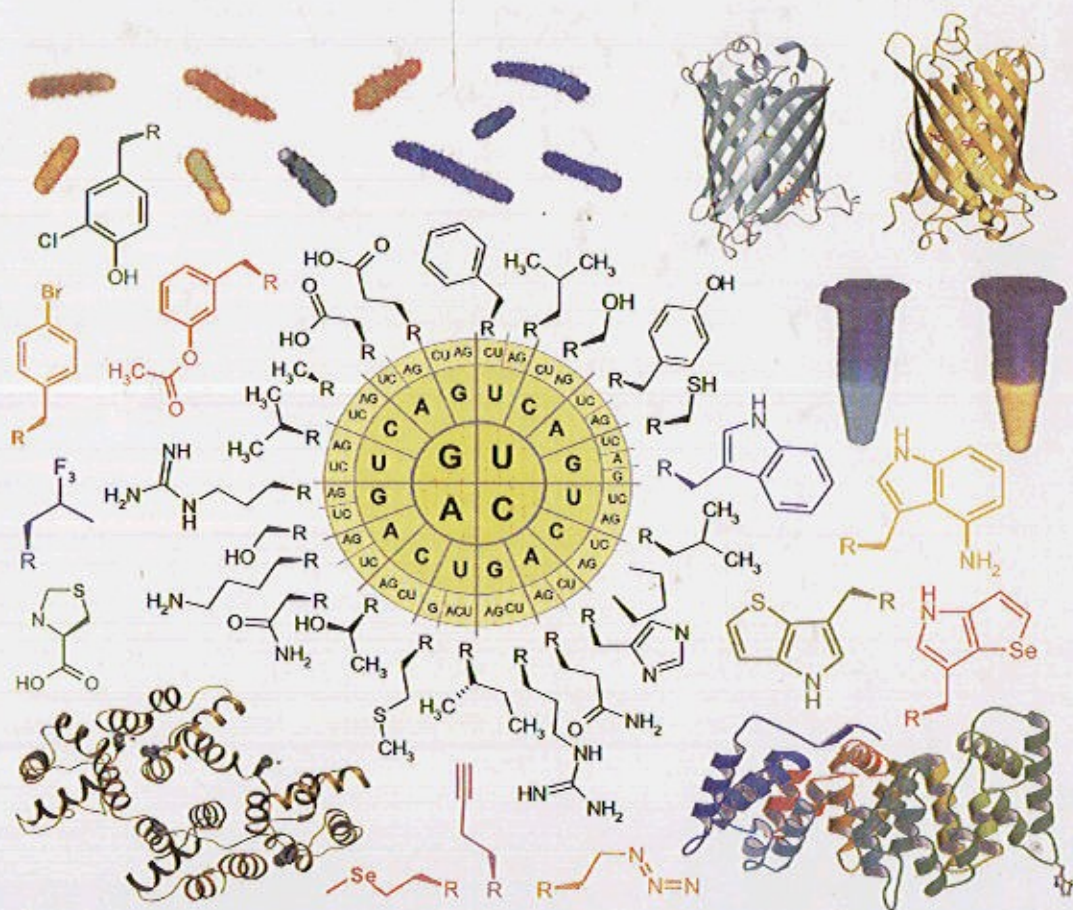


Engineering the Genetic Code

Expanding the Amino Acid Repertoire
for the Design of Novel Proteins



Contents

Foreword VII

Preface XV

1	Introduction	1
1.1	Classical Approaches to Protein Modification	1
1.2	Peptide Synthesis, Semisynthesis and Chemistry of Total Protein Synthesis	2
1.3	Chemoselective Ligations Combined with Biochemical Methods	5
1.4	Methods and Approaches of Classical Protein Engineering	5
1.5	Genetically Encoded Protein Modifications – Reprogramming Protein Translation	6
1.6	Basic Definitions and Taxonomy	8
	References	10
2	A Brief History of an Expanded Amino Acid Repertoire	13
2.1	The “Golden Years” of Molecular Biology and Triplet Code Elucidation	13
2.2	Early Experiments on the Incorporation of Amino Acid Analogs in Proteins	15
2.3	Test Tube (Cell-free) Synthesis of Proteins and Early Incorporation Experiments	18
2.4	Noncanonical Amino Acids as Tools for Studying Cell Metabolism, Physiology, Protein Processing and Turnover	19
2.5	Problem of Proofs and Formal Criteria for Noncanonical Amino Acid Incorporation	23
2.6	Recent Renaissance – Genetic Code Engineering	26
	References	28
3	Basic Features of the Cellular Translation Apparatus	31
3.1	Natural Laws, Genetic Information and the “Central Dogma” of Molecular Biology	31
3.2	Cellular Investments in Ribosome-mediated Protein Synthesis	33
3.3	Molecular Architecture of AARS	34

3.4	Structure and Function of the tRNA Molecule	40
3.5	Aminoacylation Reaction	44
3.6	AARS:tRNA Interactions – Identity Sets	45
3.7	Translational Proofreading	48
3.8	Ribosomal Decoding – A Brief Overview	50
3.9	Codon Bias and the Fidelity of Protein Synthesis	51
3.10	Preprogrammed Context-dependent Recoding: fMet, Sec, Pyl, etc.	53
3.11	Beyond Basic Coding – Posttranslational Modifications	56
	References	59
4	Amino Acids and Codons – Code Organization and Protein Structure	66
4.1	Basic Features and Adaptive Nature of the Universal Genetic Code	66
4.2	Metabolism and Intracellular Uptake of Canonical Amino Acids	68
4.3	Physicochemical Properties of Canonical Amino Acids	69
4.4	Reasons for the Occurrence of Only 20 Amino Acids in the Genetic Code	73
4.5	What Properties of Amino Acids are Best Preserved by the Genetic Code?	74
4.6	Evolutionary Legacy: Dual Nature of Conserved Code and Finite Number of Protein Folds	77
4.7	Natural Variations in Assignment of Codons of the Universal Genetic Code	79
4.7.1	Nucleoside Modifications and Codon Reassignments	81
4.8	Codon Reassignment Concepts Possibly Relevant to Code Engineering	81
4.8.1	Genome Size, Composition, Complexity and Codon Reassignments	81
4.8.2	Stop Codon Takeover, Codon Capture and Codon Ambiguity	83
	References	85
5	Reprogramming the Cellular Translation Machinery	90
5.1	Enzyme Specificity of Aminoacyl-tRNA Synthetases (AARS) and Code Interpretation	90
5.1.1	Living Cells as Platforms for Amino Acid Repertoire Expansion	90
5.1.2	Uptake, Toxicity and Metabolic Fate of Noncanonical Amino Acids	91
5.1.2.1	General Considerations	91
5.1.2.2	Amino Acid Transport	93
5.1.2.3	Metabolic Conversions and Toxicity of Analogs and Surrogates	94
5.1.3	Constraints and Levels in Code Engineering	95
5.1.4	Auxotrophism and Natural AARS with Manipulated Functions	95
5.1.4.1	Proteome-wide Replacements: “Unnatural Microorganisms”	97
5.1.4.2	Substitutions at the Level of Single Proteins – Selective Pressure Incorporation (SPI)	100
5.1.4.3	Kinetic Control – Enhanced System for Protein Translation	102
5.1.4.4	Extension of the Existing Specificities of AARS	104
5.1.4.5	AARS with an Attenuated Editing Function	106

5.1.5	Beyond Auxotrophism: Towards AARS with <i>De Novo</i> Substrate Specificity	108
5.1.5.1	Kinetic Issues of Orthogonal AARS: Catalytic Efficiency and Selectivity	111
5.2	Reassigning Coding and Noncoding Units	113
5.2.1	Sense Codon Reassignment: Most Commonly Used Substitutions	113
5.2.1.1	Tryptophan	114
5.2.1.2	Tyrosine	118
5.2.1.3	Phenylalanine	119
5.2.1.4	Histidine	122
5.2.1.5	Proline	124
5.2.1.6	Methionine	127
5.2.1.7	Leucine	130
5.2.1.8	Valine and Isoleucine	132
5.2.1.9	Arginine and Canavanine	133
5.2.1.10	Lysine and Lysyl-tRNA Synthetases – Enzymes with Many Talents	135
5.3	<i>In Vitro</i> Chemical and Enzymatic tRNA Aminoacylation	137
5.3.1	Chemical tRNA Acylation for Sense Codon Reassignment	138
5.3.2	Ribozymes, Ribosomes and MissenseSuppressions	140
5.4	Novel Codon–Anticodon Base Pairs	142
5.4.1	<i>In Vitro</i> and <i>In Vivo</i> Frameshift Suppression of 4- and 5-base Codons	142
5.4.2	Toward a Third, Noncanonical Base Pair in DNA	145
5.5	Stop Codon Takeover	146
5.5.1	The Concept of Suppression in Protein Translation	147
5.5.2	Chemical Aminoacylation of Amber Suppressor tRNA	148
5.5.3	<i>In Vitro</i> versus <i>In Vivo</i> Translation for Extension of the Amino Acid Repertoire	150
5.5.4	General Limits of Suppression-based Approaches	151
5.6	<i>In Vivo</i> Nonsense Suppression-based Methods	153
5.6.1	In Search for Orthogonal tRNA	153
5.6.2	Species-specific Aminoacylation Features and Orthogonal AARS:tRNA Pairs	154
5.6.3	Orthogonal AARS:tRNA Pairs in <i>E. coli</i>	156
5.6.4	Orthogonal Pairs in Yeast and Mammalian Cells	158
5.7	Outlook and Visions	159
5.7.1	Coupling Reprogrammed Translation with Metabolic Engineering	159
5.7.1.1	Catalytic Promiscuity and Synthetic Capacity Extension of Metabolic Pathways	159
5.7.1.2	Importing Natural Product Metabolic Pathways and Possible <i>De Novo</i> Design	160
5.7.2	Shuttle Orthogonal Pair and Hybrid Translation Systems with Codon Capture	162
	References	165

6	Implications and Insights: From Reprogrammed Translation and Code Evolution to Artificial Life	184
6.1	Code Engineering and Synthetic Biology	184
6.2	Novel Features of Protein Translation that have Emerged from Research in Code Engineering	185
6.2.1	Code Malleability, Catalytic Promiscuity and the Amino Acid “Identity” Problem	185
6.2.2	A Barrier between Allowable and Nonpermissive Amino Acids – An Index for Entry in the Genetic Code	187
6.2.3	Protein Structural Types and Amino Acid Substitution Capacity	190
6.2.4	Building a Direct Link between Metabolism and Reprogrammed Translation	192
6.3	The Amino Acid Repertoire and its Evolution	193
6.3.1	“Copernican Turn” and the Last Sacrosanct in Biochemistry	193
6.3.2	Spontaneous Terrestrial and Extraterrestrial Generation of Amino Acids	194
6.3.3	Metabolic Routes for Amino Acid Syntheses and Co-evolution Theory	197
6.4	Artificial Genetic Systems and Code Engineering	200
6.4.1	Cells with Chemically Ambiguous Proteomes – Codon Reassignment Issues	200
6.4.2	Is it Possible to Improve the Adaptive Features of the Genetic Code?	201
6.4.3	Possibilities for <i>De Novo</i> Design of Organisms with their “Own” Genetic Codes	203
6.4.4	Code Engineering and Society – Philosophical and Ethical Implications	205
6.4.5	Future Challenges, Chances and Risks	207
	References	208
7	Some Practical Potentials of Reprogrammed Cellular Translation	213
7.1	Practical Choice of Methods and Some Controversies in the Field	213
7.2	The Plasticity of the Translation Machinery, Amino Acid Generic Types and Protein Structure	215
7.3	DNA Nucleotide Analogs: From Sequencing to Expanded Code and Therapy	217
7.4	Noncanonical Amino Acids in Material Science	217
7.5	Isomorphous Replacement and Atomic Mutations in Structural Biology and Biophysics	221
7.5.1	Protein X-ray Crystallography	221
7.5.2	Atomic Mutations and Predictable Perturbations in the Protein Structure	223
7.5.3	Proteins Enriched with Chalcogen, Hydroxyl and Aza Analogs and Surrogates of Trp	225

7.6	Translationally Active Amino-Trp Analogs: Novel Spectral Windows and Protein Sensors	228
7.6.1	Providing Proteins with pH Sensitivity	228
7.6.2	Novel “Golden” Class of Autofluorescent Proteins	230
7.7	Fluorinated Amino Acids in Protein Engineering and Design	233
7.7.1	Monofluorinated Amino Acids in Protein Studies, Engineering and Design	233
7.7.2	Nonsticking Eggs and Bio-Teflon – Trifluorinated Amino Acids in Protein Engineering and Design	241
7.8	Protein Processing, Bioorthogonality and Protein Surface Diversifications	245
7.8.1	Chemoenzymatical Control of Protein Processing and Posttranslational Modifications	245
7.8.2	Staudinger–Bertozzi Ligation and “Click” Chemistry on Proteins	246
7.8.3	Tagging, Caging, Crosslinking and Photoswitching at the Protein Surface	247
7.9	Pharmacologically Active Amino Acids	249
7.9.1	Bioisosteric Compounds, Antagonists, Agonists and Antimetabolites	249
7.9.2	Neuroactive Amino Acids and their Derivatives	251
	References	253

Epilogue 261

Index 263