Journal of Science Research International (JSRI)
ISSN: 2456-6365 (Online)

DE NOVO DESIGN AND ENGINEERING OF SYNTHETIC MICROBIAL GENOMES: RECENT ADVANCES, APPLICATIONS AND FUTURE PERSPECTIVES

Meenakshi Johri*, Bindu Rajaguru, Gaurav Anamane, Riya Awasthi

Department of Biotechnology,

Pillai College of Arts, Commerce and Science (Empowered Autonomous), New Panvel, India

**Corresponding author E-mail: mjohri@mes.ac.in*

ORCID ID- https://orcid.org/0009-0002-0404-3473

Received: 01 June 2026	Revised: 19 July 2026	Accepted: 14 August 2026	Published: 31 August 2026

DOI: https://doi.org/10.5281/zenodo.22800000

Abstract:

Microbial genome engineering, from modification of individual genes toward genome minimisation, genome-wide recoding, genetic-code engineering, and de novo engineering of microbial genomes, is an emerging area of synthetic biology that aims to design, construct, and test genomes with defined genetic architectures and biological functions. Recent advances have expanded large-scale synthetic DNA assembly. This review summarises these developments with emphasis on computational genome design, identification of essential genetic functions, genome reduction, codon recoding, synthetic chromosome construction and iterative design-build-test-learn cycles. Current research shows that computational models can help prioritise candidate genome configurations, while DNA synthesis and genome engineering technologies make increasingly large genomic modifications experimentally achievable. Genome-reduced microorganisms are being investigated as simplified chassis for bio-manufacturing, recombinant protein production and metabolic engineering. At the same time, studies of reduced and recoded genomes reveal important limitations, including context-dependent gene essentiality, fitness costs, metabolic imbalance, genetic interactions and difficulties in predicting phenotypes from sequence information alone. These findings indicate that synthetic genome engineering cannot yet be treated as a purely computational design problem and requires repeated integration of modelling, construction, experimental testing and evolutionary optimisation. Future progress is likely to depend on combining whole-cell models, multi-omics data, automated genome construction, machine-learning approaches and application-specific chassis design. Together, these developments are moving synthetic biology toward increasingly rational and programmable microbial genomes.

Keywords: *Synthetic Biology, De Novo Genome Design, Genome Minimisation, Genome Recoding, Synthetic Chromosome, Genome Engineering.*

1. Introduction

Synthetic biology integrates molecular biology, genetic engineering, computational biology and biotechnology to design or redesign biological systems for defined purposes. Microorganisms are widely used as engineering platforms because their genomes can be modified and their metabolism can be redirected toward the production of useful proteins, metabolites, polymers and other bioproducts [1, 2].

Conventional microbial engineering usually targets individual genes, pathways or regulatory elements. Recent progress in DNA synthesis, genome editing, computational modelling and genome assembly has extended this approach toward genome-scale engineering [3, 4, 5]. A major milestone was the construction of an extensively recoded *Escherichia coli* genome containing thousands of designed synonymous substitutions, demonstrating that substantial genome rewriting can be compatible with cellular viability [3].

Genome minimisation is another important route toward synthetic microbial genomes. The aim is to remove genetic regions that are unnecessary for a defined condition while retaining functions required for growth and reproduction. Subsequent studies have shown, however, that gene essentiality can depend on environmental conditions and genetic background, making the concept of a single universal minimal genome difficult to establish [6, 7].

Large-scale genome engineering has also progressed through synthetic chromosome projects and improved DNA assembly strategies. These studies show that genome construction is increasingly becoming an iterative engineering process in which designed DNA is assembled, tested and revised rather than treated as a one-time construction task [8, 9].

Despite these advances, complete prediction of synthetic-genome behaviour remains difficult because cellular phenotypes arise from interactions among genes, regulatory elements, metabolic networks and environmental conditions. Genome reduction can cause fitness defects, while extensive recoding can alter cellular physiology in unexpected ways [10, 11, 7]. This review therefore examines recent advances in de novo synthetic microbial genome design and engineering, with emphasis on design principles, construction, minimisation, recoding, applications, challenges and future perspectives.

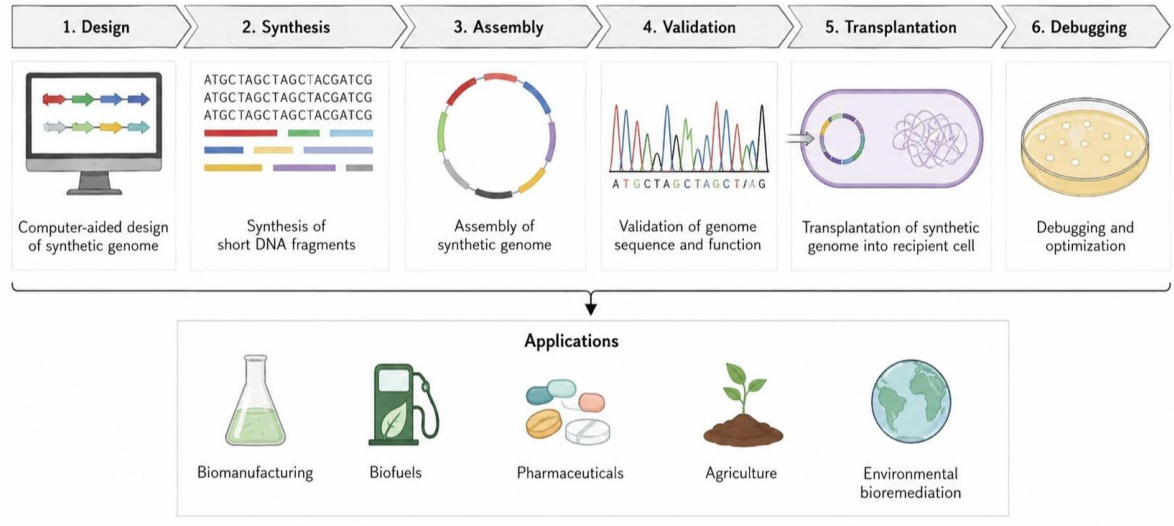

Figure 1: Overall workflow of De Novo synthetic microbial genome engineering

2. Principles of De Novo synthetic microbial genome design

2.1 Genome design and computational prediction

De novo genome design begins by defining the intended cellular functions and determining the genetic components needed to support them. Comparative genomics, essential gene datasets, metabolic models and whole-cell models can be used to evaluate possible genome configurations before experimental construction [4, 7]. Whole-cell modelling has been applied to generate candidate minimal genomes and to estimate the effects of alternative gene combinations, reducing the number of designs that need to be tested experimentally [4]. However, computational predictions remain dependent on the quality and scope of the underlying biological models.

2.2 Genome minimization

Genome minimisation removes dispensable genes or genomic regions to obtain a simpler and potentially more controllable chassis. Reduction can decrease competing functions and simplify subsequent engineering, but the boundary between essential and dispensable DNA is context-dependent [6]. Recent studies also show that reduced genomes may initially suffer growth or metabolic defects, which can sometimes be improved through adaptive evolution [10, 12].

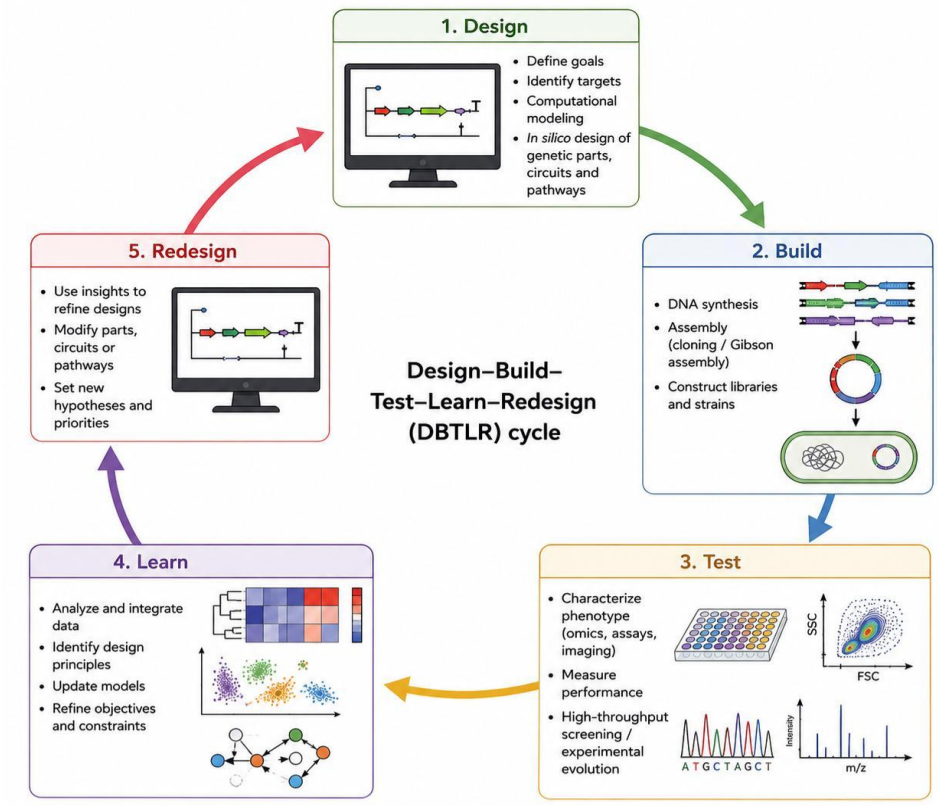

Figure 2: Design-Build-Test-Learn-Redesign (DBTLR) cycle

2.3 Codon compression and genome recoding

Genome recoding changes the nucleotide sequence of a genome according to a planned coding scheme. Fredens et al. demonstrated extensive synonymous recoding of *E. coli* and produced a viable four-megabase synthetic genome using a reduced codon set [3]. Later genetic-code engineering studies demonstrated that altering codon

interpretation can provide biological isolation and resistance to some forms of genetic exchange [11]. Recoding therefore expands genome engineering beyond deletion or insertion toward deliberate rewriting of information-processing rules.

2.4 Genome architecture and synthetic chromosomes

Genome design also involves the organization of genes, regulatory sequences, and chromosome regions. Standardized genome architectures have been proposed to make bacterial genome engineering more modular [5]. Synthetic chromosome studies in *Saccharomyces cerevisiae* further demonstrate that large genomic regions can be redesigned, assembled, and repeatedly tested to understand chromosome-level interactions [8,9]. Although yeast is eukaryotic, these studies provide useful engineering principles for large-scale genome construction.

2.5 design-build-test-learn cycles

A recurring principle in synthetic genomics is iteration: a genome is computationally designed, DNA is synthesized and assembled, the resulting organism is tested, and experimental data are used to improve the next design. This cycle is particularly important when genome-scale changes produce phenotypes that cannot be predicted accurately from individual gene functions [4,8,10].

3. Construction and synthesis of synthetic microbial genomes

3.1 DNA synthesis and genome assembly

Construction of a synthetic genome requires accurate synthesis of designed DNA fragments followed by their assembly into progressively larger genomic units. Large-scale synthesis projects have demonstrated that genome sections can be assembled and introduced into cells, extending genome engineering from local editing to chromosome-scale replacement [3,13]. The practical difficulty increases with genome size because sequence accuracy, assembly efficiency and verification must all be maintained.

3.2 Genome replacement and verification

Following DNA assembly, synthetic genomic regions must be incorporated into an appropriate cellular background and verified by sequencing and functional testing. The 2019 recoded *E. coli* study demonstrated that large numbers of designed sequence changes can be incorporated into a viable bacterial genome [3]. More recent genome-synthesis work has emphasized iterative correction of problematic regions, highlighting the importance of verification and troubleshooting in synthetic-genome construction [14].

3.3 genetic-code engineering

An advanced form of genome construction involves altering the genetic code itself. Swapping codon meanings can create genetic isolation from organisms using the standard code and can reduce the compatibility of engineered organisms with some biological information-transfer processes [3]. Such systems have potential value for biocontainment, although stability, escape and physiological effects must be evaluated experimentally.

4. Genome reduction and minimal microbial genomes

4.1 Identification of essential and dispensable regions

Genome reduction depends on identifying genetic regions that can be removed without preventing growth under the intended conditions. Essentiality can be assessed using comparative genomics, experimental deletion, transposon-based approaches and computational models [7,15]. The resulting definition is conditional rather than universal because genes that are dispensable in one environment may become important in another.

4.2 Reduced genomes as biotechnology chassis

Reduced genomes are being investigated as chassis for biotechnology because removing competing pathways may simplify cellular metabolism and redirect resources toward an intended product. Studies in *E. coli* have examined genome reduction together with polyhydroxybutyrate production, while work in *Paenibacillus polymyxa* and *Bacillus subtilis* has explored reduced strains for application-specific production [2,16,17]. These studies indicate that the useful level of genome reduction depends on the desired phenotype rather than simply on achieving the smallest possible genome.

4.3 Fitness costs and evolutionary recovery

Genome streamlining can reduce growth fitness because deleted functions may interact with metabolism and cellular regulation. Experimental evolution of genome-reduced *E. coli* has shown that substantial portions of the initial growth defect can be recovered during prolonged laboratory evolution [12]. Similarly, systemic analysis of the reduced *E. coli* DGF-298 strain identified metabolic imbalance associated with genome reduction and demonstrated that model-guided engineering could mitigate part of the defect [18].

4.4 Emerging genome-trimming strategies

Recent work is moving toward more systematic genome trimming. A 2026 CRISPR-based strategy called CREAT combines targeted cleavage and homology-arm walking to identify dispensable genomic regions and accelerate repeated genome reduction. The approach achieved substantial genome reduction in *Saccharolobus islandicus* and was also adapted to *Bacillus subtilis* and *E. coli* [19]. Such methods may reduce the trial-and-error burden associated with conventional genome minimization.

5. Genome recoding and reprogramming

Genome recoding differs from genome reduction because its primary objective is to rewrite genomic sequences rather than simply remove DNA. Synonymous codon replacement can alter genome architecture while preserving the encoded protein sequence, but recent research shows that extensive recoding can affect transcription, translation and cellular fitness [3,14]. Genetic-code engineering can additionally create biological barriers between engineered organisms and natural genetic systems [11]. These findings suggest that recoding should be treated as a systems-level intervention requiring extensive validation rather than as a collection of independent synonymous substitutions.

Table 1: Comparison of genome minimization and synthetic genome construction in microbial genome engineering

Sr. No.	Parameter	Genome Minimization	Synthetic genome construction
1.	Objective	To reduce the genome size by removing genes or DNA regions that are not essential for basic cellular functions.	To construct a genome with a planned genetic composition using designed and synthesized DNA sequences.
2.	Major Strategy	Identification and deletion of non-essential genes, intergenic regions and other dispensable DNA sequences.	Computational genome design followed by DNA synthesis, fragment assembly, verification and introduction into a suitable host.

3.	Effect on Genome	Produces a smaller and more streamlined genome while retaining genes required for cell survival and growth.	Produces a newly assembled or extensively redesigned genome with selected genetic features.
4.	Purpose	To simplify the genetic system and improve its usefulness as a microbial chassis for research and biotechnology.	To study genome organization and function and to create microorganisms with predetermined or modified characteristics.
5.	Application	Development of genome-reduced microbial chassis, metabolic engineering and production of useful biomolecules.	Synthetic biology, genome-function studies, microbial engineering and development of customized biological production systems.
6.	Limitation	It can be difficult to determine whether a gene is truly non-essential because its importance may change under different environmental conditions.	Large-scale DNA synthesis and assembly are technically challenging, expensive and require extensive genome verification.

6. Technologies Supporting Synthetic Microbial Genome Engineering

6.1 CRISPR-based engineering

CRISPR systems provide programmable genome modification and can support targeted deletions, insertions and multiplex editing. Their value in synthetic-genome projects is particularly high when many genomic regions must be modified in a controlled sequence.

6.2 Recombineering and multiplex editing

Recombineering enables designed DNA changes to be introduced through homologous recombination and can be combined with CRISPR-based selection. Multiplex strategies are useful when many edits must be introduced during genome reduction or recoding.

6.3 Computational and whole-cell modelling

Whole-cell models integrate multiple cellular processes and can prioritize candidate genome configurations before construction [4]. Their main limitation is that current models do not capture every genetic interaction, regulatory effect, or environmental influence.

6.4 multi-omics-guided genome engineering

Transcriptomics, proteomics, metabolomics, and related datasets can help identify why a synthetic genome performs poorly and can guide redesign. Recent recoding work illustrates how multi-omics measurements can be combined with genome synthesis and directed evolution to troubleshoot problematic synthetic regions [14].

7. Applications of synthetic microbial genomes

7.1 microbial cell factories

Genome-reduced microorganisms can be developed as simplified cell factories in which competing pathways are reduced, and resources are redirected toward a defined product. Recent studies report applications including polymer accumulation and recombinant protein production [2,13,16].

7.2 Recombinant protein and biopharmaceutical production

Genome-minimized Bacillus strains have been investigated for secretion of functional antibody fragments, demonstrating the potential of reduced-genome hosts for recombinant protein production [13]. Application-

specific chassis may be particularly useful when unwanted proteolysis, metabolic competition or genetic instability limits production in conventional hosts.

7.3 Biocontainment and Genetic Isolation

Genome recoding and genetic-code alteration can make engineered microorganisms less compatible with natural biological systems. Such approaches may reduce viral susceptibility or restrict horizontal transfer of genetic information [11]. However, containment must be evaluated under realistic evolutionary conditions.

7.4 Fundamental study of cellular organization

Synthetic genomes also provide experimental systems for studying essential genes, gene interactions, chromosome architecture, and the relationship between genotype and phenotype [4,8,20]. Thus, the field has both applied and fundamental scientific value.

8. Challenges

8.1 context-dependent essentiality

Genes that appear dispensable under one growth condition may become important under another, limiting the idea of a universal minimal genome [7].

8.2 Genetic interactions and unpredictable phenotypes

Large numbers of simultaneous changes can interact non-additively, producing phenotypes that cannot be predicted from individual edits [9,14].

8.3 Fitness and metabolic costs

Genome reduction can decrease growth and create metabolic imbalances, although evolutionary adaptation and model-guided engineering can partly restore performance [10,12,18].

8.4 DNA synthesis, assembly and verification

Large synthetic genomes require high sequence accuracy, efficient assembly and extensive verification, making construction technically demanding [3,13,14].

8.5 computational limitations

Whole-cell and genome-scale models can prioritize designs but cannot yet reproduce all molecular interactions and environmental effects of living cells [4,7].

8.6 Biosafety and containment

Recoded and genetically isolated organisms may provide useful containment strategies, but stability and the possibility of evolutionary escape require continued assessment [11].

9. Future perspectives

9.1 AI-Assisted genome design

Machine-learning approaches combined with large experimental datasets may improve genotype–phenotype prediction and help prioritize genome designs. Their greatest value is likely to come from integration with mechanistic and whole-cell models.

9.2 Multi-omics integration

Combining genome, transcriptome, proteome, and metabolome information should improve diagnosis of defects in synthetic genomes and guide targeted redesign [14,18].

9.3 Automated design-build-test-learn

Automation of DNA assembly, genome editing, phenotyping and data analysis could shorten the time between genome design and validated biological function [8,9].

9.4 Evolution-assisted engineering

Experimental evolution can be deliberately incorporated after genome construction to recover growth and stability. Recent genome-reduction studies provide direct evidence that adaptation can compensate for part of the fitness cost caused by streamlining [10,12].

9.5 Application-specific synthetic chassis

Future synthetic microorganisms are likely to be designed for particular products, environments, or biocontainment requirements rather than according to a single universal definition of minimality [13,16,18].

10. Conclusion

De novo synthetic microbial genome engineering has progressed from genome minimization and individual genetic modifications toward deliberate genome-scale design, synthesis, recoding and testing. Recent studies demonstrate that large portions of microbial genomes can be redesigned, but they also show that genome sequence alone is not sufficient to predict cellular behaviour. Context-dependent essentiality, genetic interactions, fitness costs and metabolic effects remain major barriers to fully predictable synthetic genomes. The most promising path forward is therefore an iterative integration of computational modelling, DNA synthesis, genome engineering, multi-omics, experimental testing and evolutionary optimization. With continued development of these approaches, application-specific synthetic microbial genomes may become increasingly useful for biotechnology, bio-manufacturing, biological containment and fundamental studies of cellular organization.

References

1. Vernyik, V., Karcagi, I., Tímár, E., (2020). Exploring the fitness benefits of genome reduction in *Escherichia coli* by a selection-driven approach. *Scientific Reports, 10*, 7345.
2. Ma, S., Su, T., Liu, J., (2022). Random genome reduction coupled with polyhydroxybutyrate biosynthesis to facilitate its accumulation in *Escherichia coli*. *Frontiers in Bioengineering and Biotechnology, 10*, 978211.
3. Fredens, J., Wang, K., de la Torre, D., (2019). Total synthesis of *Escherichia coli* with a recoded genome. *Nature, 569*, 514–518. https://doi.org/10.1038/s41586-019-1192-5
4. Rees-Garbutt, J., Chalkley, O., Landon, S., (2020). Designing minimal genomes using whole-cell models. *Nature Communications, 11*, 836. https://doi.org/10.1038/s41467-020-14545-0
5. Bayer, C. N., Rennig, M., Ehrmann, A. K., & Nørholm, M. H. H. (2021). A standardized genome architecture for bacterial synthetic biology. *Nature Communications, 12*, 5876.
6. Hutchison, C. A., III, Chuang, R.-Y., Noskov, V. N., (2016). Design and synthesis of a minimal bacterial genome. *Science, 351*(6280), aad6253. https://doi.org/10.1126/science.aad6253
7. Antczak, M., Michaelis, M., & Wass, M. N. (2019). Environmental conditions shape the nature of a minimal bacterial genome. *Nature Communications, 10*, 3100. https://doi.org/10.1038/s41467-019-10837-2
8. Foo, J. L., Kitano, S., Susanto, A. V., (2023). Establishing chromosomal design-build-test-learn through a synthetic chromosome and its combinatorial reconfiguration. *Cell Genomics, 3*(11), 100435. https://doi.org/10.1016/j.xgen.2023.100435

9. Zhao, Y., Coelho, C., Hughes, A. L., (2023). Debugging and consolidating multiple synthetic chromosomes reveals combinatorial genetic interactions. *Cell, 186*, 5220–5236.e16.
10. Moger-Reischer, R. Z., Glass, J. I., Wise, K. S., (2023). Evolution of a minimal cell. *Nature, 620*, 122–127. https://doi.org/10.1038/s41586-023-06288-x
11. Nyerges, Á., Vinke, S., Flynn, R., (2023). A swapped genetic code prevents viral infections and gene transfer. *Nature, 615*, 720–727. https://doi.org/10.1038/s41586-023-05824-z
12. Hitomi, K., Ishii, Y., & Ying, B.-W. (2024). Experimental evolution for the recovery of growth loss due to genome reduction. *eLife, 13*, RP93520. https://doi.org/10.7554/eLife.93520
13. Schilling, T., (2025). Toward antibody production in genome-minimized *Bacillus subtilis* strains. *ACS Synthetic Biology, 14*(3), 740–755. https://doi.org/10.1021/acssynbio.4c00688
14. Nyerges, Á., Chiappino-Pepe, A., Budnik, B., (2024). Synthetic genomes unveil the effects of synonymous recoding. *bioRxiv*. https://doi.org/10.1101/2024.06.16.599206
15. China, R., (2022). Reduction of the bacterial genome by transposon-mediated random deletion. *ACS Synthetic Biology, 11*(2), 668–677.
16. Ravagnan, G., Lesemann, J., Müller, M.-F., (2024). Genome reduction in *Paenibacillus polymyxa* DSM 365 for chassis development. *Frontiers in Bioengineering and Biotechnology, 12*, 1378873. https://doi.org/10.3389/fbioe.2024.1378873
17. Matteau, D., Champie, A., Grenier, F., & Rodrigue, S. (2023). Complete sequence of the genome-reduced *Escherichia coli* DGF-298. *Microbiology Resource Announcements, 12*. https://doi.org/10.1128/MRA.00665-23
18. Champie, A., Lachance, J.-C., Sastry, A., (2024). Diagnosis and mitigation of the systemic impact of genome reduction in *Escherichia coli* DGF-298. *mBio, 15*, e00873-24. https://doi.org/10.1128/mbio.00873-24
19. Yuan, G., Gao, Z., Qi, Y., (2026). CREAT: A CRISPR-based genome trimming strategy for systematic identification of dispensable regions and rapid genome reduction. *Advanced Science*. https://doi.org/10.1002/advs.76042
20. Sun, L., Wise, K. S., (2021). Genetic requirements for cell division in a genomically minimal cell. *Cell, 184*(9), 2430–2440.
21. Zürcher, D., (2023). Continuous synthesis of *E. coli* genome sections and Mb-scale human DNA assembly. *Nature, 619*, 555–562. https://doi.org/10.1038/s41586-023-06268-1