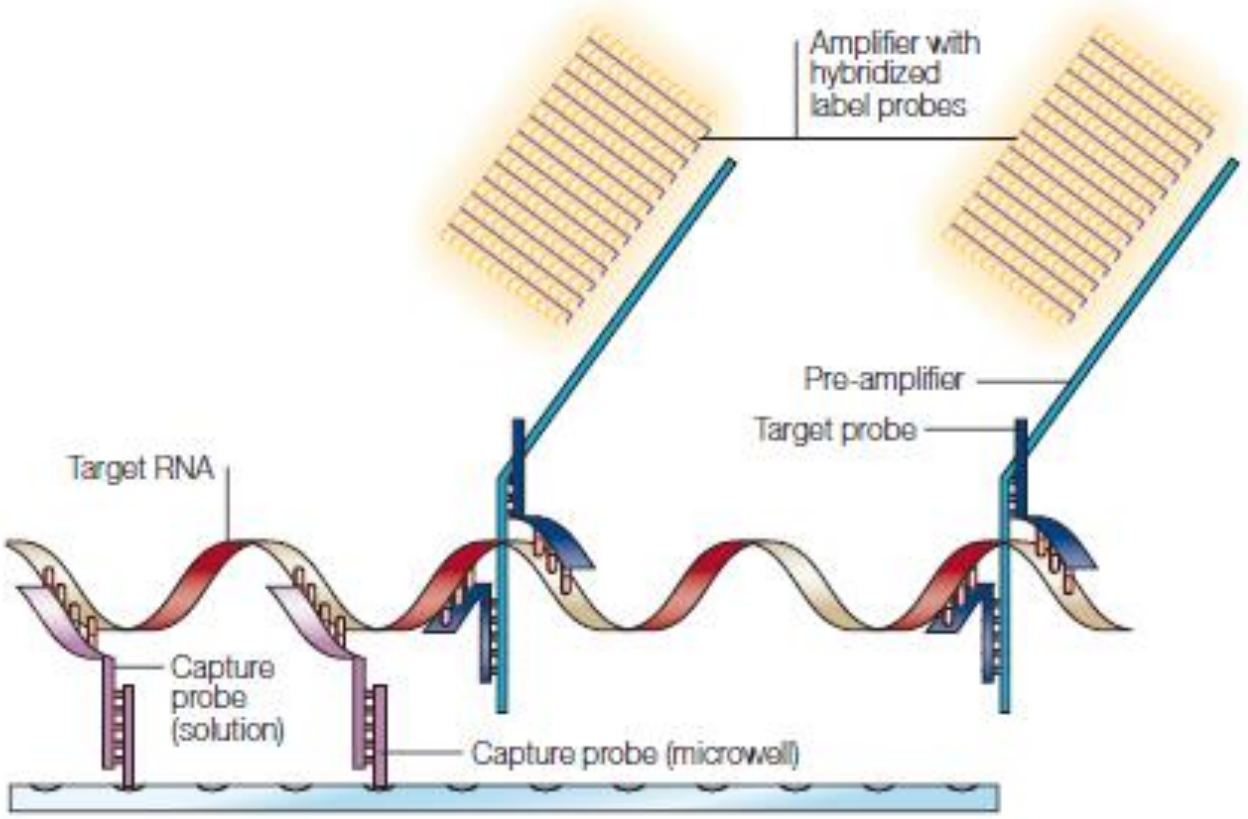


Figure 2 | **Branched DNA assay developed by scientists at Chiron and Bayer Diagnostics.** The target RNA molecule to be detected (the analyte) is attached to the plastic of a microwell (bottom) by the hybridization of the analyte to a **series of capture probes**. This complex then captures, through hybridization, a target probe, which in turn hybridizes to a **pre-amplifier molecule**, thereby 'sandwiching' the analyte between the capture probe and the pre-amplifier. The pre-amplifier captures a branched DNA dendrimer (amplifier) that contains several signalling molecules on each branch. As a consequence of the branching, a single analyte assembles a large number of signalling molecules in the microwell. These assays use the expanded genetic alphabet shown in FIG. 1. When standard nucleotides were used to assemble the signalling nanostructure, significant noise was seen, because non-target DNA that was present in the biological sample was captured by the probes in the microwell even in the absence of analyte. Incorporating components of the artificial genetic alphabet in the dendrimer reduced the noise. As a consequence, the assay now helps manage the care of some 400,000 patients annually, **detecting as few as eight molecules** of the analyte DNA in a sample.

Nature Reviews Genetics 6, 533-543 (July 2005)

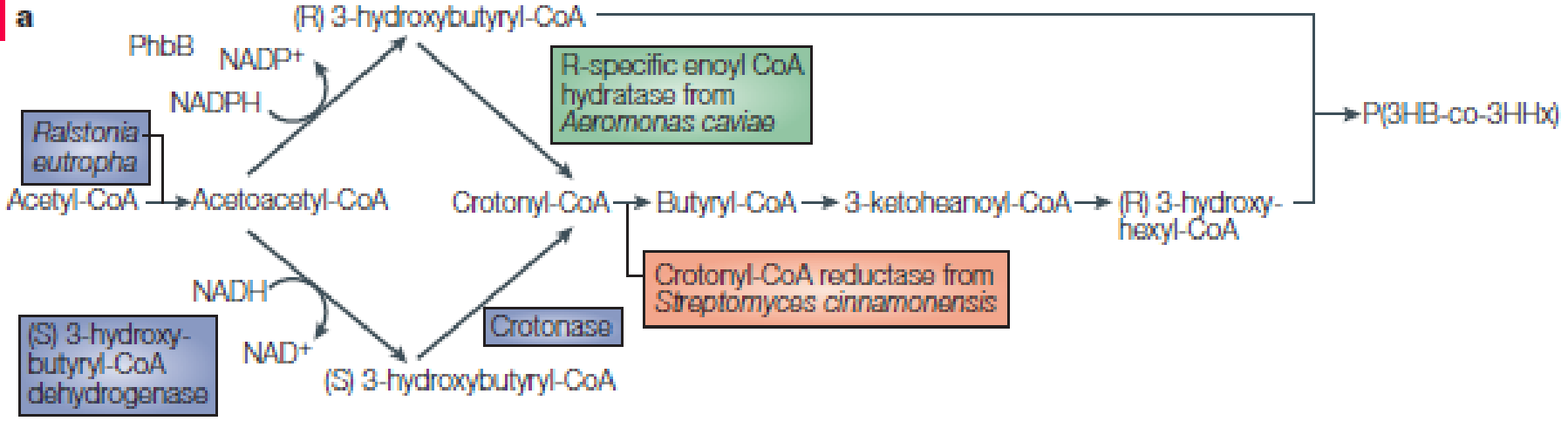
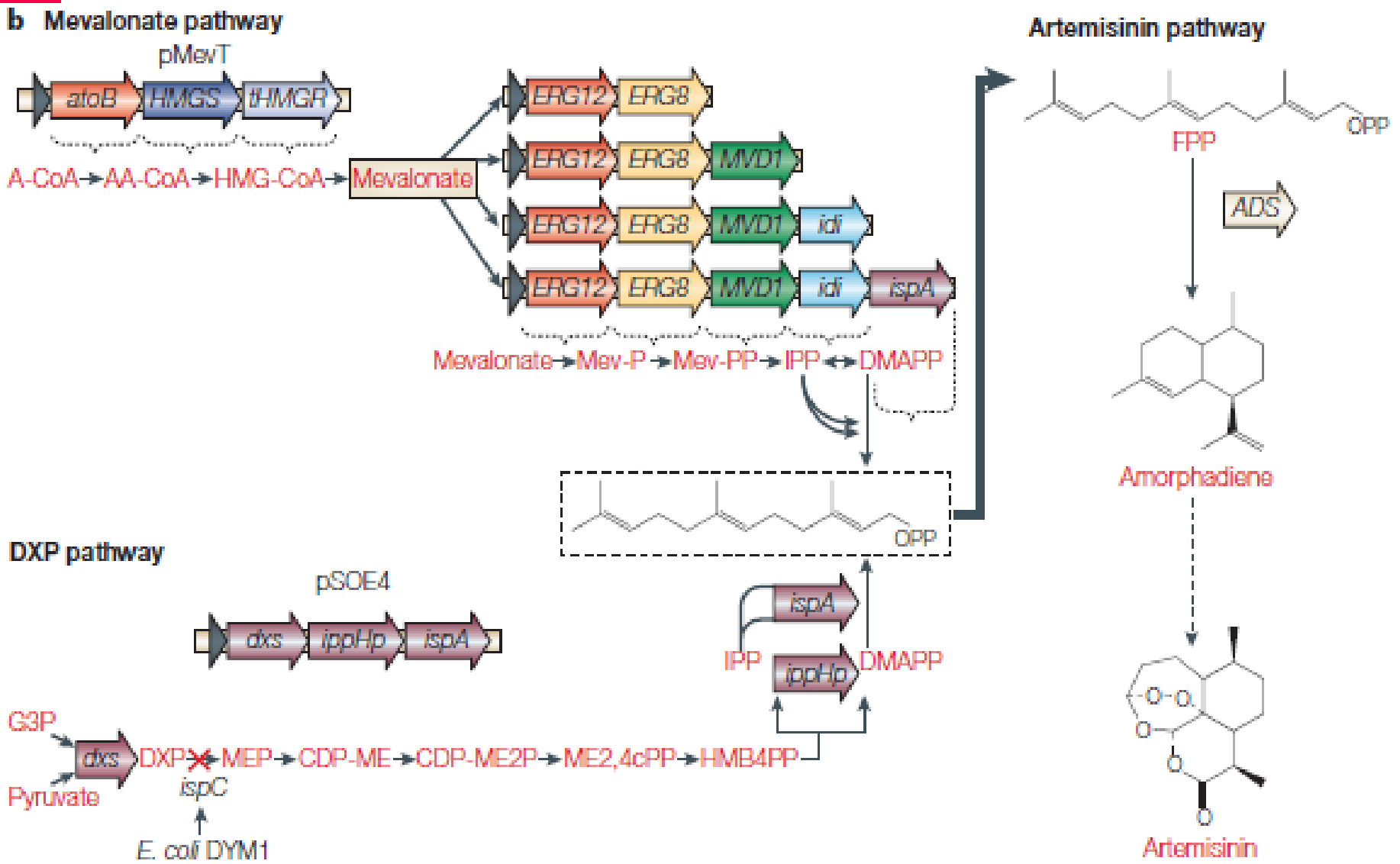
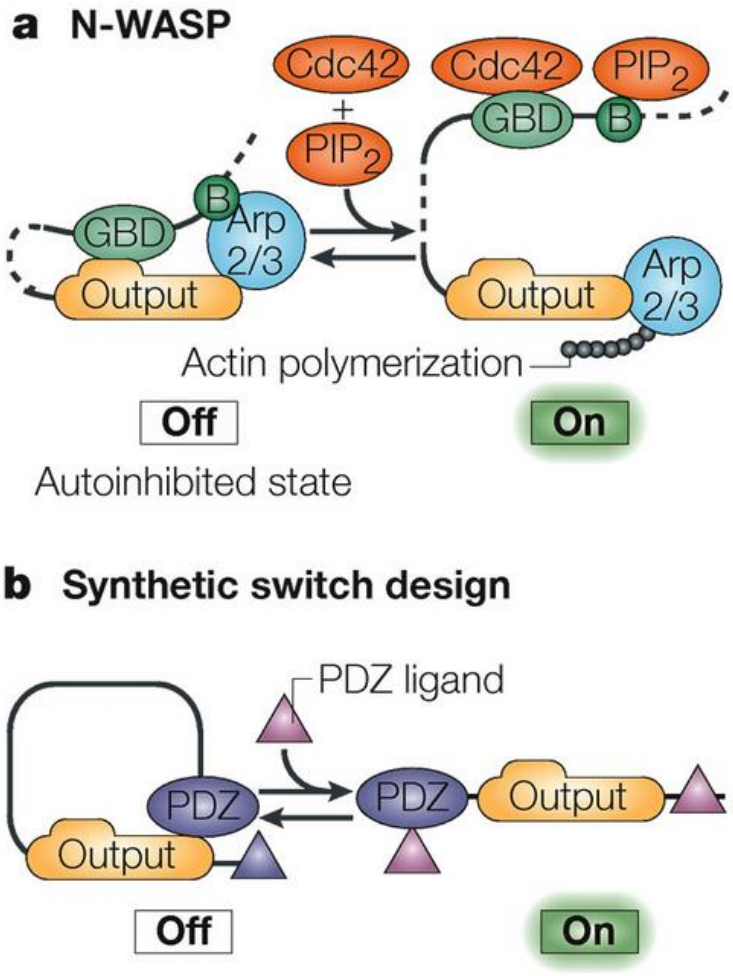


Figure 4 | **Using proteins as interchangeable parts in synthetic biology. a** | The combination of enzymes from three sources in a *Ralstonia eutropha* host generated a strain that produced large amounts of a poly(3-hydroxybutyrate-co-3-hydroxyhexanoate) polymer from fructose. All enzymes from *Ralstonia eutropha* are shown in black, whereas those from *Areomonas caviae* and *Streptomyces cinnamonensis* are shown in green and red, respectively. Modified, with permission, from REF. 76 © American Chemical Society (2002). **b** | The combination of enzymes from three sources in an *Escherichia coli* host generated a strain of the bacterium that produced a precursor for artemisinin, an antimalarial drug. The challenge of this experiment lay in the need to curtail the pathway to recognize and avoid metabolite toxicity, while optimizing the yield of the desired product. The general methodology for the pathway design was to use an engineered mevalonate pathway that is absent in *E. coli*, rather than the DXP (1-deoxy-D-xylulose 5-phosphate) pathway that is native to the organism. The synthetic operons used are depicted, and the engineered pathway metabolites are shown in red. In the engineered mevalonate pathway, the fan of genes from *ERG12* to *ispA* exist on multiple plasmids to tune the pathway for optimization of the product while avoiding metabolite toxicity. As depicted at the bottom of the figure, the *E. coli* strain DYM1, a strain deficient in isoprenoid synthesis, was used, because the DXP pathway was found to limit product yield, probably owing to an unrecognized link between the pathway and physiological control elements in the organism. Enzymes used (isolated from *Saccharomyces cerevisiae* unless otherwise noted): *ADS*, amorphadiene synthase; *atoB*, acetoacetyl-CoA thiolase (*E. coli*); *dxs*, 1-deoxy-D-xylulose 5-phosphate synthase; *ERG12*, mevalonate kinase; *ERG8*, phosphomevalonate kinase; *HMGS*, HMG-CoA synthase; *idi*, IPP isomerase (*E. coli*); *ippHp*, IPP isomerase (*H. pluvialis*); *ispA*, FPP synthase (*E. coli*); *ispC*, 1-deoxy-D-xylulose 5-phosphate reductoisomerase; *MVD1*, mevalonate pyrophosphate decarboxylase; *tHMGR*, truncated HMG-CoA reductase. Pathway intermediates: AA-CoA, acetoacetly-CoA; A-CoA, acetyl-CoA; CDP-Me, 4-diphosphocytidyl-2-C-methyl-D-erythritol; CDP-ME2P, 4-diphosphocytidyl-2-C-methyl-D-erythritol 2-phosphate; DMAPP, dimethylallyl pyrophosphate; DXP, 1-deoxy-D-xylulose 5-phosphate; FPP, farnesyl pyrophosphate; G3P, glyceraldehyde 3-phosphate; HMB4PP, 1-hydroxy-2-methyl-2-(E)-butenyl 4-pyrophosphate; HMG-CoA, hydroxymethylglutaryl-CoA; IPP, isopentenyl pyrophosphate; Mav-P, mevalonate 5-phosphate; ME-2,4cPP, 2-C-methyl-D-erythritol 2,4-cyclopyrophosphate; MEP, 2-C-methyl-D-erythritol 4-phosphate; Mev-PP, mevalonate pyrophosphate. Adapted, with permission, from *Nature Biotechnology* REF. 77 © (2003) Macmillan Magazines Ltd.

Nature Reviews Genetics 6, 533-543 (July 2005)



Nature Reviews Genetics 6, 533-543 (July 2005)



Engineering regulatory circuits

The engineering of molecular circuits has recently been reviewed by Hasty *et al.*81 ; this article should be consulted for more details of the cases briefly reviewed below, and for further examples. In one experiment in yeast, Lim and his co-workers rewired Mapk signalling93 and an actin regulatory switch known as N-WASP 94,95. N-WASP carries an output region that (in isolation) stimulates the polymerization of actin by binding the actin-related protein (Arp) 2/3 complex (see part a of the figure). This activity is autoinhibited, however, by means of two modules from N-WASP, a highly basic (B) motif and a guanosine 5'- triphosphatase (GTPase)-binding domain (GBD). These bind to two activating inputs, the phosphoinositide Pip2 and the activated GTPase Cdc42, where the respective binding disrupts autoinhibition. Because the two activating inputs function cooperatively, N-WASP behaves as an 'AND' gate; the output is positive (actin is polymerized) only if both Pip2 and Cdc42 inputs are present.

Lim and co-workers then attempted to reprogram the input control of N-WASP. They tethered an unrelated modular domain–ligand pair (a PDZ domain and its cognate C-terminal peptide ligand) to the end of the N-WASP output domain (see part b of figure). They expected this to create an unnatural autoinhibitory interaction that could be relieved by the competitive binding of an external PDZ ligand. Under basal conditions, this synthetic switch was not activating. The unnatural autoinhibitory interaction was then removed by adding the PDZ ligand. The maximal activity was close to that of the isolated output domain. The result was the same switch, but one that was turned on by a different activator.

The synthetic biologist then attempted to construct AND gates by tethering two unnatural modular domain-ligand pairs to the N-WASP output domain, replacing the two natural modules (B and GBD). A total of 34 constructs were synthesized and examined for their gating behaviour.

As expected for any real chemical system, the output depended on the concentrations of the input molecules, and was not predictable in an engineering sense. Some switches showed little basal repression. Others could not be activated. A few showed antagonistic gating, with one input activating and the other repressing. A few showed 'OR' gating, where one or the other input generated a positive output. Nevertheless, about half of the synthetic constructs showed positive AND gating, where both inputs were required to generate a positive output. This input-output relationship was analogous to that seen with native N-WASP, but was obtained from a synthetic system, by using unnatural activators. Figure modified, with permission, from REF. 95 c **(2004) Elsevier Science.**

Nature Reviews Genetics 6, 533-543 (July 2005)

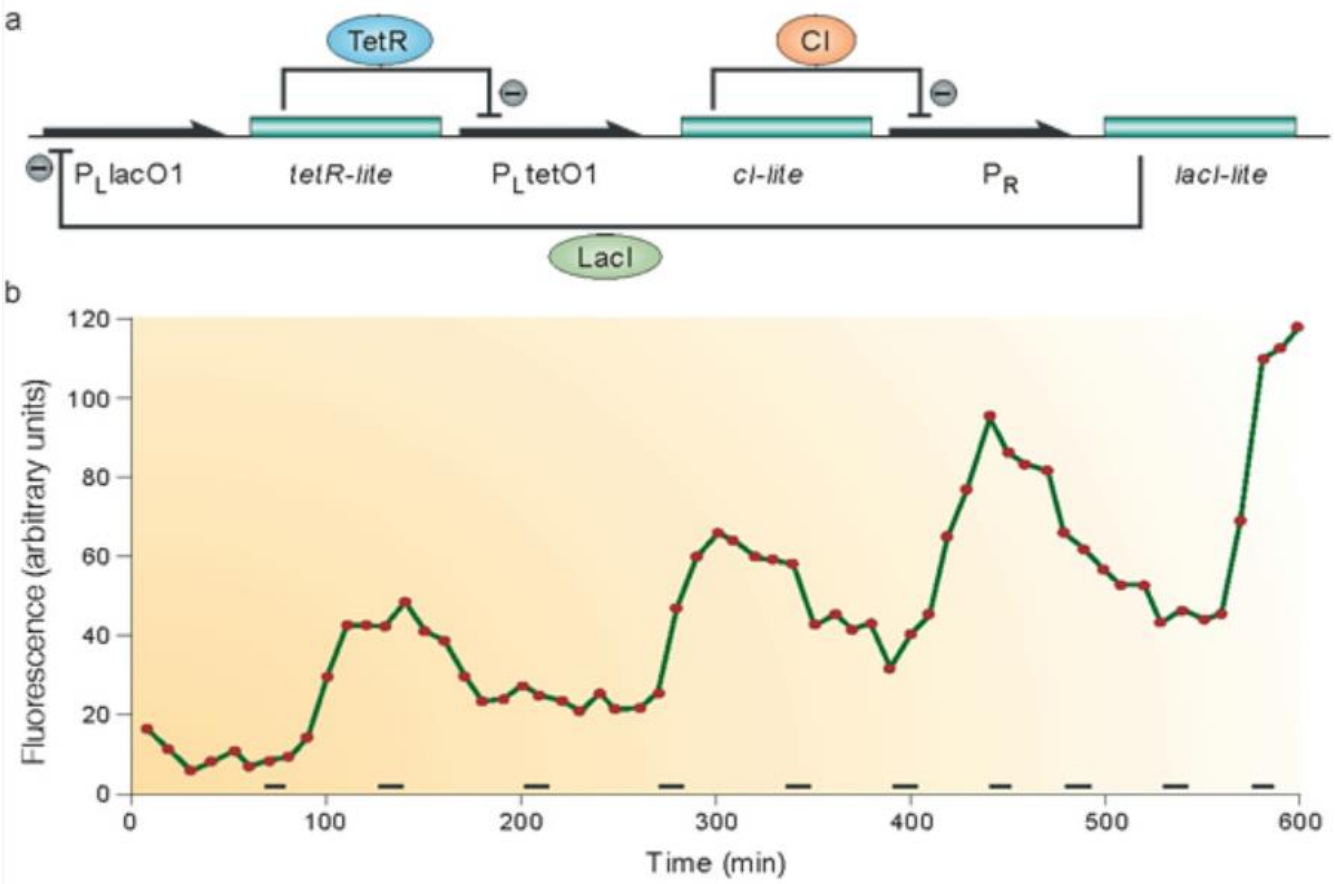
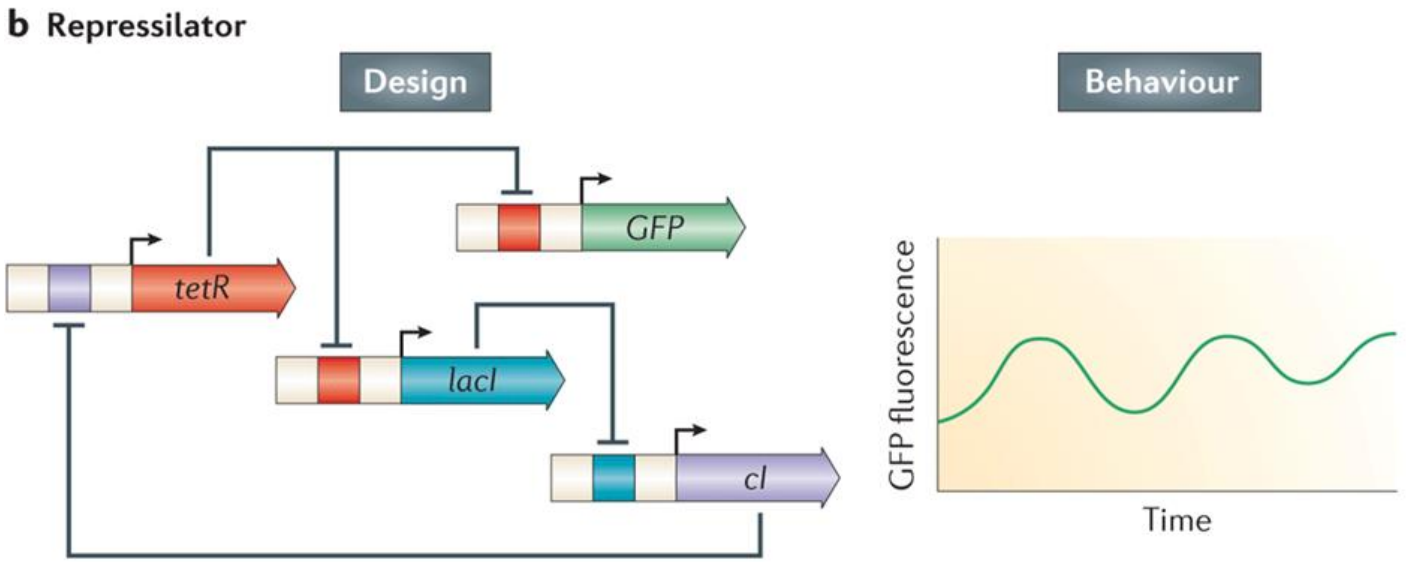


Figure 5 | **The design and application of the repressilator.**
a | Schematic showing the regulation pattern that forms the basis of a repressilator..
b | The luminescence pattern of a reporter plasmid that carries *GFP* under the transcriptional control of the PLtetO1 promoter, when the reporter construct is transferred to an *Escherichia coli* in the presence of the repressilator.

Nature Reviews Genetics 6, 533-543 (July 2005)

Figure 5 | **The design and application of the repressilator. a** | Schematic showing the regulation pattern that forms the basis of a repressilator. Three gene–promoter pairs are arranged so that the product derived from the expression of the gene following a promoter is a repressor for the next promoter in the cycle. Black connecting lines show that promoter PLlacO1 controls the transcription of the gene *tetR-lite*, the tetracycline repressor protein TetR represses PLtetO1, which is the next promoter in the sequence. PLtetO1 in turn controls the transcription of *cI-lite*, and the protein CI represses the promoter PR. Finally, PR controls the expression of *lacI-lite*, and the lactose repressor protein LacI represses PLlacO1, completing the cycle. The suffix '*-lite*' refers to the presence of tags that increase the degradation rate of the proteins. **b** | The luminescence pattern of a reporter plasmid that carries *GFP* under the transcriptional control of the PLtetO1 promoter, when the reporter construct is transferred to an *Escherichia coli* in the presence of the repressilator. As the experimental trace shows, the oscillation of the TetR repressor expressed from the repressilator results in the time dependent oscillation of *GFP* expression. Bars at the bottom of the diagram show the timing of cell division events. The period of the oscillations is longer than the cell division time, and the cycle of oscillations continues in the subsequent generations. Adapted, with permission, from *Nature* REF. 16 © (2000) Macmillan Magazines Ltd.



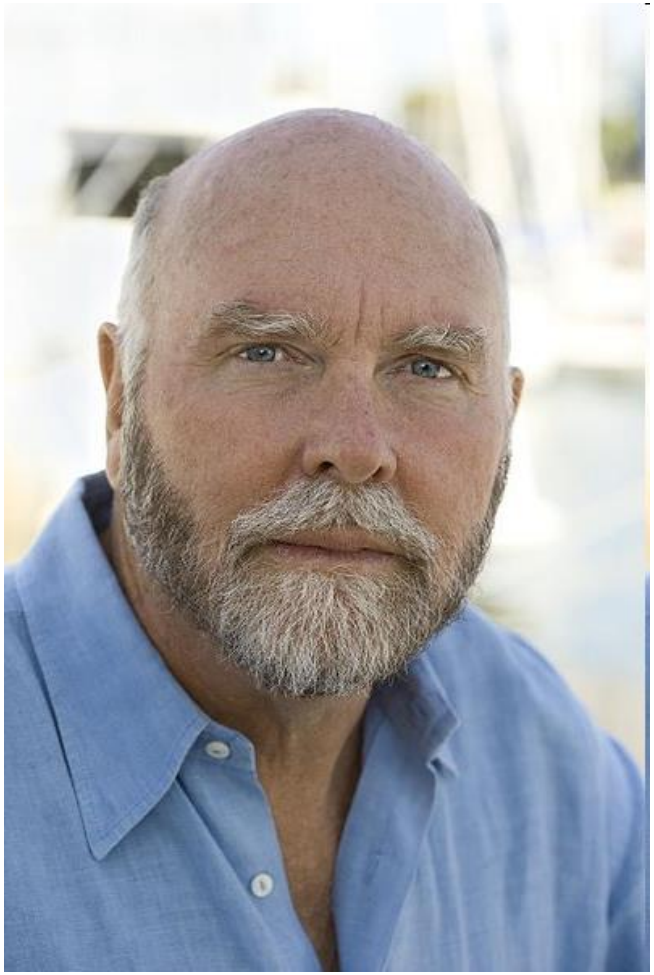
The repressilator. The circuit is constructed from three repressor–promoter interactions (between cI, LacI and TetR repressors and their associated promoters), which are linked together to form a ring-shaped network, in which TetR regulates a GFP-reporter node. When analysed at the single-cell level using time-lapse fluorescence microscopy, the circuit exhibits periodic oscillations in GFP expression, which persist for a number of generations; however, oscillations become dampened after a few periods and are generally noisy, with individual cells showing high variability in both the amplitude and period of their oscillations.

Nature Reviews Microbiology 12, 381–390 (2014) doi:10.1038/nrmicro3239

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,1 John I. Glass, Carole Lartigue, Vladimir N. Noskov, Ray-Yuan Chuang, Mikkel A. Algire, Gwynedd A. Benders, Michael G. Montague, Li Ma, Monzia M. Moodie, Chuck Merryman, Sanjay Vashee, Radha Krishnakumar, Nacyra Assad-Garcia, Cynthia Andrews-Pfannkoch, Evgeniya A. Denisova, Lei Young, Zhi-Qing Qi, Thomas H. Segall-Shapiro, Christopher H. Calvey, Prashanth P. Parmar, Clyde A. Hutchison, Hamilton O. Smith, J. Craig Venter

Science, 329(5987):52-56, **2010**



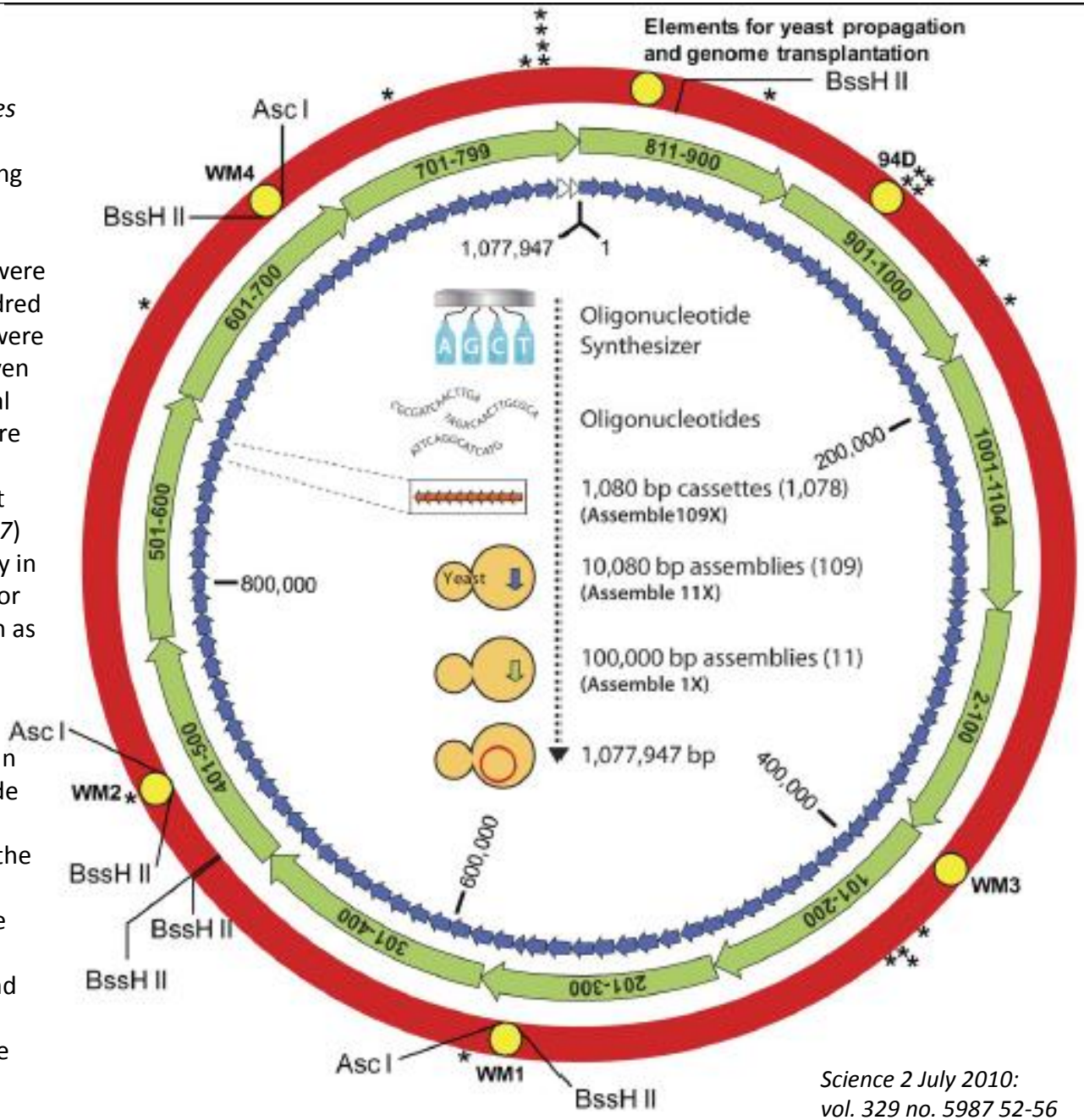


Fig. 1. The assembly of a synthetic *M. mycoides* genome in yeast. A synthetic *M. mycoides* genome was assembled from 1,078 overlapping DNA cassettes in three steps. In the first step, 1,080-bp cassettes (orange arrows), produced from overlapping synthetic oligonucleotides, were recombined in sets of 10 to produce one hundred nine ~10-kb assemblies (blue arrows). These were then recombined in sets of 10 to produce eleven ~100-kb assemblies (green arrows). In the final stage of assembly, these eleven fragments were recombined into the complete genome (red circle). With the exception of 2 constructs that were enzymatically pieced together in vitro (*27*) (white arrows), assemblies were carried out by in vivo homologous recombination in yeast. Major variations from the natural genome are shown as yellow circles. These include 4 watermarked regions (WM1-WM4), a 4-kb region that was intentionally deleted (94D), and elements for growth in yeast and genome transplantation. In addition, there are 20 locations with nucleotide polymorphisms (asterisks). Coordinates of the genome are relative to the first nucleotide of the natural *M. mycoides* sequence. The designed sequence is 1,077,947 bp. The locations of the Asc I and BssH II restriction sites are shown. Cassettes 1 and 800-810 were unnecessary and removed from the assembly strategy (*11*). Cassette 2 overlaps cassette 1104 and cassette 799 overlaps cassette 811.

Science 2 July 2010: vol. 329 no. 5987 52-56

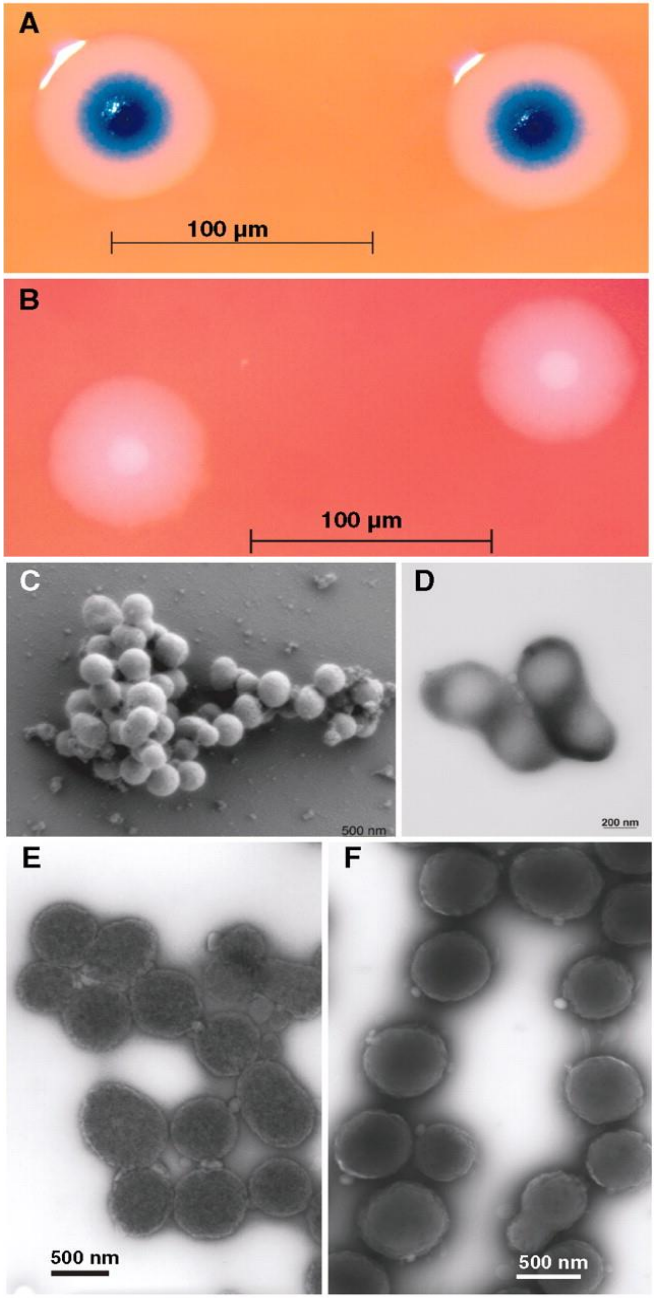
Table 1

Genomes that have been assembled from 11 pieces and successfully transplanted. Assembly 2-100, 1; assembly 101-200, 2; assembly 201-300, 3; assembly 301-400, 4; assembly 401-500, 5; assembly 501-600, 6; assembly 601-700, 7; assembly 701-799, 8; assembly 811-900, 9; assembly 901-1000, 10; assembly 1001-1104, 11. WM, watermarked assembly.

Genome assembly	Synthetic fragments	Natural fragments
Reconstituted natural genome	None	1–11
2/11 semisynthetic genome with one watermark	5 WM, 10	1–4, 6–9, 11
8/11 semisynthetic genome without watermarks	1–4, 6–8, 11	5, 9, 10
9/11 semisynthetic genome without watermarks	1–4, 6–8, 10–11	5, 9
9/11 semisynthetic genome with three watermarks	1, 2 WM, 3 WM, 4, 6, 7 WM, 8, 10–11	5, 9
10/11 semisynthetic genome with three watermarks	1, 2 WM, 3 WM, 4, 5 WM, 6, 7 WM, 8, 10–11	9
11/11 synthetic genome, 811-820 correction of *dnaA*	1, 2 WM, 3 WM, 4, 5 WM, 6, 7 WM, 8, 9–11	None
11/11 synthetic genome, 811-900 correction of *dnaA*	1, 2 WM, 3 WM, 4, 5 WM, 6, 7 WM, 8, 9–11	None

Science 2 July 2010:
vol. 329 no. 5987 52-56

Fig. 5. **Images of *M. mycoides* JCVI-syn1.0 and WT *M. mycoides*.**
To compare the phenotype of the JCVI-syn1.0 and non-YCp WT strains, we examined colony morphology by plating cells on SP4 agar plates containing X-gal. Three days after plating, the JCVI-syn1.0 colonies are blue because the cells contain the *lacZ* gene and express beta-galactosidase, which converts the X-gal to a blue compound (**a**). The WT cells do not contain *lacZ* and remain white (**b**). Both cell types have the fried egg colony morphology characteristic of most mycoplasmas. EMs were made of the JCVI-syn1.0 isolate using two methods. (**c**) For scanning EM, samples were post-fixed in osmium tetroxide, dehydrated and critical point dried with CO2, and visualized using a Hitachi SU6600 SEM at 2.0 keV. (**d**) Negatively stained transmission EMs of dividing cells using 1% uranyl acetate on pure carbon substrate visualized using JEOL 1200EX CTEM at 80 keV.
To examine cell morphology, we compared uranyl acetate stained EMs of *M. mycoides* JCVI-syn1.0 cells (**e**) with EMs of WT cells made in 2006 that were stained with ammonium molybdate (**f**). Both cell types show the same ovoid morphology and general appearance. EMs were provided by Tom Deerinck and Mark Ellisman of the National Center for Microscopy and Imaging Research at the University of California at San Diego.



Science 2 July 2010: vol. 329 no. 5987 52-56

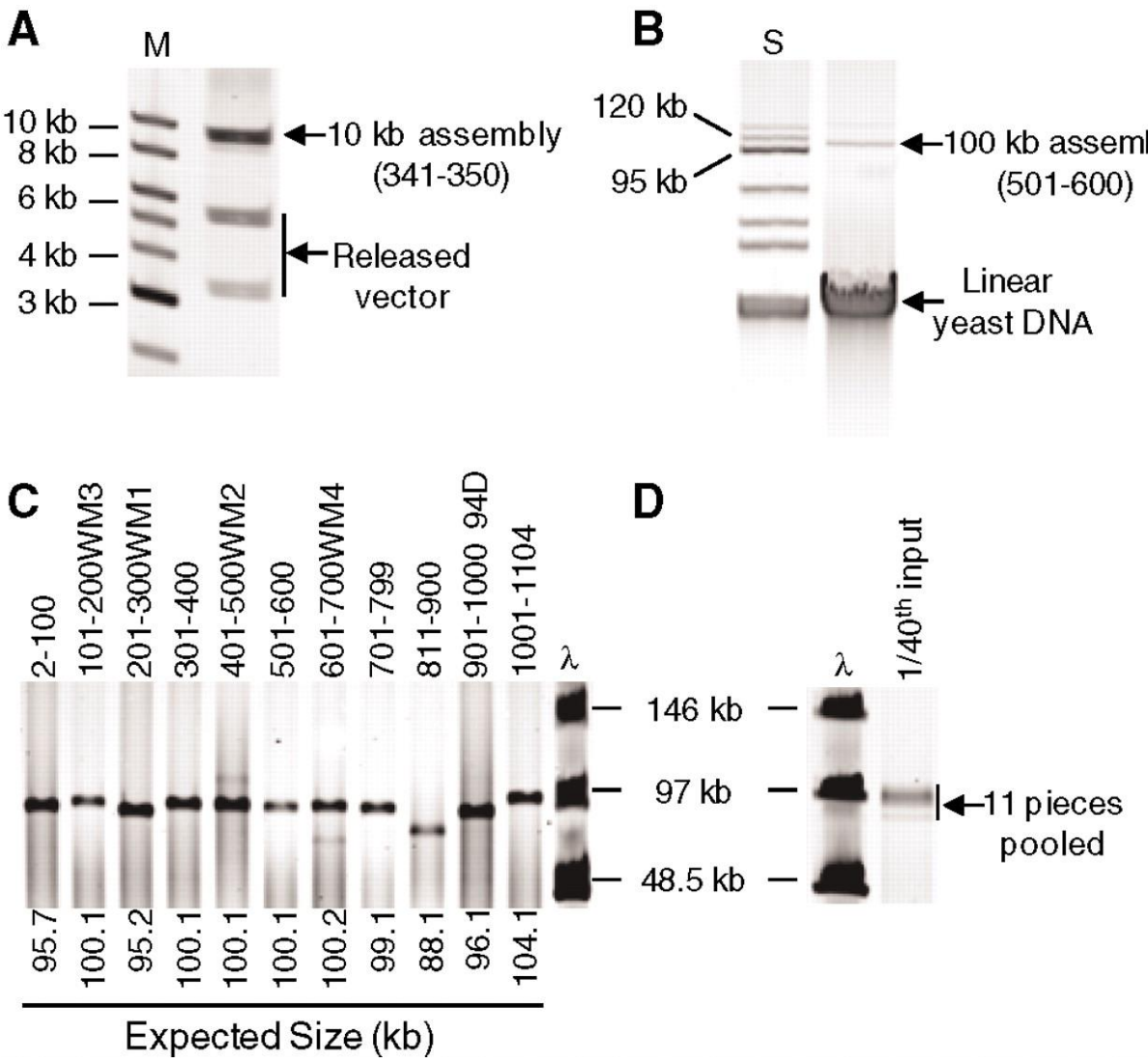
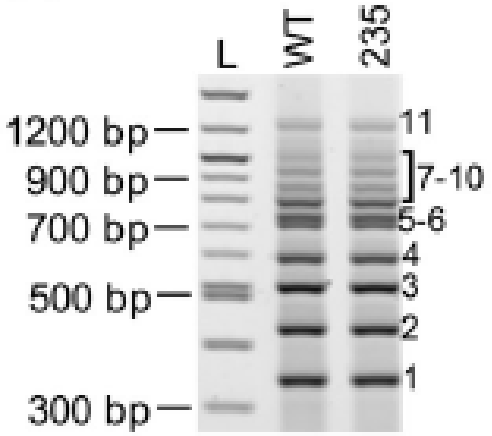


Fig. 2. Analysis of the assembly intermediates. (**a**) Not I and Sbf I double restriction digestion analysis of assembly 341-350 purified from *E. coli*. These restriction enzymes release the vector fragments (5.5 kb and 3.4 kb) from the 10-kb insert. Insert DNA was separated from the vector DNA on a 0.8% E-gel (Invitrogen). M indicates the 1-kb DNA ladder (New England Biolabs; NEB). (**b**) Analysis of assembly 501-600 purified from yeast. The 105-kb circles (100-kb insert plus 5-kb vector) were separated from the linear yeast chromosomal DNA on a 1% agarose gel by applying 4.5 V/cm for 3 hours. S indicates the BAC-Tracker supercoiled DNA ladder (Epicentre). (**c**) Not I restriction digestion analysis of the eleven ~100-kb assemblies purified from yeast. These DNA fragments were analyzed by FIGE on a 1% agarose gel. The expected insert size for each assembly is indicated. λ indicates the lambda ladder (NEB). (**d**) Analysis of the 11 pooled assemblies shown in (**c**) following topological trapping of the circular DNA and Not I digestion. One fortieth of the DNA used to transform yeast is represented.

A



B

Strain	Digest	Fragment # and size (kb)
WT	AscI	No sites
WT	BssHII	(4) 668 (5) 419
Syn235	AscI	(1) 685 (2) 233 (3) 160
Syn235	BssHII	(6) 533 (7) 233 (8) 152 (9) 126 (10) 34

C

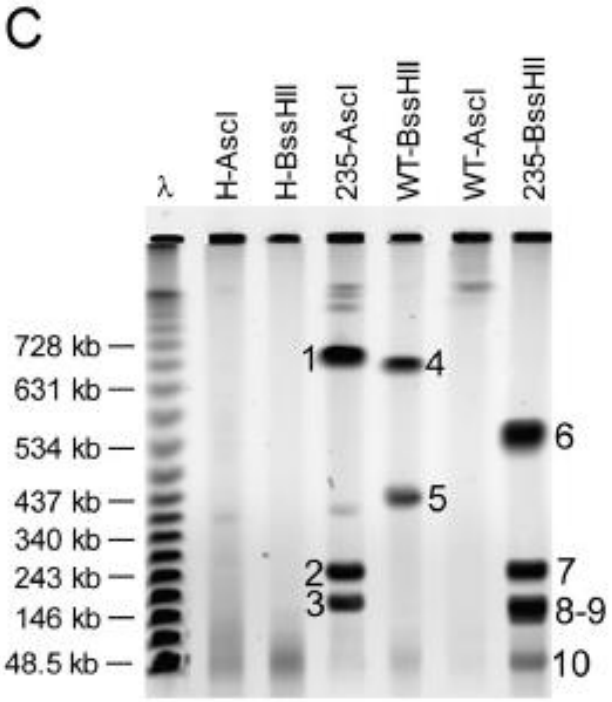
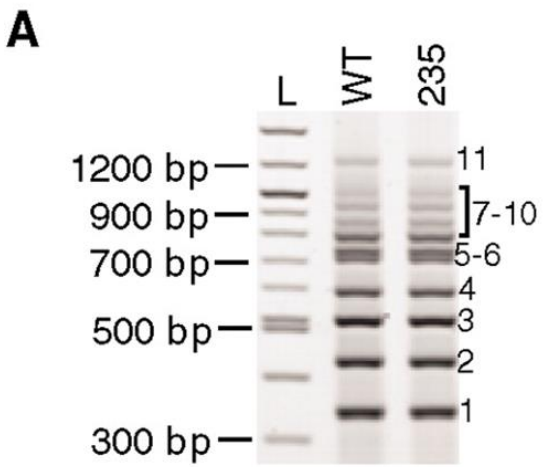


Fig. 3. Characterization of the synthetic genome isolated from yeast. (**a**) Yeast clones containing a completely assembled synthetic genome were screened by multiplex PCR with a primer set that produces 11 amplicons; one at each of the 11 assembly junctions. Yeast clone sMmYCp235 (235) produced the 11 PCR products expected for a complete genome assembly. For comparison, the natural genome extracted from yeast (WT) was also analyzed. PCR products were separated on a 2% E-gel (Invitrogen). L indicates the 100-bp ladder (NEB). (**b**) The sizes of the expected Asc I and BssH II restriction fragments for natural (WT) and synthetic (Syn235) *M. mycoides* genomes. (**c**) Natural (WT) and synthetic (235) *M. mycoides* genomes were isolated from yeast in agarose plugs. In addition, DNA was purified from the host strain alone (H). Agarose plugs were digested with Asc I or BssH II and fragments were separated by clamped homogeneous electrical field (CHEF) gel electrophoresis. Restriction fragments corresponding to the correct sizes are indicated by the fragment numbers shown in (**b**).

Science 2 July 2010: vol. 329 no. 5987 52-56



B

Strain	Digest	Fragment # and size (kb)
WT	AscI	No sites
WT	BssHII	(4) 668 (5) 419
Syn235	AscI	(1) 685 (2) 233 (3) 160
Syn235	BssHII	(6) 533 (7) 233 (8) 152 (9) 126 (10) 34

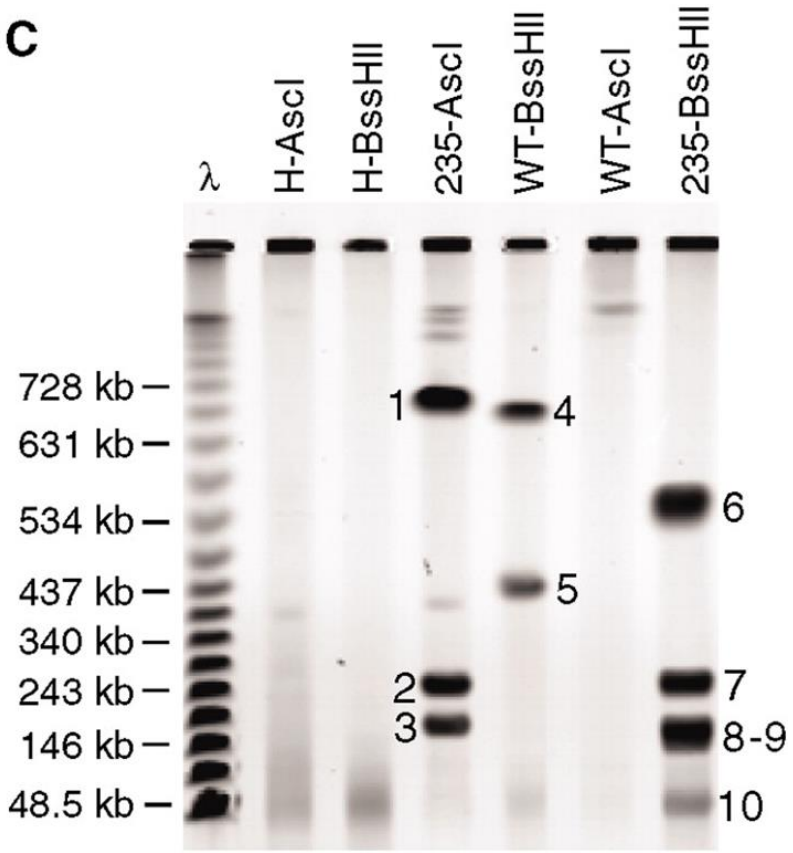


Fig. 3. **Characterization of the synthetic genome isolated from yeast.** (a) Yeast clones containing a completely assembled synthetic genome were screened by multiplex PCR with a primer set that produces 11 amplicons; one at each of the 11 assembly junctions. Yeast clone sMmYCp235 (235) produced the 11 PCR products expected for a complete genome assembly. For comparison, the natural genome extracted from yeast (WT) was also analyzed. PCR products were separated on a 2% E-gel (Invitrogen). L indicates the 100-bp ladder (NEB). (b) The sizes of the expected Asc I and BssH II restriction fragments for natural (WT) and synthetic (Syn235) M. mycoides genomes. (c) Natural (WT) and synthetic (235) M. mycoides genomes were isolated from yeast in agarose plugs. In addition, DNA was purified from the host strain alone (H). Agarose plugs were digested with Asc I or BssH II and fragments were separated by clamped homogeneous electrical field (CHEF) gel electrophoresis. Restriction fragments corresponding to the correct sizes areindicated by the fragment numbers shown in (b).

Science 2 July 2010: vol. 329 no. 5987 52-56

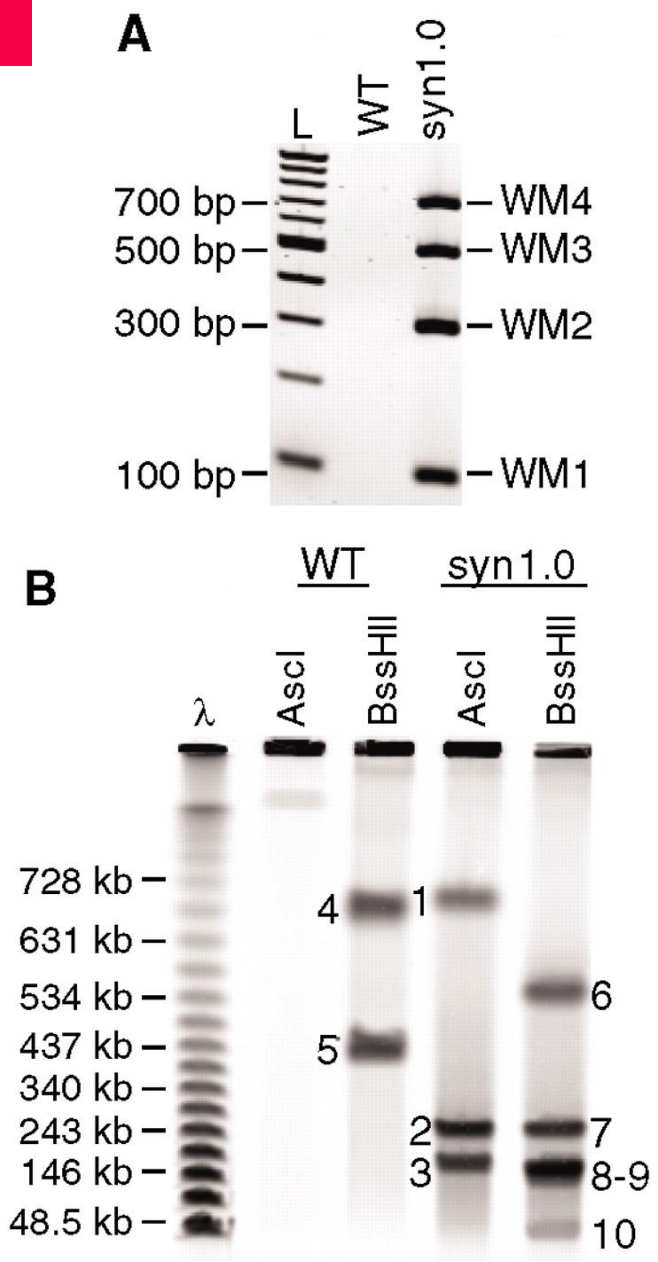
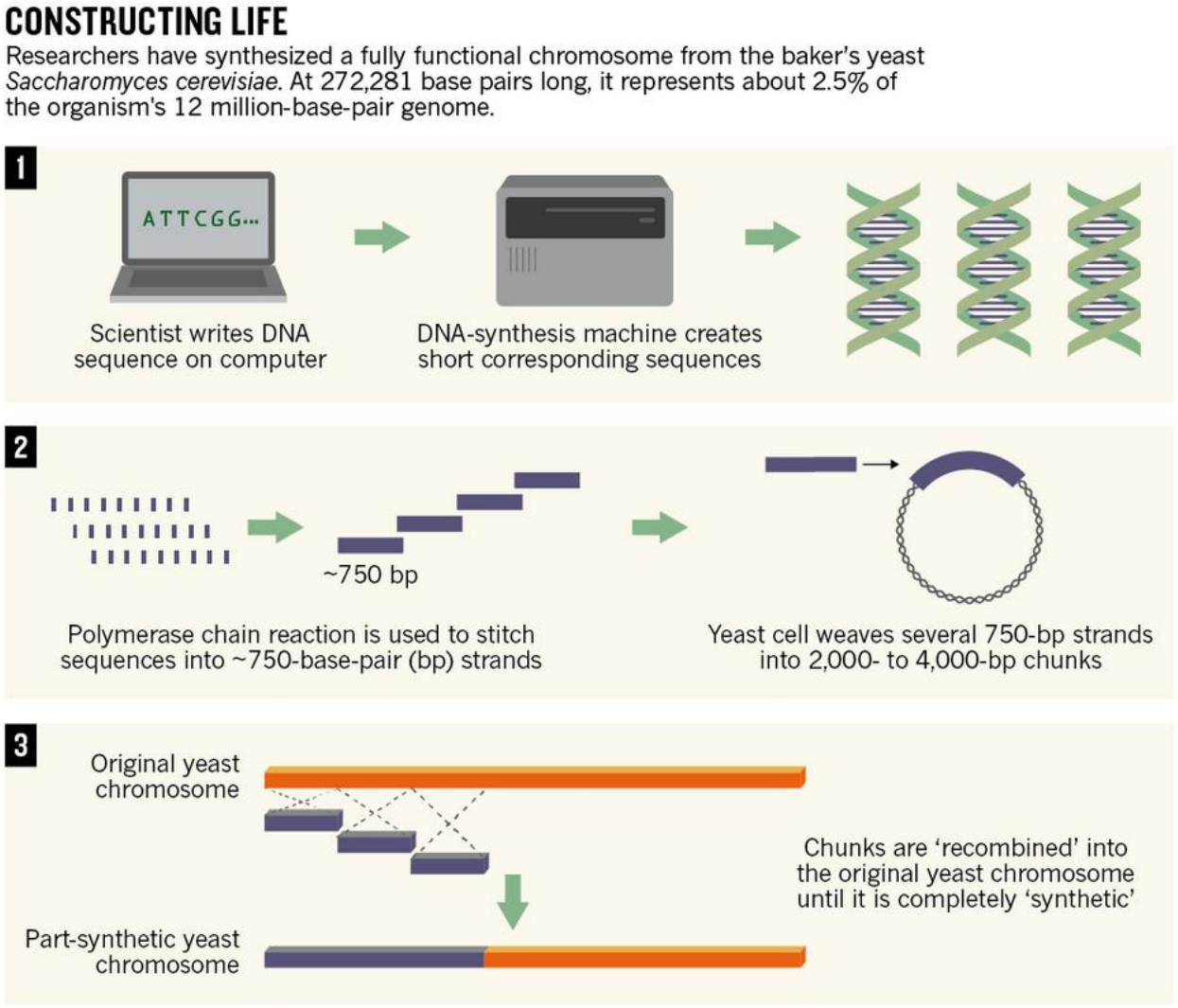


Fig. 4. **Characterization of the transplants.** (**a**) Transplants containing a synthetic genome were screened by multiplex PCR with a primer set that produces 4 amplicons; one internal to each of the four watermarks. One transplant (syn1.0) originating from yeast clone sMmYCp235 was analyzed alongside a natural, non-synthetic genome (WT) transplanted out of yeast. The transplant containing the synthetic genome produced the 4 PCR products whereas the WT genome did not produce any. PCR products were separated on a 2% E-gel (Invitrogen). (**b**) Natural (WT) and synthetic (syn1.0) *M. mycoides* genomes were isolated from *M. mycoides* transplants in agarose plugs. Agarose plugs were digested with Asc I or BssH II and fragments were separated by CHEF gel electrophoresis. Restriction fragments corresponding to the correct sizes are indicated by the fragment numbers shown in Fig. 3b.

Science 2 July 2010: vol. 329 no. 5987 52-56

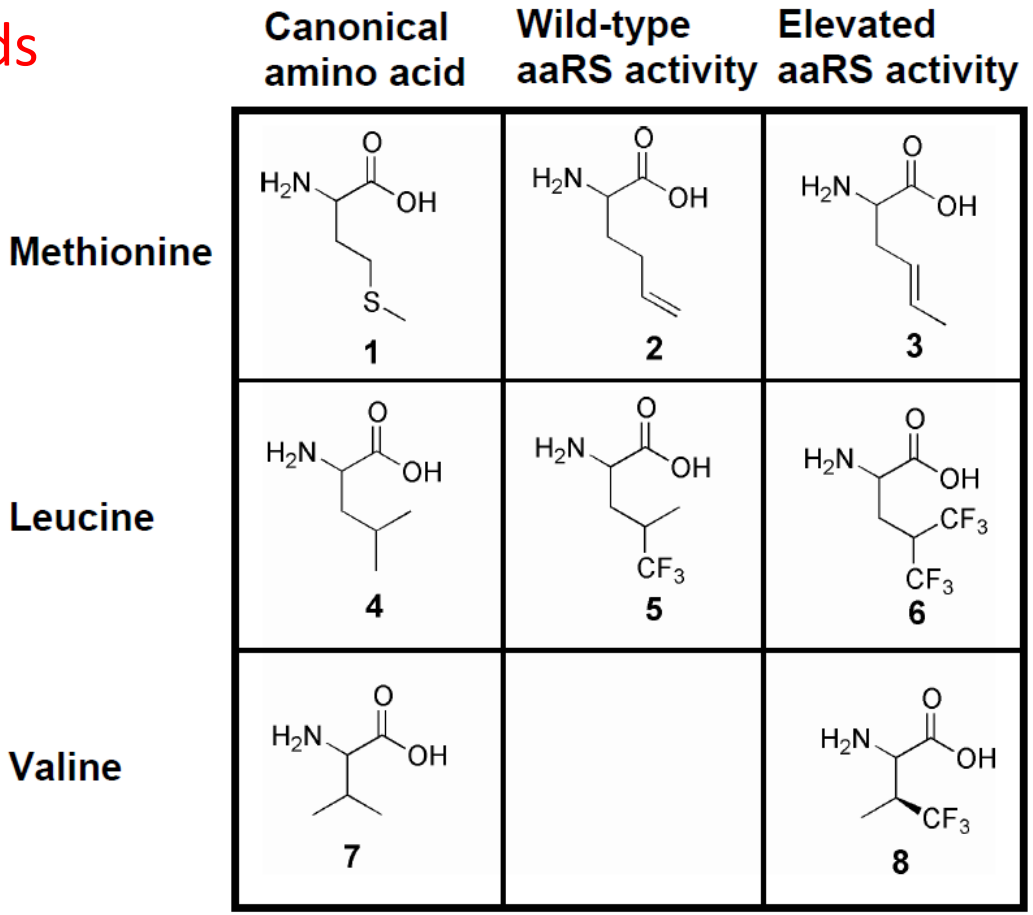
First synthetic yeast chromosome revealed

US-based project recruited dozens of undergraduates to stitch DNA fragments together.



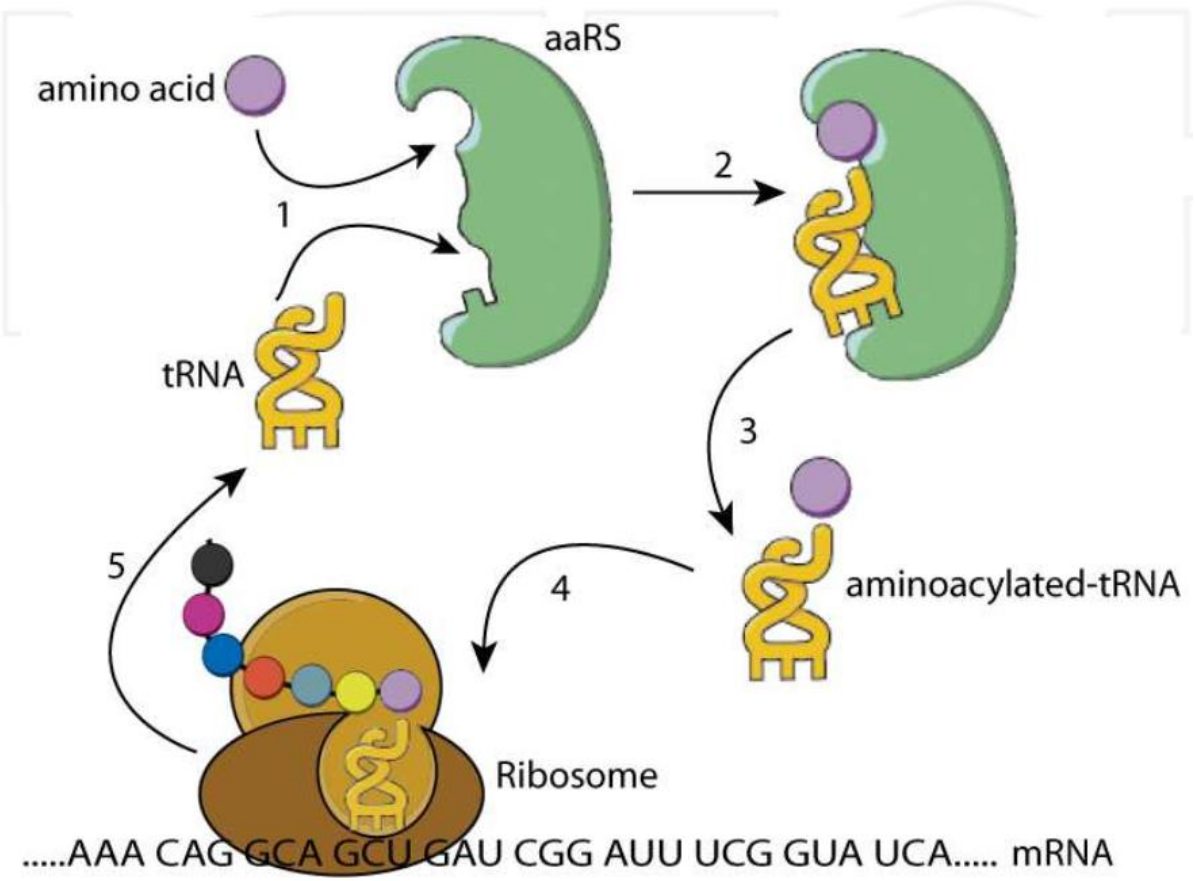
Nature doi:10.1038/nature.2014.14941

Non-canonical Amino Acids in Polypeptides



Sense codon reassignment by increasing aminoacyl-tRNA synthetase (aaRS) activity. Some non-canonical amino acids such as homoallylglycine (2) and trifluoroleucine (5) can replace their canonical counterparts in *E. coli* with wild-type aaRS activity. Elevating the appropriate aaRS activity of the cell results in the ability to incorporate analogues such as trans-crotylglycine (3), hexafluoroleucine (6), and trifluorovaline (8).

Link, A. J.; Mock, M. L.; Tirrell, D. A. Current Opinion in Biotechnology 2003, 14, 603-609 and Link, A. J.; Tirrell, D. A. Methods 2005, 36, 291-298.



Natural protein translation system. Aminoacyl-tRNA synthetases (aaRSs) play a key role in maintaining the fidelity of protein translation by demonstrating strict specificity for their cognate amino acid and their tRNA. (1) Each aaRS recognizes unique structural elements of its cognate tRNA(s) and amino acid. (2) The aaRS joins the amino acid to the tRNA, forming the aminoacyl-tRNA (3) that is released into the cytosol. (4) As the ribosome moves along an mRNA molecule, codons are paired with aminoacylated tRNAs containing the complementary anticodons. The ribosome catalyses the sequential peptide-bond formation of amino acids delivered by tRNAs, generating the nascent protein. At the end of translation, release factors occupy stop codons and trigger the termination of protein synthesis. (5) Deaminoacylated tRNA is recycled for further use.

Protein Engineering with Non-Natural Amino Acids; Aijun Wang, Natalie Winblade Nairn, Marcello Marelli and Kenneth Grabstein, Allozyne; USA ; www.intechopen.com

Methods for incorporation of NNAAs into recombinant proteins

Reassigned sense codon

The simplest method to introduce NNAAs into proteins exploits the promiscuity of aaRSs of the host cells. The successful substitution of a canonical amino acid with a NNAA relies on the use of auxotrophic expression hosts deficient in the biosynthesis of that amino acid. Employment of such hosts limits competition from the canonical amino acid for the reassigned sense codon

Wobble codon

Bias codon

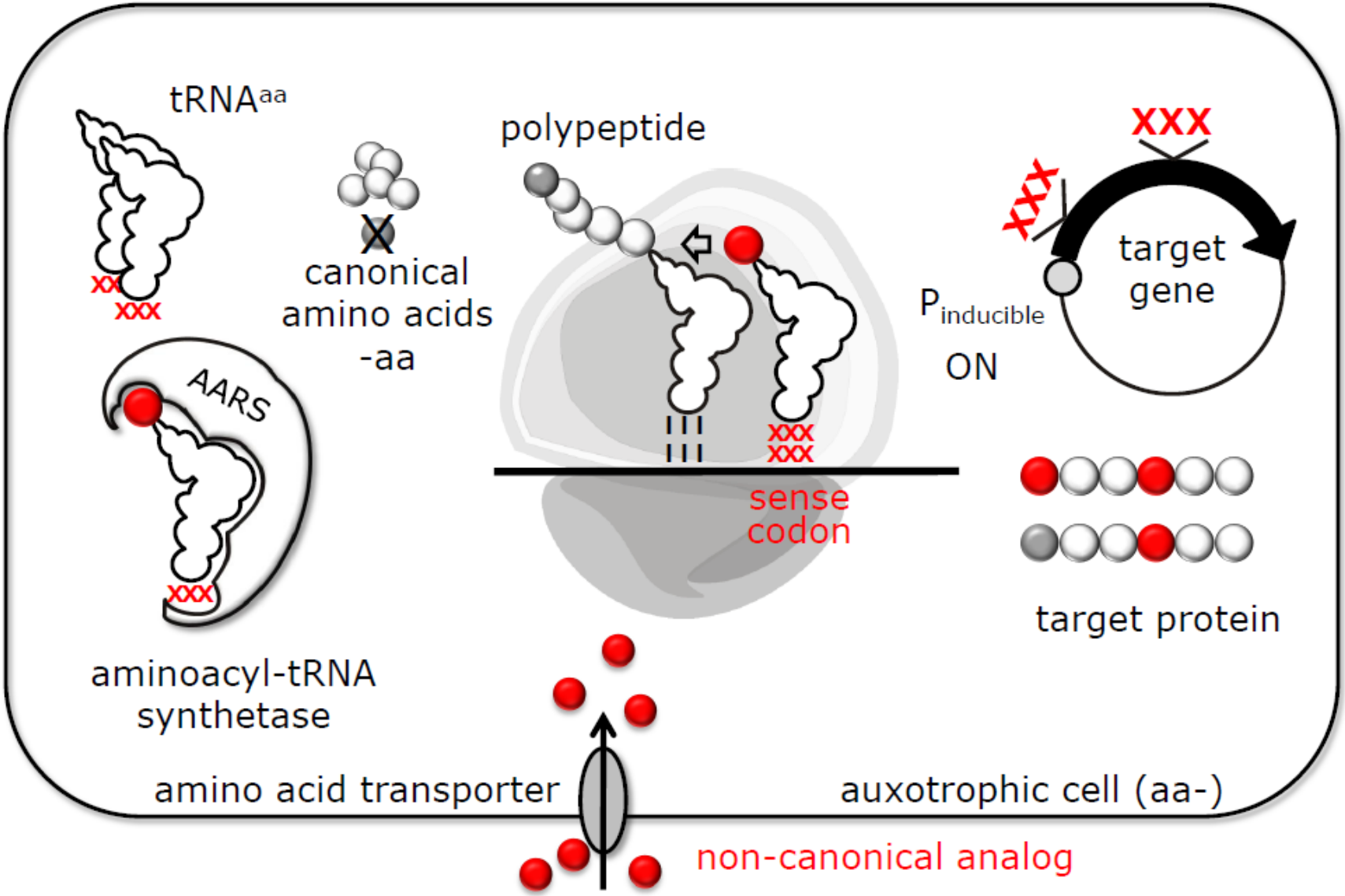
degenerate codons are not used with equal frequency

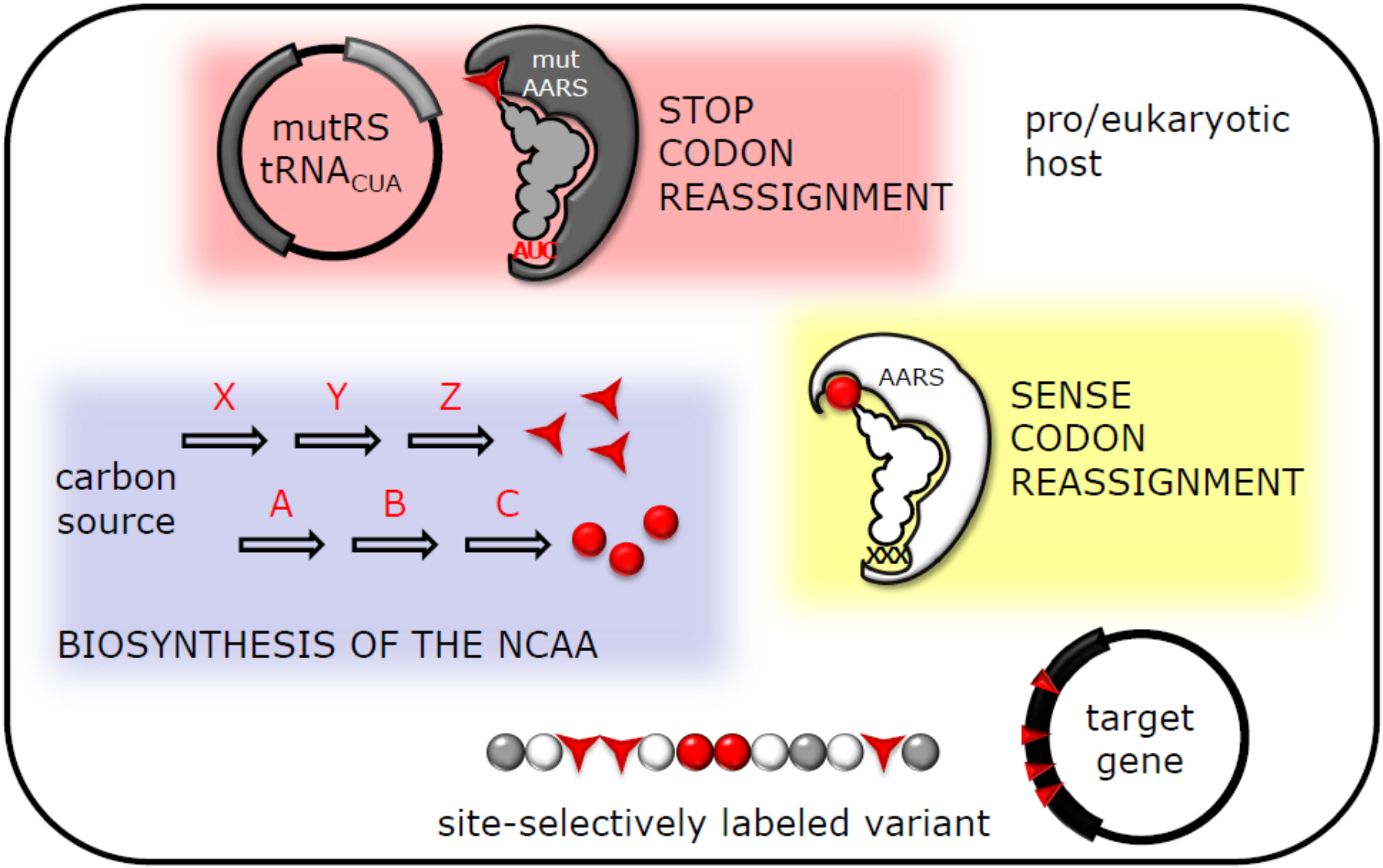
Stop codon

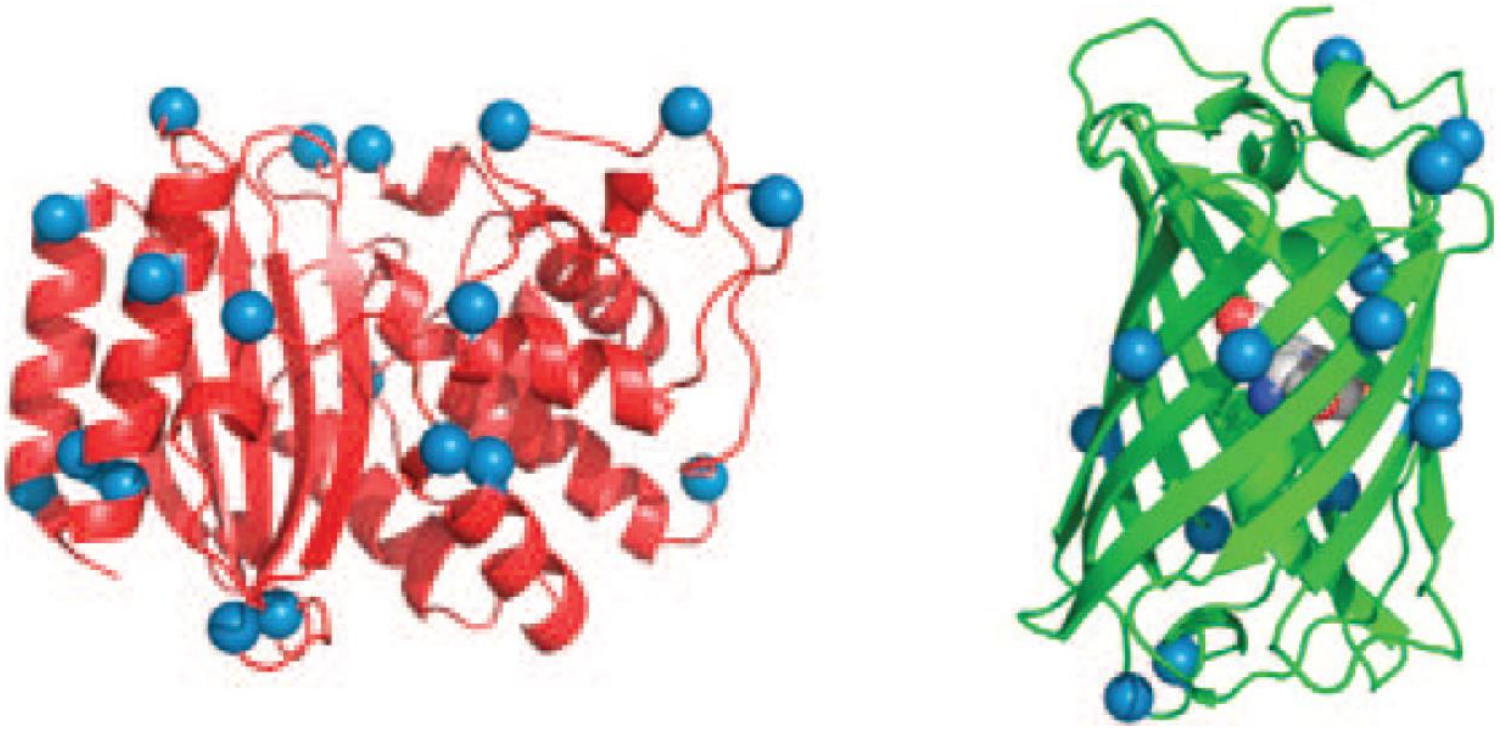
Four-base-pair codon or non-natural base pair

Reserved codon in artificial cells

Sense codon reassignment



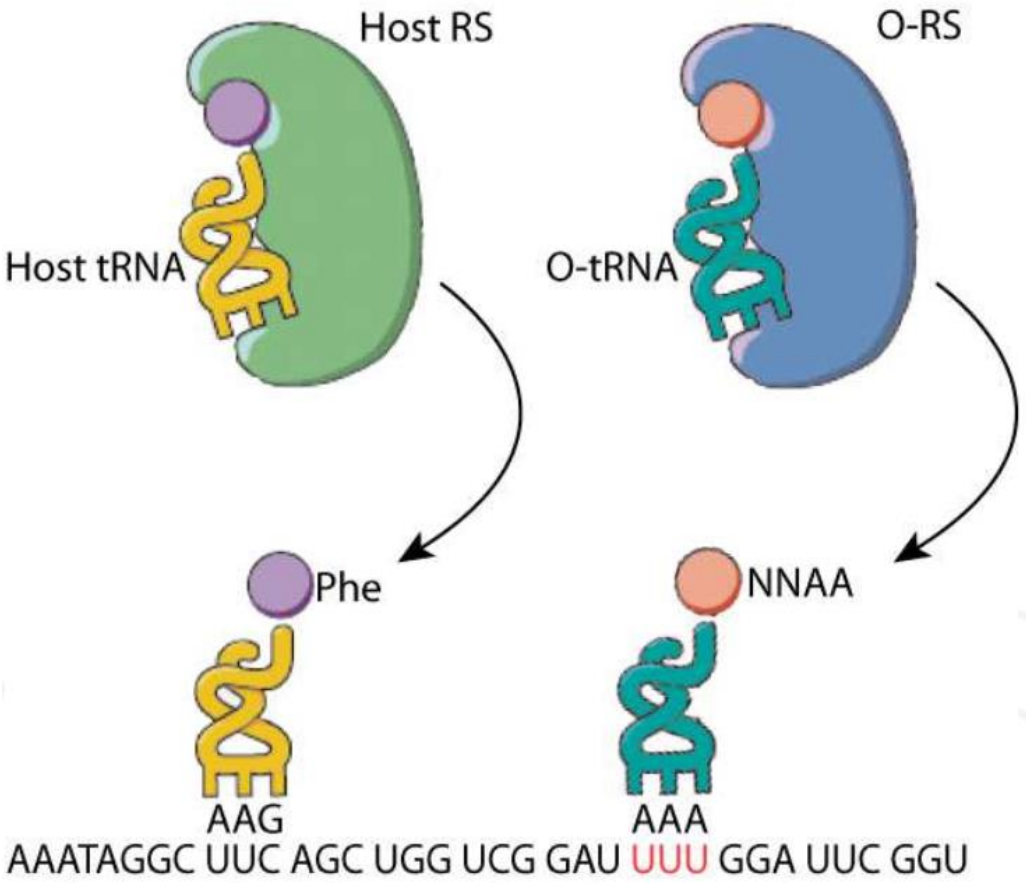

mutRS
tRNA$_{CUA}$
mut
AARS
AUC
STOP
CODON
REASSIGNMENT
pro/eukaryotic
host
X
Y
Z
carbon
source
A
B
C
BIOSYNTHESIS OF THE NCAA
AARS
XXX
SENSE
CODON
REASSIGNMENT
target
gene
site-selectively labeled variant



The result of screening TriNEx TAG-libraries. Spheres indicate where the noncanonical amino acid was tolerated. The full sequence space was not investigated, and the results shown emerged from just 80 sequenced clones. Note the significant distribution of TAG codons throughout the genes.

Current Opinion in Structural Biology 2011, 21:481–487

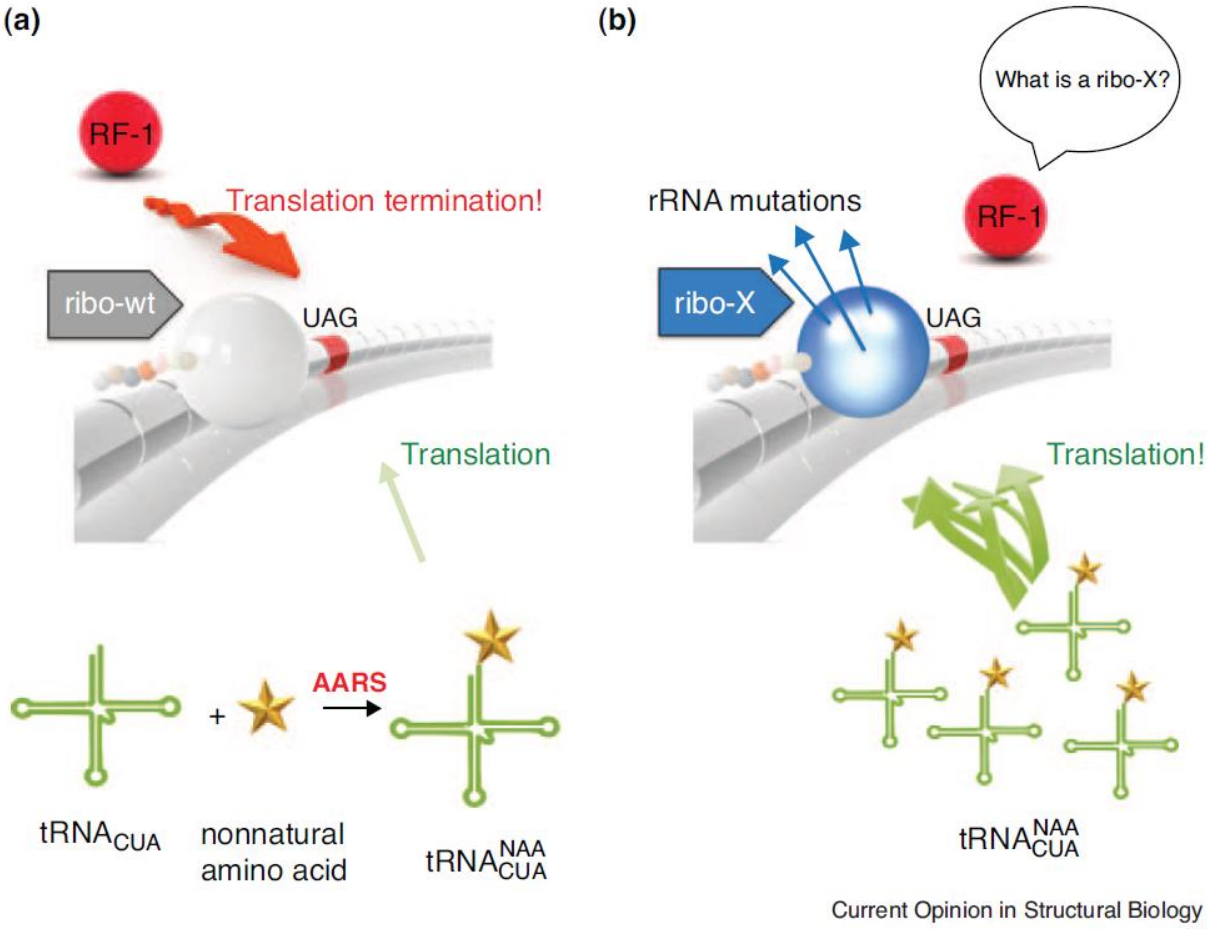
Wobble codon



Site-specific integration of NNAAs at wobble codons. Phe is normally encoded by two codons UUC and UUU, with both codons recognized by the tRNAGAA. By expressing an orthogonal aaRS/tRNAAAA (O-RS and O-tRNA) pair, with specificity for a NNAA and containing the “AAA” anticodon, efficient introduction of the NNAA at UUU codons can be achieved.

Protein Engineering with Non-Natural Amino Acids; Aijun Wang, Natalie Winblade Nairn, Marcello Marelli and Kenneth Grabstein, Allozyne; USA ; www.intechopen.com

Stop codon



(a) Schematic representation illustrating the competitive process resulting from the recoding of UAG as a sense codon using evolved aaRS/$tRNA_{CUA}$ pairs; note the use of a wild type ribosome. (b) The orthogonal ribosomes (ribo-X) recognize unique Shine–Dalgarno initiator sequences (not shown) in the mRNA. Increased UAG suppression, relative to the scheme shown in (a) is observed. Decreased recognition between ribo-X and release factor (RF-1) has been offered as an explanation.

Current Opinion in Structural Biology 2011, 21:481–487

The pyrrolysine translational machinery as a genetic-code expansion tool

Tomasz Fekner[1], Michael K Chan[1, 2]

In 2002, the internal UAG (amber) codon present in the monomethylamine methyltransferase gene of *Methanosarcina barkeri* [1], a methanogenic archaeon, was shown to encode pyrrolysine (**1**, Pyl, Figure 1) [2 and 3], establishing it as the 22nd proteinogenic amino acid [4 and 5]. Apart from the engagement of UAG which is usually a nonsense (stop) codon, pyrrolysine behaves like a typical canonical amino acid and is charged directly into its cognate amber suppressor tRNA ($tRNA^{Pyl}$) by its own pyrrolysyl-tRNA synthetase (PylRS) [6 and 7].

Current Opinion in Chemical Biology

Volume 15, Issue 3, June 2011, Pages 387–391

Molecular Diversity

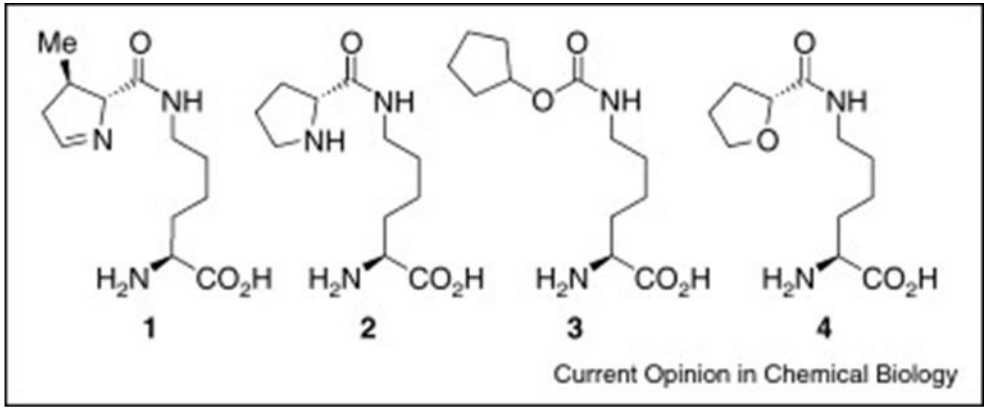
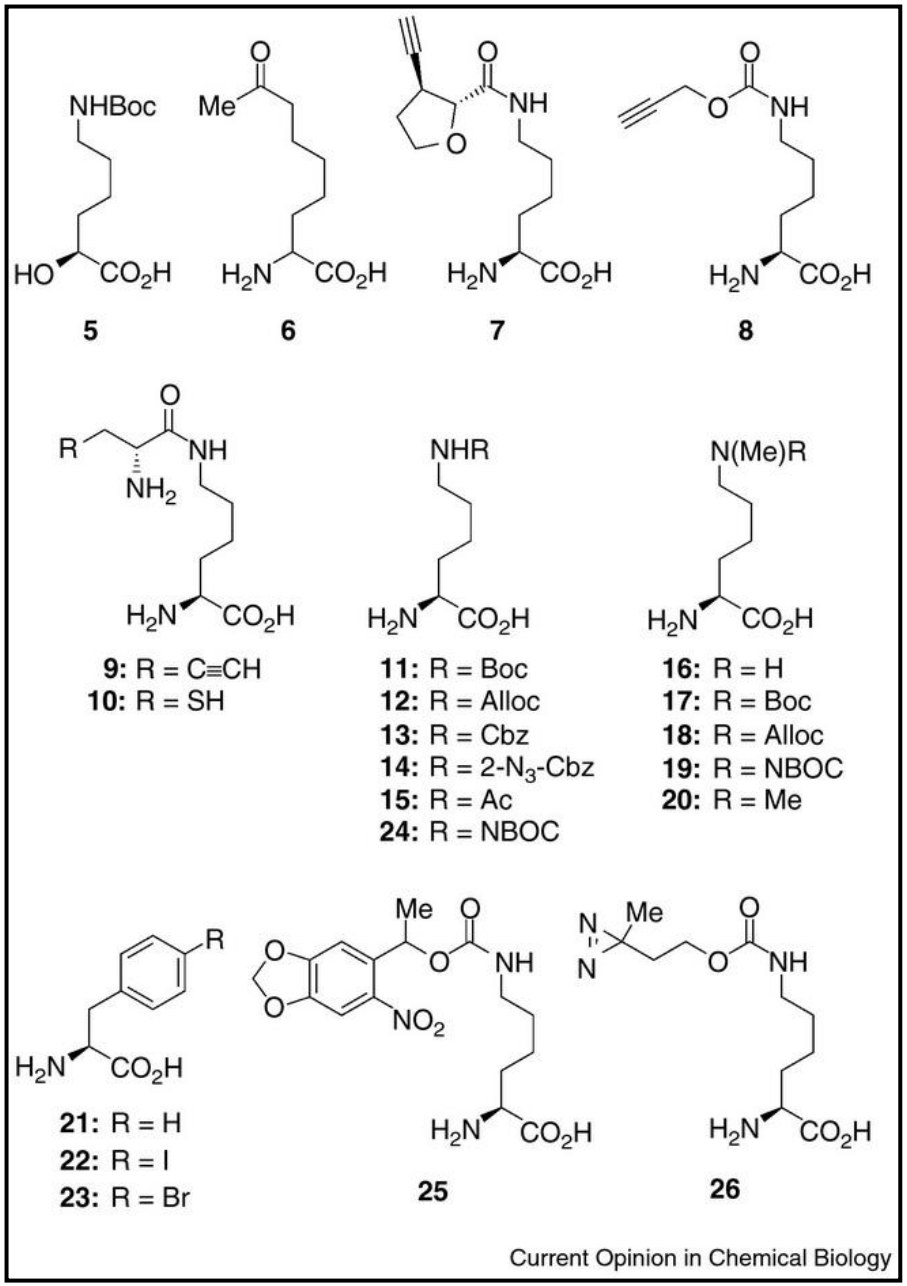
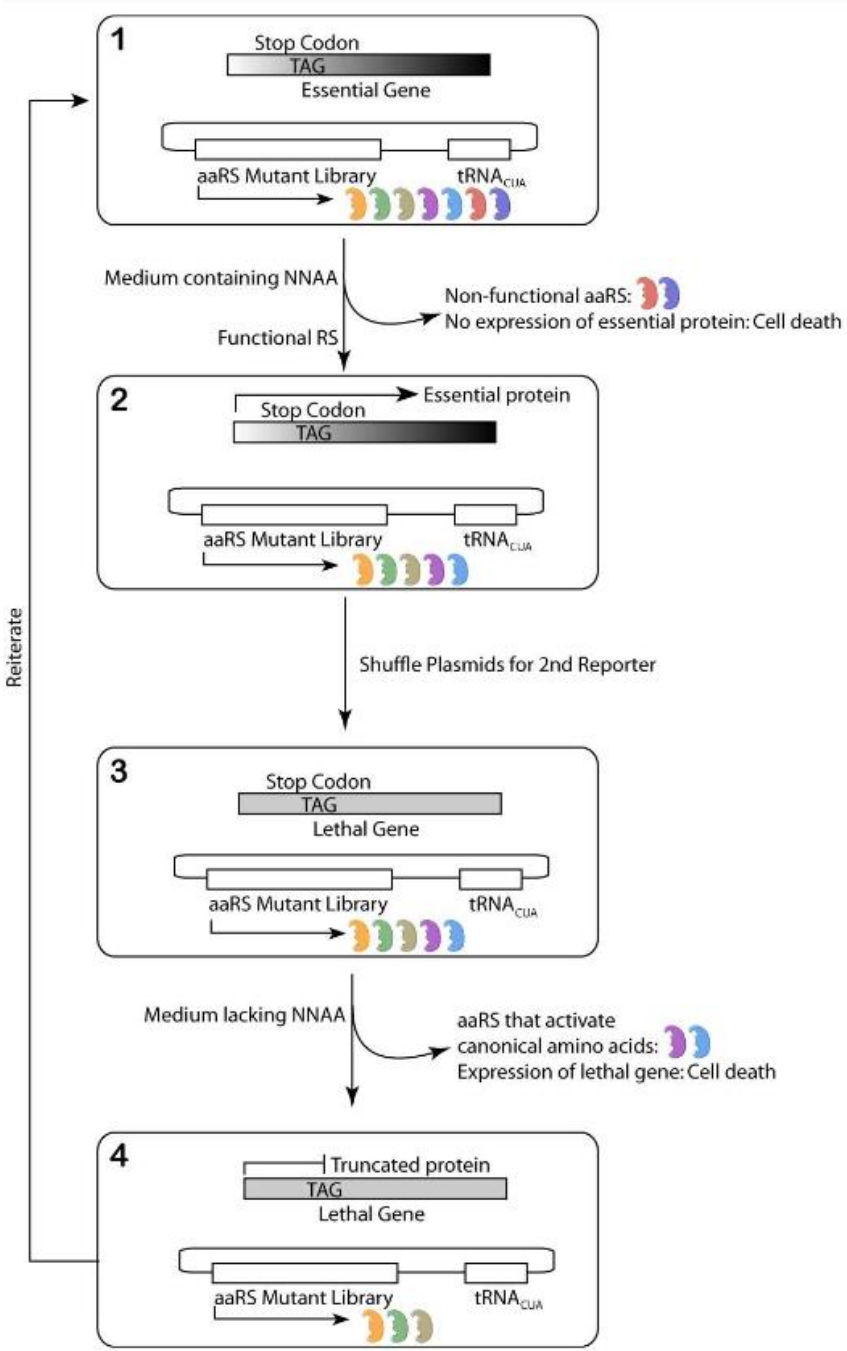


Figure 1 Pyrrolysine (1) and its original surrogates 2 – 4 .

Figure 2.

Representative examples of compounds incorporated into recombinant proteins using the pyrrolysine incorporation system (MmPylRS–tRNAPyl for 5, 7, 9–14, 19, 21–24; MbPylRS–tRNAPyl for 6, 8, 15, 17–18, 25–26; MbPylRS–MmtRNAPyl for 19. Note that 16 and 20 were not incorporated directly and that two systems were used for 19). Boc: tert-butoxycarbonyl, Alloc: allyloxycarbonyl, Cbz: benzyloxycarbonyl, 2-N3-Cbz: 2-azidobenzyloxycarbonyl, Ac: acetyl, and NBOC: 2-nitrobenzyloxycarbonyl.

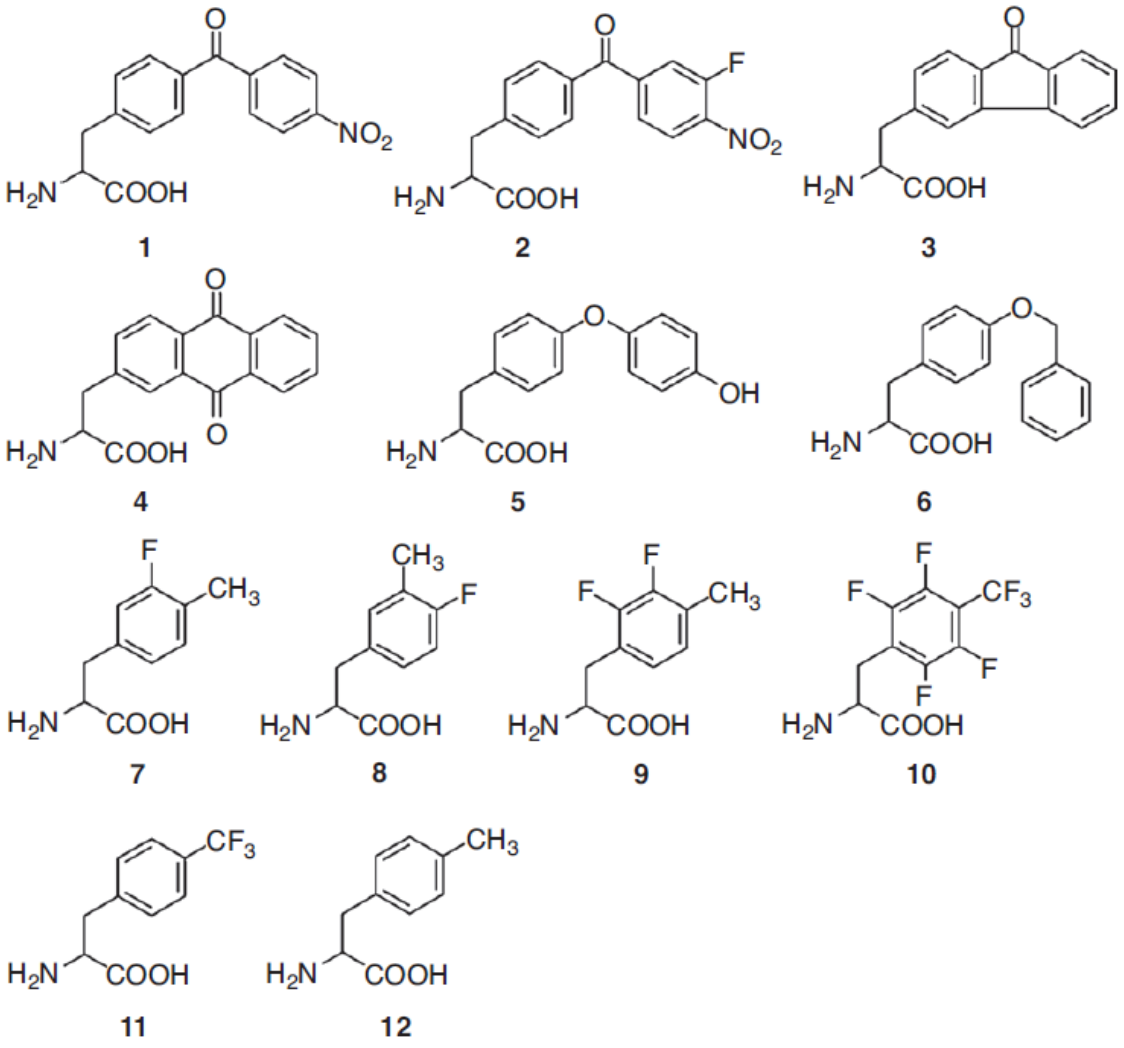




Engineering of aa-tRNA Synthetases

An alternating positive and negative screen for evolving orthogonal aaRSs and tRNAs. (1) A library of aaRS mutants and its cognate tRNACUA are introduced into an expression host containing a conditionally essential gene (e.g. antibiotic resistance marker) with an amber stop codon interrupting its open reading frame and grown in medium containing NNAA. (2) Cells containing functional aaRSs express the full-length essential gene and are able to survive under selective conditions (e.g., presence of antibiotic). (3) Surviving cells are harvested and the aaRS/tRNACUA variants are introduced into cells containing a gene (with an amber codon) that when expressed is lethal (e.g. Barnase), and grown in the absence of NNAA. (4) Cells containing orthogonal NNAA aaRS/tRNACUA pairs express truncated non-functional Barnase and get selected for the next round of screening.

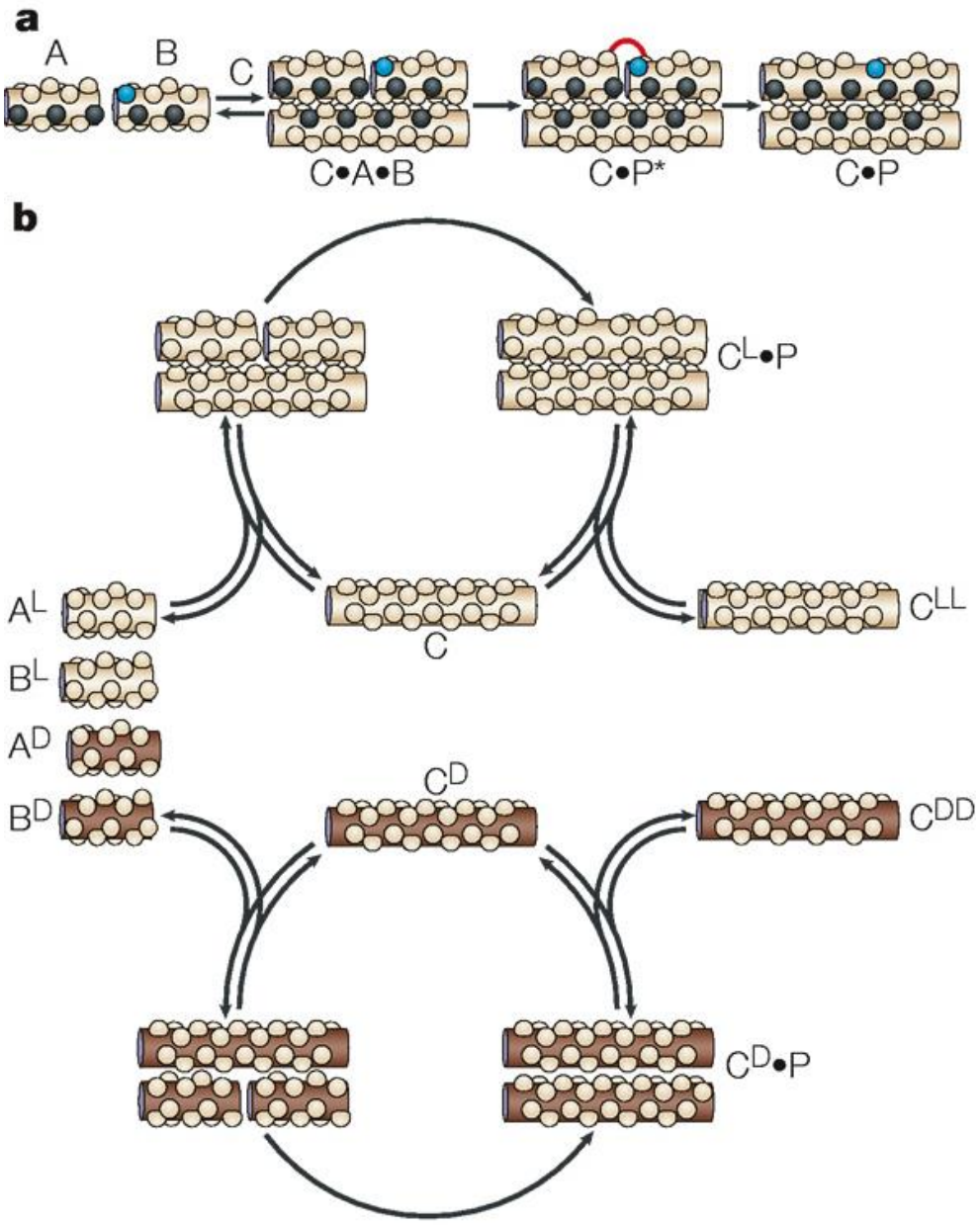
Protein Engineering with Non-Natural Amino Acids; Aijun Wang, Natalie Winblade Nairn, Marcello Marelli and Kenneth Grabstein, Allozyne; USA ; www.intechopen.com



Noncanonical amino acids encoded using aaRSs resulting from screening for promiscuity of previously evolved aaRSs directed evolution.

Current Opinion in Structural Biology 2011, 21:481–487

Figure 3 | **A self-templating system built from peptide units.**
a | A ***de novo* designed peptide ligase**. α-helical peptides A and B bind to the electrostatically complementary α-helical peptide C to form the C•A•B ternary complex, which is composed of two coiled coils. **Peptide A** has a **modified amino terminus** that reacts with the **chemically modified carboxyl termi**nus of peptide B on formation of the ternary complex. The reaction of peptides A and B is a ligation that forms the product C•P (C•P* represents the chemical reaction between A and B to produce P).
b |
Peptide replicator schematic based on the reaction illustrated in part **a**, showing the reaction of peptide A with peptide B on formation of a ternary complex with peptide C. Peptides AL, BL, and CLL are composed of **L-amino acids**, whereas peptides AD, BD, and CDD are composed of **D-amino acids**. Peptides CLL and CDD are produced autocatalytically in a template-directed fashion through the reaction of precursors AL with BL, and AD with BD, respectively. Therefore, this replicator is stereochemically selective, only producing products (CLL and CDD) that are isomerically pure. Part **b** modified, with permission, from REF. 65 © American ChemicalSociety (2001) and from *Nature* REF. 66 © (2001) Macmillan Magazines Ltd.



Nature Reviews Genetics 6, 533-543 (July 2005)

Genome Engineering

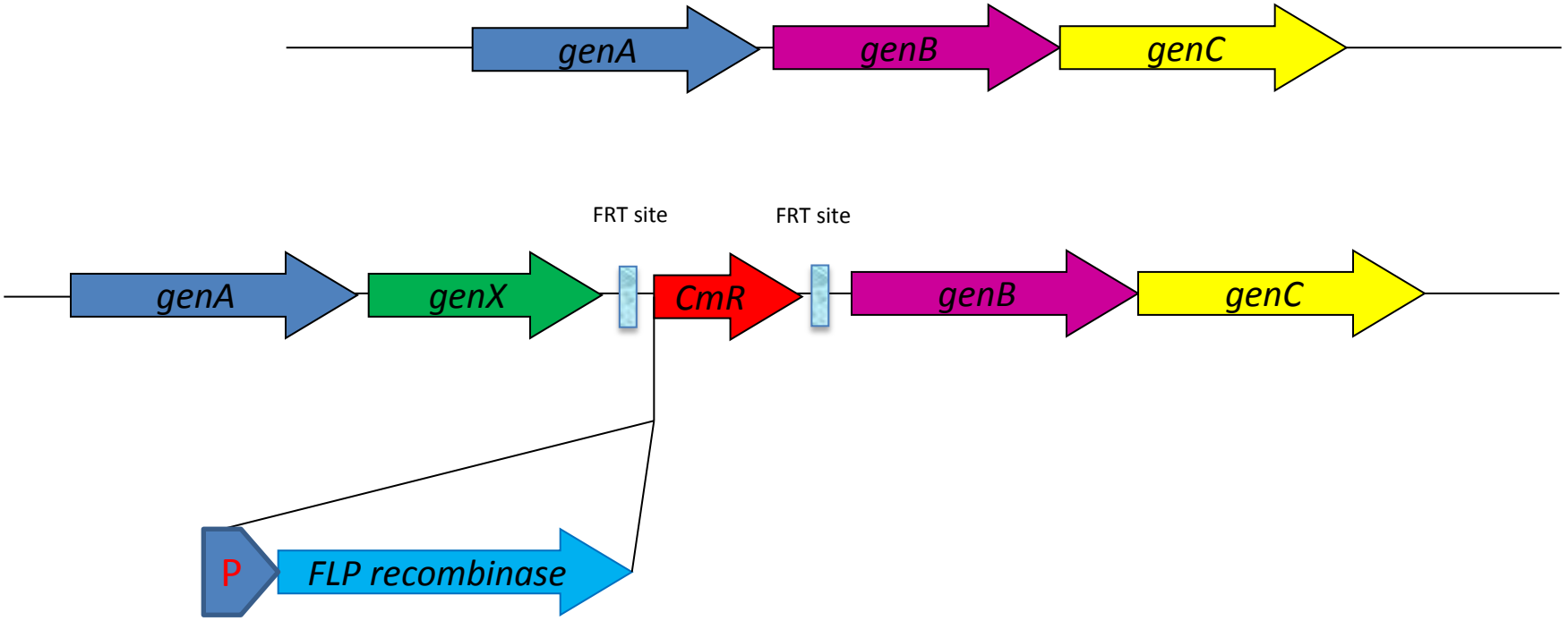
FRT/ Flp

- Marker free deletion or integration

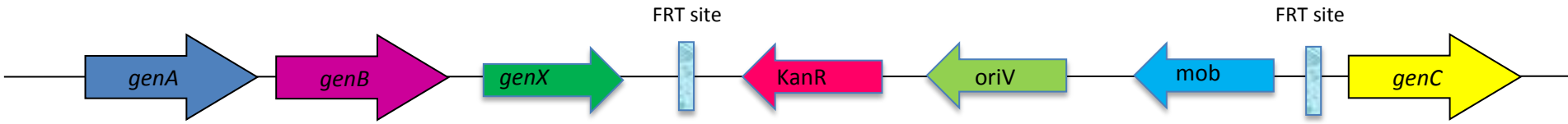
CRISPR-Cas9 system

- CRISPR: clustered regularly interspaced short palindromic repeats
- CRISPR systems are adaptable immune mechanisms used by bacteria to protect themselves from foreign nucleic acids (viruses, plasmids)
- CRISPR/Cas9 system from *S. pyogenes* → adapted for inducing sequnce specific double strand breaks and targeted genome editing

FRT/ Flp



Inducible Promoter



Genome Editing Glossary

Cas = CRISPR-associated genes

Cas9, Csn1 = a CRISPR-associated protein containing two nuclease domains, that is programmed by small RNAs to cleave DNA

crRNA = CRISPR RNA

dCAS9 = nuclease-deficient Cas9

DSB = Double-Stranded Break

gRNA = guide RNA

HDR = Homology-Directed Repair

HNH = an endonuclease domain named for characteristic histidine and asparagine residues

Indel = insertion and/or deletion

NHEJ = Non-Homologous End Joining

PAM = Protospacer-Adjacent Motif

RuvC = an endonuclease domain named for an *E. coli* protein involved in DNA repair

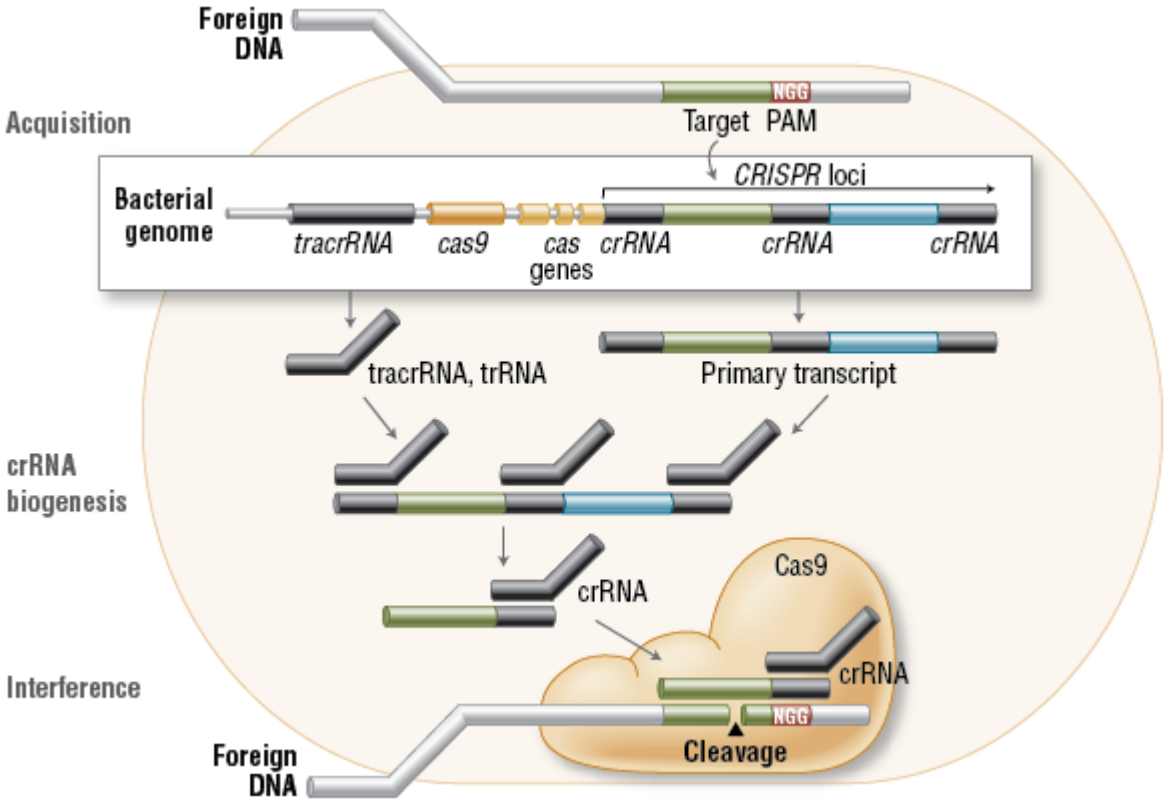
sgRNA = single guide RNA

tracrRNA, trRNA = trans-activating crRNA

TALEN = Transcription-Activator Like Effector Nuclease

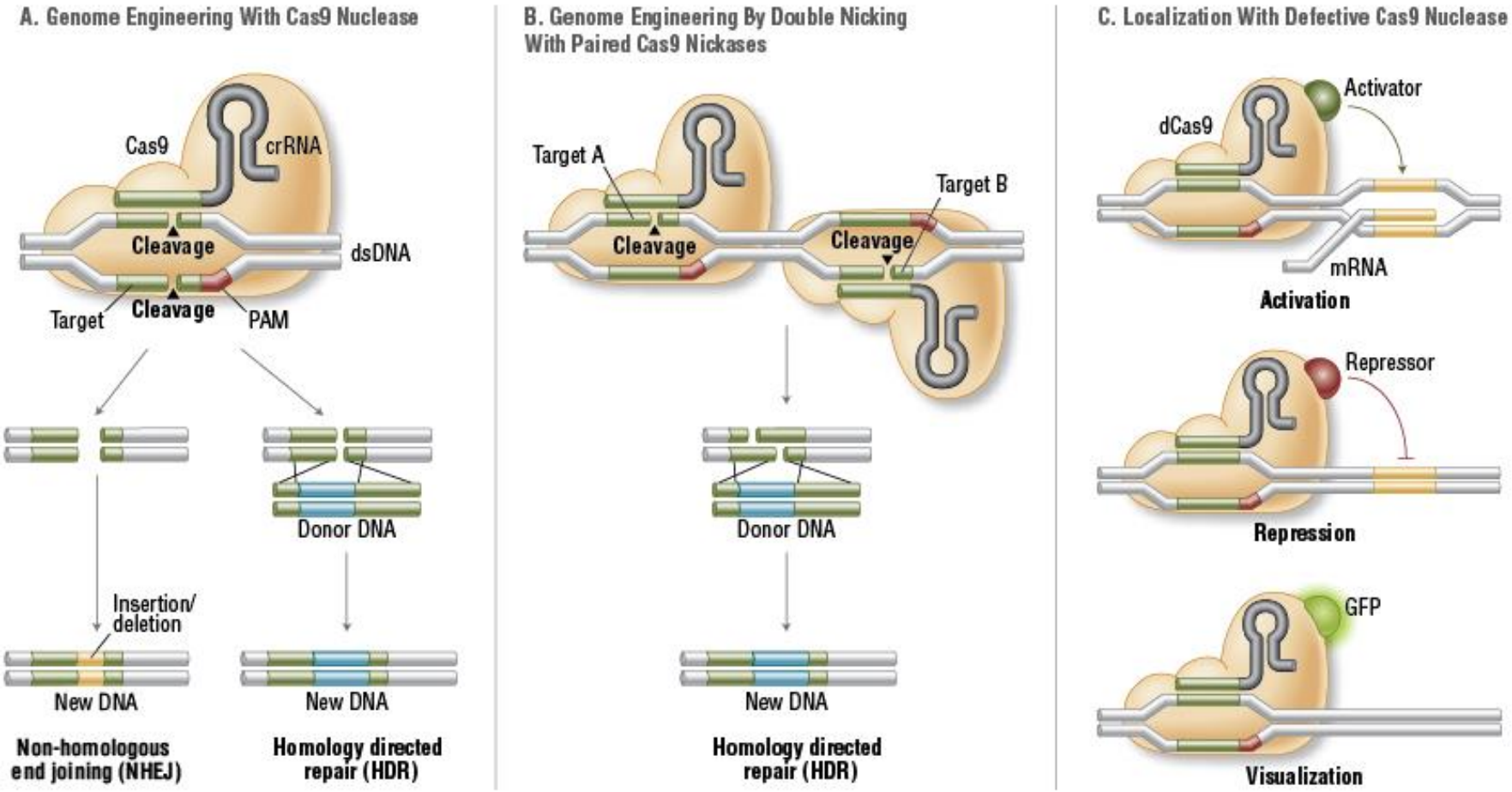
ZFN = Zinc-Finger Nuclease

https://www.neb.com/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology



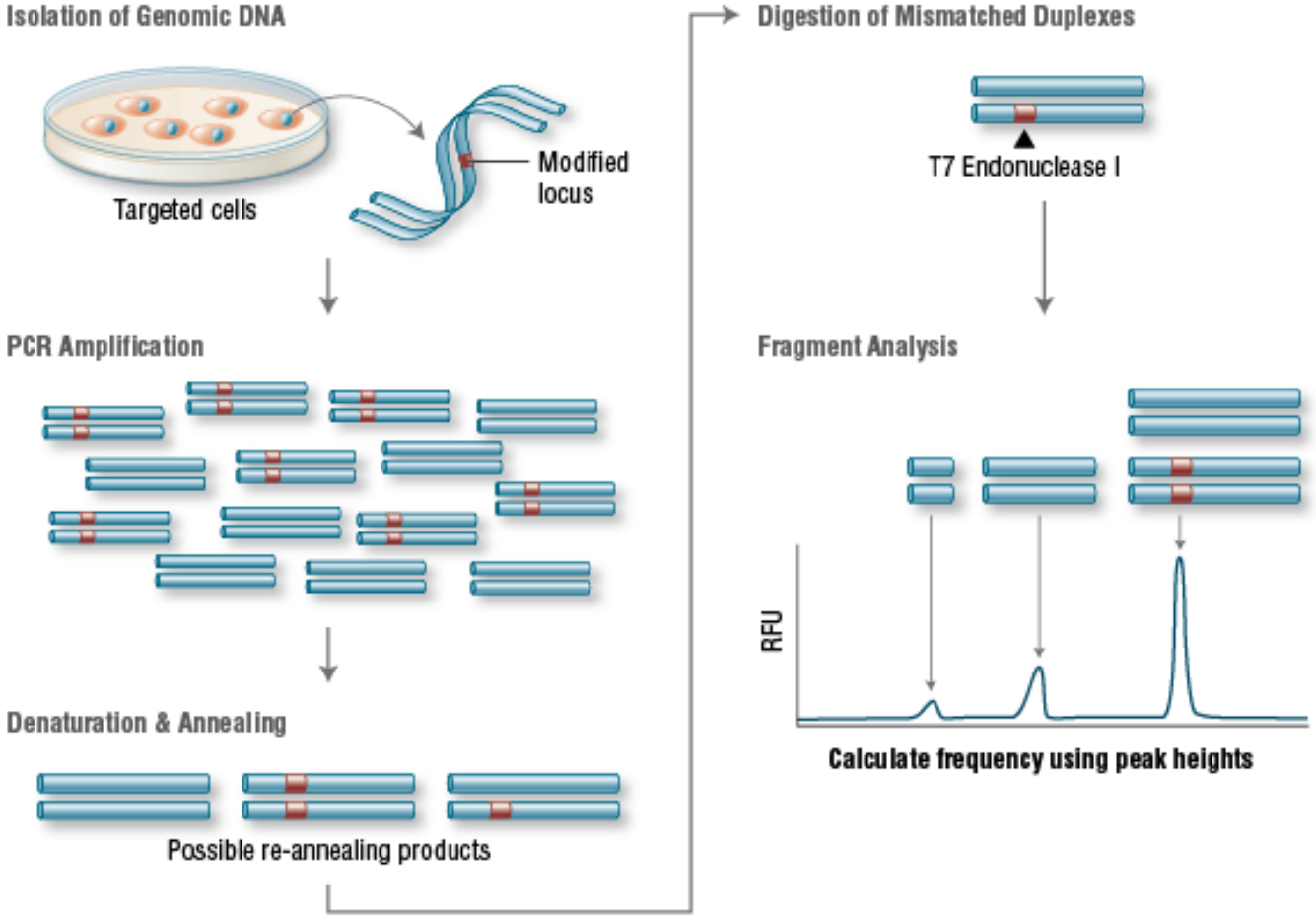
In the acquisition phase, foreign DNA is incorporated into the bacterial genome at the CRISPR loci. CRISPR loci is then transcribed and processed into crRNA during crRNA biogenesis. During interference, Cas9 endonuclease complexed with a crRNA and separate tracrRNA cleaves foreign DNA containing a 20-nucleotide crRNA complementary sequence adjacent to the PAM sequence. (Figure not drawn to scale.)

https://www.neb.com/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology



A. Wild-type Cas9 nuclease site specifically cleaves double-stranded DNA activating double-strand break repair machinery. In the absence of a homologous repair template non-homologous end joining can result in indels disrupting the target sequence. Alternatively, precise mutations and knock-ins can be made by providing a homologous repair template and exploiting the homology directed repair pathway.
B. Mutated Cas9 makes a site specific single-strand nick. Two sgRNA can be used to introduce a staggered double-stranded break which can then undergo homology directed repair.
C. Nuclease-deficient Cas9 can be fused with various effector domains allowing specific localization. For example, transcriptional activators, repressors, and fluorescent proteins.

https://www.neb.com/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology



Genomic DNA is amplified with primers bracketing the modified locus. PCR products are then denatured and re-annealed yielding 3 possible structures. Duplexes containing a mismatch are digested by T7 Endonuclease I. The DNA is then electrophoretically separated and fragment analysis is used to calculate targeting efficiency.

https://www.neb.com/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology