Preface

The title of this book could have been “The RNA World, 4th edition,” because it is very much the descendent of three previous volumes published by Cold Spring Harbor Laboratory (CSHL) Press. Instead, we’ve chosen “RNA Worlds” to reflect the book’s dual purpose. On the one hand, the volume covers the exciting diversity of form and function of RNA in the present day world, including RNA functions that have been discovered only recently and are still emerging. On the other hand, the volume maintains a major focus on the ancient RNA world that is thought to have predated genetically encoded proteins, DNA, and organisms we know about. As documented here, evidence for the reality of such a primordial role for RNA is increasing, although challenges remain in demonstrating model RNA-based replicases and in uncovering feasible origins of precursors for RNA synthesis. Nevertheless the plasticity of RNA stemming from its 2’ hydroxyl group and single-stranded nature, which permits diverse folding in contrast to its storage molecule cousin, DNA, continues to fascinate.

Even as we see health and economic benefits of biological research, including RNA research, pursuit of intellectual curiosity remains fundamentally important for the human psyche as it was for explorers in previous eras. Curiosity about our origins seems particularly deep-seated, as reflected by the complex tapestry of explanations in the religious heritages of diverse societies. But the scientific explanations of life’s origins seem more ennobling to us than, for instance, the idea of the original human female being derived from a rib of a man.

The evidence is overwhelming that not only all humans but all known life on earth had a common origin. This would not be a surprise to Darwin who wrote in his “Origin of Species” book one and a half centuries ago that “probably all organic beings which have ever lived on this earth have descended from some one primordial form.” To what extent these conclusions have been providing a bulwark against the intolerance of fundamentalist beliefs is, of course, part of a “well-worn” debate.

These considerations contribute to the broad interest in theories and experiments concerning a primordial RNA world. How best to focus interest on the key roles that RNA played in the origins of life and continues to play in present day species diversity and function? It is crucial to continue the present research directions in biology to reveal the bonanza yet in store and to encourage its distillation and dissemination to a wider audience.

Even the pioneers of the RNA world concept did not foresee that ribozymes had survived to the present day. Is it conceivable that some RNA-based organisms currently exist on our planet? After all, without ribosomes, such an “organism” may be even smaller than organisms that rely on protein synthesis. Such an organism might be uniquely present in some minute deep-rock niches, and maybe in environments that could not support predatory DNA-based organisms. Such a ribo-organism might even have evolved strategies to resist predation. While microorganisms are now known to live deeper in the earth and in more places than previously appreciated, the possibility of RNA-based organisms whose ancestor arose after the general advent of protein-based life does not seem to have been widely considered. Search of stable shales with appropriate pore sizes would likely yield new organisms with an unquantifiable possibility of the discovery of RNA-based organisms that would even require modification of the statement above about common ancestry. As referenced in the chapter by Benner and colleagues, this is just one of the possibilities for radically different life on earth whose potential existence merits investigation. Such a project would be much cheaper than the quest to find life on Mars and, if successful, just as wonderful scientifically.

Since the third edition of “The RNA World” book, a key pioneer of RNA world studies and author in three previous editions of the book, Leslie Orgel, has passed away. We will not forget his rigorous approach to origin-of-life experiments and his warm humor.

We thank Richard Sever and John Inglis of CSHL Press for wise advice and continuing support. It is a pleasure also to acknowledge the care and understanding of the project coordinator at CSHL, Inez Sialiano, in making this book a reality, and Susan Roberts for help with the cover graphics.

J.F. Atkins
R.F. Gesteland
T.R. Cech
Index

A

Adenine, abiotic synthesis, 33
S-Adenosylhomocysteine (SAH), riboswitch ligand, 69–71, 85
S-Adenosylmethionine (SAM) riboswitch ligand, 69–71
tandem riboswitches, 72–74
Adenovirus, noncoding RNA function, 166, 169
AFM. See Atomic force microscopy
Age-related macular degeneration (ARMD), aptamer therapy, 334–335
Ago. See Argonaute
Alanyl nucleic acid (ANA) abiotic synthesis, 37
polymer structure, 36
Alu RNA, RNA polymerase II loading mediation, 285–287
6-Amino-5-nitropyridin-2-one, six letter polymerase chain reaction, 15–16
ANA. See Alanyl nucleic acid
Ancient RNA
conservation and persistence, 45
evidence, 44–45
persistence, 45
Anisomycin, peptidyl transferase reaction inhibition, 131–132
Aptamer. See SELEX; SOMAmers
Arabinoside, abiotic synthesis, 33–34
ArcZ, small RNA regulator function, 219–220
Argonaute (Ago)
ancestral function, 250–251
phylogenetic analysis in three clades, 244–245
Piwi, 250, 258
RNA binding modes, 247–248
RNA interference role, 244, 257–258
small RNA-guided cleavage, 248–250
structure and function, 245–247
ARMD. See Age-related macular degeneration
Assemble program, interactive molecular modeling, 314–315, 317
Atomic composition, CHNOPS consistency, 45
Atomic force microscopy (AFM), RNA folding studies, 323–324, 328
Azithromycin, peptidyl transferase reaction inhibition, 131–132

B

Bacterial small RNA regulators (sRNAs). See also CRISPRs
abundance of types, 216
antisense small RNA regulators, 217
base pairing with limited complementarity functions

C

Cas. See CRISPRs
CCND1 histone modification mediation, 284–285
signal sensing, 288
Cellular life
membranes as compartment boundaries, 52–53
phosphoramide nucleic acids, 57–59
protocell assembly, 59
encapsulated templated replication, 59–60
prospects for complete model, 60–61
RNA-catalyzed RNA replication, 56–57
vesicle division pathways, 55–56
vesicle growth pathways, 53–55
outer membrane protein synthesis repression, 219
overview, 219–220
transcription factor synthesis modulation, 219
Hfq function, 218
Hfq-independent base pairing, 218–219
mechanisms, 217–219
outcomes, 217–218
evolution
capture of random transcription for purpose of regulation, 226
recent evolution
conserved gene neighborhood, 225
conserved regulation, 225
gene duplication, 225
horizontal transfer, 225–226
target-imposed constraints, 225
sRNA-mRNA connections, 226
sRNA-tRNA connections, 226
functional overview, 216
mimics in regulation, 221–222
mRNA dual function, 221
prospects for study, 227
protein activity modulation
CarB, 223
discovery, 223
6S RNA, 222–223
riboswitches, 221
tmRNA intrinsic functions, 223–224
Bayes’s theorem, RNA world applicability, 45
Borate, premetabolic cycle, 14–15
Brtr small nuclear ribonucleoprotein structure, 193–194
spliceosome function, 191

355
Index

chbBC, small RNA regulator function, 221–222
Chirality, early life, 35
Chloramphenicol, peptidyl transferase reaction inhibition, 131
Chromatin. See Heterochromatin
Cid12, heterochromatic RNA interference at fission yeast centromeres, 259
Class I ligase ribozyme. See RNA ligase ribozyme
Clusters of regularly interspersed short palindromic repeats. See CRISPRs
CMT3, heterochromatic RNA interference, 262
CMV. See Cytomegalovirus
Contemporary RNA world, overview, 3–4
COSMIC LOPER, overview, 48
CRISPRs
  Cas protein functions, 235
  examples, 233
  functional overview, 232
  loci and Cas genes, 232, 234
  mechanism of action
    integration of new spacers, 235–236
    target interference, 237–238
    transcription and processing, 236–237
    prospects for study, 240
  RNA interference homology, 238–240
CsrB, small RNA regulator function, 223
Ctcf, X-chromosome inactivation role, 271
CyaR, small RNA regulator function, 226
Cytomegalovirus (CMV), noncoding RNA, 174
  discovery, 174–175
  functions, 168, 170
  microRNAs, 173–174
  ribonucleoprotein complexes, 168, 170
  transcripts, 166, 168
Erl1, silencing on chromosome arms, 261
Error threshold, RNA replicate, 26–27
Erythromycin, peptidyl transferase reaction inhibition, 131
Evf2, histone modification mediation, 283–284
F
FAD
  continuity with RNA world, 47–48
  initial darwinian ancestor modern descendant, 46
fbp1+, RNA polymerase II loading mediation, 287
Fidelity
  early RNA replicate, 26–27
  translation, 133
Flavin mononucleotide (FMN), riboswitch ligand, 67–68, 85
Fluorescence resonance energy transfer (FRET)
  folding pathways of purine riboswitch aptamer, 83–84
  single-molecule studies of RNA reactions, 324–329
  vesicle growth studies, 53–54
FMN. See Flavin mononucleotide
Folding. See Modules, RNA; Secondary structure, RNA; Tertiary structure, RNA
Frameshifting. See Ribosome
G
GlcN6P. See Glucosamine-6-phosphate
 glmS ribozyme
  catalytic mechanism, 98–99
  glucosamine-6-phosphate cofactor, 98–99
  structural overview, 95–96
GlmE, small RNA regulator function, 221–222
Glucosamine-6-phosphate (GlcN6P)
  glmS ribozyme cofactor, 98–99
  riboswitch ligand, 66–67
Glyceraldehyde-2-phosphate, abiotic synthesis, 33
Glycoaldehyde phosphate, abiotic synthesis, 33
Glycol nucleic acid (GNA)
  abiotic synthesis, 37
  polymer structure, 36
GNA. See Glycol nucleic acid
Group II intron
  associated proteins
    DEAD-box proteins, 113
    intron-encoded protein component, 104, 106–108, 111–112
    LtrA, 112
    recruitment, 113
  catalytic mechanism, 108–109
  degenerate intron splicing, 106
  evolution, 116–117
  lineages, 108
  mobility
    DNA target site recognition, 115
    retrohoming
      mechanisms, 115
    reverse splicing into DNA, 113–115
    retrotransposition into new sites, 115–116
  targetrons,
phylogenetic distribution, 104
RNA structure, 104–106
three-dimensional structure, 109–111
trans-splicing, 106
twintron formation, 106

H
Hairpin ribozyme
catalytic mechanism, 96–97
fluorescence resonance energy transfer studies, 324–325
RNase H homology, 97
structural overview, 95–96
Hammerhead ribozyme
catalytic mechanism, 98
structural overview, 95–96, 98
HDV ribozyme. See Hepatitis delta virus ribozyme
Hepatitis delta virus (HDV) ribozyme, structural overview, 95–96, 99–100
Herpes simplex virus-1 (HSV-1), noncoding RNA, 174
Herpesvirus saimiri U RNA (HSUR)
function, 171
structure, 170–171
transcripts, 170–171
Heterochromatin
definition, 256
histone modification mediation by long noncoding RNA
CCND1, 284–285
Esf2, 283–284
HOTAIR, 285
SRA, 283
histone modifications, 256–257
position effect variegation, 256
RNA interference
filamentous fungi, 262–263
fission yeast centromeres, 258–261
silencing on chromosome arms, 261–262
Hfq, small RNA regulator base pairing with limited complementarity role, 218
Histones. See Heterochromatin
HIV. See Human immunodeficiency virus
HOTAIR, histone modification mediation, 285
HOX genes, HOTAIR regulation, 285
Hrtr1, heterochromatic RNA interference at fission yeast centromeres, 259
HSUR. See Herpesvirus saimiri U RNA
HSV-1. See Herpes simplex virus-1
Human immunodeficiency virus (HIV), reverse transcriptase RNA-binding studies, 326–328
Hydrolysis, susceptibility of RNA bonds, 12

I
IAPV. See Israeli acute paralysis virus
IDA. See Initial darwinian ancestor
Imidazo[1,2-c]pyrimidin-5(1H)-one, six letter polymerase chain reaction, 15–16
Initial darwinian ancestor (IDA)
AMP-containing enzymatic cofactors as modern descendants, 45–46
origin of life context, 47
time vantage, 44
Initiation factors, deletion studies, 142
Internal transesterification reaction, small self-cleaving ribozymes, 94–95
Internally transcribed spacer (ITS), phylogenetic reconstruction, 303
Intron-encoded protein. See Group II intron
Israeli acute paralysis virus (IAPV), RNA interference targeting, 350
ITS. See Internally transcribed spacer

K
Kaposi's sarcoma-associated herpesvirus (KSHV)
microRNA, 173
PAN RNA function, 171–172
Kasugamycin, resistance mechanisms, 142
Khps1, DNA methylation mediation, 285
KSHV. See Kaposi's sarcoma-associated herpesvirus

L
L1 ligase. See RNA ligase ribozyme
L22, Epstein-Barr virus-encoded RNA binding, 170
La, Epstein-Barr virus-encoded RNA binding, 168
Last universal common ancestor (LUCA), RNA fragments, 11
Leakage, premetabolic cycle, 15
Life. See also Cellular life
definition, 8
RNA-first view on origin
nonenzymatic replication of RNA, 24–25
nucleotide biosynthesis, 31–32
polynucleotide abiotic synthesis, 22–24
RNA replicate, 25–31
RNA-later view on origin
alternate genetic systems, 35–38
nucleotide synthesis, 32–35
Linezolid, peptidyl transferase reaction inhibition, 131
Long noncoding RNA. See also Polycomb proteins;
X-chromosome inactivation
advantages in function
sequence specificity, 275–276
transcription site localization, 275
cell structure integrity role, 288
classification, 280
expression regulation, 281
functional overview, 280
processing, 281–282
promoter-associated RNA functions, 287–288
prospects for study, 288
protein complex recruitment, 282–283
signal sensing, 288
strandedness, 280
Long noncoding RNA (Continued)
subcellular localization, 280–281
transcriptional boundary marking, 288
transcriptional regulation mechanisms
DNA methylation, 285
histone modification, 283–285
RNA polymerase II loading, 285–287
transcription factors, 287
transcription interference, 287
LtrA, group II intron association, 112
LUCA. See Last universal common ancestor

M
M1CB, herpesvirus targeting, 173
Macugen, historical perspective, 334
Malaria, RNA interference targeting, 350
Manip program, interactive molecular modeling, 314
Meiotic silencing of unpaired DNA (MSUD), RNA interference pathway, 251
Membrane
compartment boundaries, 52–53
vesicles
division pathways, 55–56
growth pathways, 53–55
1-Methyl-adenine derivatives, polymerization, 24
2-Methylimidazolide derivatives, polymerization, 24
MHV-68. See Murine herpesvirus-68
MicA, small RNA regulator function, 219–220
MicroRNA, viral function, 172–174
Modules, RNA
crystal structure comparisons, 315–316
examples, 313
large assembly modeling, 313
searching, 312–313
Montmorillonite, nucleoside 5’-phosphorimidazole polymerization, 24
mRNA
splicing. See Spliceosome
translation. See Ribosome
MSUD. See Meiotic silencing of unpaired DNA
Murine herpesvirus-68 (MHV-68), noncoding RNA, 175

N
NAD/P
abiatic synthesis of NAD, 46
continuity with RNA world, 47–48
initial darwiniian ancestor modern descendant, 45–47
replication chemistry, 48–49
NEAT RNAs, cell structure integrity role, 288
NMR. See Nuclear magnetic resonance
NPRON, transcription factor mediation, 287
NS3, helicase mechanism studies, 328
Nuclear magnetic resonance (NMR), ligand-induced conformational change studies in riboswitches, 86
Nucleolin, Epstein-Barr virus-encoded RNA binding, 170
Nucleoside 5’-phosphorimidazoles, polymerization, 23–24
Nucleotides
abiatic synthesis, 32–35
activation in polymerization, 20–21, 25, 32
synthetic ribozymes, 31–32

O
OmrB, small RNA regulator function, 219–220
OmtA, small RNA regulator function, 219–220
Optical tweezers
force measurements, 322–323
ribosome-mRNA movement studies, 328–329
OxyS, small RNA regulator function, 219–220

P
P15AS, DNA methylation mediation, 285
PABPC1, See Poly(A)-binding protein C1
PAN RNA. See Polyadenylated nuclear RNA
PCR. See Polymerase chain reaction
Peptide nucleic acid (PNA)
abiotic synthesis, 37
polymer structure, 36
Peptidyl transferase center. See Ribosome
Perchlorate, solar system distribution, 15
PEV. See Position effect variegation
Phosphoramidate nucleic acids, cellular life origins, 57–59
5-Phosphoribosyl-1-pyrophosphate (PRPP), ribozyme synthesis limitations, 32
piRNA, functions, 250–251, 258
Piwi. See Argonaute
PKR. See Protein kinase R
PNA. See Peptide nucleic acid
Poly(A)-binding protein C1 (PABPC1), PAN RNA association, 172
Polyadenylated nuclear (PAN) RNA, Kaposi’s sarcoma-associated herpesvirus function, 172
Polycomb proteins
long noncoding RNA advantages in regulation, 274–276
RNA recruitment, 272–274, 282
Polymerase chain reaction (PCR), six letter polymerase chain reaction, 15–16
Position effect variegation (PEV), heterochromatin silencing, 256
Pre-mRNA splicing. See Spliceosome
Primordial RNA world, overview, 2–3
p-RNA, structure, 35–36
Protein kinase R (PKR), Epstein-Barr virus-encoded RNA binding, 168
Protein synthesis. See also Ribosome
duplicator RNA and origins of ribosomal decoding site, 148–150
evolutionary driving force for translation from RNA world, 150–151, 162–163
RNA role, 8, 10, 22, 126–127, 142–144, 157–158
stop tRNAs and type I release factor evolution, 148
translation steps, 156–157
Protocell. See Cellular life
Prp2, spliceosome function, 191
Prp5
heterochromatic RNA interference at fission yeast centromeres, 260
spliceosome function, 190–191
Prp8
heterochromatic RNA interference at fission yeast centromeres, 260
spliceosome function, 191, 195
Prp10, heterochromatic RNA interference at fission yeast centromeres, 260
Prp16, spliceosome function, 191
Prp19, spliceosome function, 188, 191
Prp22, spliceosome function, 191–192
Prp28, spliceosome function, 191
Prp43, spliceosome function, 192
PRPP. See 5-Phosphoribosyl-1-pyrophosphate
Pseudoknot
frameshifting studies, 325–326
telomerase RNA, 208–209
R2 retrotransposon, RNA structure elucidation, 297–298
RCNMV. See Red clover necrotic mosaic virus
Rdp1, heterochromatic RNA interference at fission yeast centromeres, 259
Red clover necrotic mosaic virus (RCNMV), noncoding RNA, 175
Release factors, stop tRNAs and type I release factor evolution, 148
Release factors, tRNA mimics, 129–130
RepA
Polycomb protein recruitment, 274–275
signal sensing, 288
X-chromosome inactivation role, 269–270
Ribosome. See also Protein synthesis
aminoacyl-tRNA selection, 144–146
antibiotic inhibition of peptidyl transferase reaction, 130–133
complexes with both A- and P-site substrates bound, 127–128
crystal structures, 125–126, 156
duplicator RNA and origins of ribosomal decoding site, 148–150
elongation factor structure and function, 136–137
evolutionary drive for translation from RNAs, 150–152, 162–163
fidelity of translation, 133
induced-fit activation of peptide synthesis, 128–129
mRNA movement studies with optical tweezers, 328–329
peptidyl-tRNA protection from hydrolysis, 128–129
protein structures and functions, 148
pseudoknot frameshifting studies, 325–326
release factors as tRNA mimics, 129–130
tRNA
role in mRNA decoding, 134–136
tRNA interactions in P site, 146–147
subunits and functions, 124–125
transition state intermediate and stabilization, 128
translocation
E site conservation and role, 162
RNA molecular mechanics, 147
subunit interactions, 161–162
tRNA energy storage, 162
tRNA
binding sites in small ribosomal subunit, 133–134
CCA end binding to E site of large subunit, 130
decoding
evolutionary role of minor groove recognition, 158–159
minor groove recognition by RNA, 158
structural changes, 159
dynamics analysis with fluorescence resonance energy transfer, 329
stop tRNAs and type I release factor evolution, 148
Riboswitch
S-adenosylmethionine riboswitch, 69–71
classes, 64–65, 68–69
domains, 80
effector molecule recognition, 84–85
functional overview, 64, 80
gene regulation
mechanisms, 65–66, 80
temporal sensitivity, 86–87
ligand binding affinity and kinetics of function, 67–68
terms, 75
ribozyme control mechanisms, 67
RNA world functions, 89
small RNA regulators, 221
structure
elements, 80–82
folding pathways of purine riboswitch aptamer, 82–84
ligand-induced conformational change, 83–86
models for structural switching, 87–89
tandem switches, 71–75
RIG-I, helicase mechanism studies, 328
RITS. See RNA-induced initiation of transcriptional silencing complex
RNA III, small RNA regulator function, 221
RNAi. See RNA interference
RNA-induced initiation of transcriptional silencing complex (RITS), heterochromatic RNA interference at fission yeast centromeres, 259–260
RNA interference (RNAi). See also Argonaute
CRISPR homology, 238–240
eukaryote distribution, 257
heterochromatic RNA interference
filamentous fungi, 262–263
fission yeast centromeres, 258–261
silencing on chromosome arms, 261–262

359
Index

RNA interference (RNAi) (Continued)
immune function, 250
mechanisms, 257–258
pest control applications, 350
screening in vivo
Caenorhabditis elegans, 345–346
developmental biology applications, 349
Drosophila melanogaster, 346–349
plants and crop improvement, 350
reagents, 344–345
vertebrate models, 349–350
therapeutic prospects, 350–351
RNA ligase ribozyme
magnesium dependence, 30
polymerization reaction and optimization, 30
structures, 29–30
RNA polymerase II
heterochromatic RNA interference at fission yeast centromeres, 260
long noncoding RNA mediation of loading, 285–287
RNA–protein complexes, modeling, 313
RNA replicase ribozyme
accuracy and survival, 26–27
cellular life origins, 56–57
chicken-and-egg paradox, 27–29
ligase evolution, 29–31
RNA-first view of life, 25
RNA technology, as third RNA world, 4
RprA, small RNA regulator function, 219–220
rRNA. See Ribosome
RyB, small RNA regulator function, 219–220
RyB, small RNA regulator function, 219–220

S
SAH. See S-Adenosylhomocysteine
SAM. See S-Adenosylmethionine
Secondary structure, RNA
comparative sequence analysis
automation
align then fold approach, 300
fold and align approach, 299
fold then align approach, 300
programs, 300
overview, 311–312
crystal structure comparisons, 315–316
evolution studies, 302–304
folding space restraints
combined methods for determination, 297–298
comparative structures, 297
experimental findings, 296–297
overview, 295–296
free energy minimization, 294
functional RNA discovery, 300–301
overview of analysis, 310–311
partition functions and probabilities, 294–295
phylogenetic reconstruction, 303
prediction programs, 295
prospects for study, 301–302
SELEX
clinical applications of aptamers, 334–335
historical perspective, 334–335
proteomics for aptamer optimization, 335–336
RNA function discovery, 338–339
RNA world applications of aptamers, 337–338
SOMAmers
plasma proteome probing at high throughput and specificity, 337
specificity, 336
structure, 336
SF1, spliceosome function, 190
SF3 proteins, spliceosome function, 188, 190
SgrS, small RNA regulator function, 221
Six letter polymerase chain reaction, 15–17
Small nuclear ribonucleoproteins (snRNPs), spliceosome structures
electron microscopy, 193–194, 197–199
high-resolution structures, 194–195
U1 small nuclear ribonucleoprotein, 195–196
Small RNA regulators. See Bacterial small RNA regulators; CRISPRs
Small self-cleaving ribozymes
active site mechanisms, 96
glmS ribozyme, 98–99
hairpin ribozyme, 96–98
hammerhead ribozyme, 98
hepatitis delta virus ribozyme, 99–100
internal transesterification reaction, 94–95
structural overview, 95–96
snRNPs. See Small nuclear ribonucleoproteins
SOMAmers
plasma proteome probing at high throughput and specificity, 337
specificity, 336
structure, 336
Spliceosome
cis-acting elements and catalytic steps of splicing, 182–184
conformational two-state model for catalytic center, 187
intron- and exon-defined assembly pathways, 183, 185
overview of pre-mRNA splicing, 182
protein
composition dynamics, 187–189
flexibility of metazoan spliceosomes, 189–190
posttranslational modifications, 192–193
RNA footprinting studies, 197
splice site recognition role, 190
structural rearrangement facilitation, 190–192
RNA–RNA interactions, 185–187
small nuclear ribonucleoprotein structures
electron microscopy, 193–194, 197–199
high-resolution structures, 194–195
U1 small nuclear ribonucleoprotein, 195–196
U2-dependent spliceosome mechanism, 183–184
SRA, histone modification mediation, 283
SreA, small RNA regulator function, 221
SreB, small RNA regulator function, 221
sRNAs. See Bacterial small RNA regulators
Sub2, spliceosome function, 190–191
SUNH4/KYP, heterochromatic RNA interference, 262
Sutherland’s synthesis, prebiotic origins of RNA, 13
Systematic evolution of ligands by exponential enrichment.
See SELEX; SOMAmers

T
Targetron, group II intron, 116
Telomerase
  evolutionary origins of components, 211–212
  functional overview, 206
  protein and RNA component interplay, 206, 210
  RNA
    evolutionary divergence, 206–207
    fine-tuning of repeat synthesis activity, 209–210
    holoenzyme biogenesis and regulation motifs, 210–211
    motif gain-of-function in ribonucleoprotein context, 212
    nuclear assembly and addressing in vertebrates, 211
    pseudoknot, 208–209
    stem terminus element, 209
    template 5′-boundary enforcement, 210
    template, 207–208
    TERT-binding motifs, 210
    yeast function, 211
Template-directed replication
  multi-stem-loops structures, 28
  nonenzymatic replication of RNA, 24–25
TER. See Telomerase
Tertiary structure, RNA
  constraint extraction, 312
  crystal structure comparisons, 315–316
  large assembly modeling, 313
Thiamin pyrophosphate (TPP), riboswitch ligand, 66–67,
85, 86
Thioester peptide nucleic acid (tPNA)
  abiotic synthesis, 37
  polymer structure, 36–37
4-Thiouridylate, ribozyme synthesis, 31–32
Threose
  instability, 12
  ribose substitution, 15
  threose nucleotide polymers, 35–36
tmRNA, functions, 223–224
tPNA. See Thioester peptide nucleic acid
TPP. See Thiamin pyrophosphate
Translation. See Protein synthesis; Ribosome
tRNA. See Ribosome
Tsix
  Polycomb protein binding, 274
  signal sensing, 288
  X-chromosome inactivation role, 270–272

U
Unnatural nucleic acids, potential in RNA world,
15–16

V
Vascular endothelial growth factor (VEGF), Macugen
inhibition, 334
VEGF. See Vascular endothelial growth factor
Vesicle. See Membrane
Viral noncoding RNA
  adenovirus, 166, 169
  cytomegalovirus, 174
  Epstein-Barr virus, 166, 168, 170, 174–175
  herpes simplex virus-1, 174
  herpesvirus saimiri, 170–171
  Kaposi's sarcoma-associated herpesvirus, 171–172
  microRNAs, 172–174
  murine herpesvirus-68, 175
  overview, 166–168
  prospects for study, 175
  red clover necrotic mosaic virus, 175
  West Nile virus, 175

W
West Nile virus (WNV), noncoding RNA, 175
WNV. See West Nile virus
Work, Gibbs free energy change, 323

X
X-chromosome inactivation (XCI)
  long noncoding RNA advantages, 274–276
  noncoding RNA regulators, 269–270
  overview, 268
  RNA in pairing, counting, and choice,
  270–272
  signal sensing by RNA, 288
  symmetry break by Tsix RNA, 272
  unique features, 268–269
XCI. See X-chromosome inactivation
Xist
Polycomb protein regulation, 272–274
signal sensing, 288
X-chromosome inactivation role,
269–270
Xite, X-chromosome inactivation role, 271
The RNA Worlds in Context

Thomas R. Cech

Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, University of Colorado, Boulder, Colorado 80309-0215

Correspondence: thomas.cech@colorado.edu

SUMMARY

The chapters in this collection discuss not one RNA world, but two. The first is the primordial RNA world, a hypothetical era when RNA served as both information and function, both genotype and phenotype. The second RNA world is that of today’s biological systems, where RNA plays active roles in catalyzing biochemical reactions, in translating mRNA into proteins, in regulating gene expression, and in the constant battle between infectious agents trying to subvert host defense systems and host cells protecting themselves from infection. This second RNA world is not at all hypothetical, and although we do not have all the answers about how it works, we have the tools to continue our interrogation of this world and refine our understanding. The fun comes when we try to use our secure knowledge of the modern RNA world to infer what the primordial RNA world might have looked like.

Outline

1 The primordial RNA world
2 The contemporary RNA world
3 The world of RNA technology and medical applications

References
1 THE PRIMORDIAL RNA WORLD

The term “RNA world” was first coined by Gilbert (1986), who was mainly interested in how catalytic RNA might have given rise to the exon—intron structure of genes. But the concept of RNA as a primordial molecule is older, hypothesized by Crick (1968), Orgel (1968), and Woese (1967). Noller subsequently provided evidence that ribosomal RNA is more important than ribosomal proteins for the function of the ribosome, giving experimental support to these earlier speculations (Noller and Chaires 1972; Noller 1993). The discovery of RNA catalysis (Kruger et al. 1982; Guerrier-Takada et al. 1983) provided a much firmer basis for the plausibility of an RNA world, and speculation was rekindled. The ability to find a broad range of RNA catalysts by selection of RNAs from large random-sequence libraries (SELEX) (Ellington and Szostak 1990; Tuerk and Gold 1990; Wright and Joyce 1997) fueled the enthusiasm, and made it possible to conceive of a ribo-organism that carried out a complex metabolism (Benner et al. 1989). The widely accepted order of events is the evolution of an RNA world and from the RNA world to contemporary biology is summarized in Figure 1.

Did an RNA world exist? Some of the most persuasive arguments in favor of an RNA world are as follows. First, RNA is both an informational molecule and a biocatalyst—both genotype and phenotype—whereas protein has extremely limited ability to transmit information (as with prions). Thus, RNA should be capable of replicating itself, and indeed RNA can perform the sort of chemistry required for RNA replication (Cech 1986). Second, it is more parsimonious to conceive of a single type of molecule replicating itself than to posit that two different molecules (such as a nucleic acid and a protein capable of replicating that nucleic acid) were synthesized by random chemical reactions in the same place at the same time. Third, the ribosome uses RNA catalysis to perform the key activity of protein synthesis in all extant organisms, so it must have done so in the Last Universal Common Ancestor (LUCA). Fourth, other catalytic activities of RNA—activities that RNA would need in an RNA world but that have not been found in contemporary RNAs—are generally already present in large combinatorial libraries of RNA sequences and can be discovered by SELEX. Fifth, RNA clearly preceded DNA, because multiple enzymes are dedicated to the biosynthesis of the ribonucleotide precursors of RNA, whereas deoxyribonucleotide biosynthesis is a derivative of ribonucleotide synthesis, requiring only two additional enzymatic activities (thymidylate synthase and ribonucleotide reductase.) Finally, a primordial RNA world has the attractive feature of continuity; it could evolve into contemporary biology by the sort of events that are well preceded, whereas it is unclear how a self-replicating system based on completely unrelated chemistry could have been supplanted by RNA.

Opinions vary, however, as to whether RNA comprised the first autonomous self-replicating system or was a derivative of an earlier system. Benner et al. (this collection) and Robertson and Joyce (this collection) are circumspect, noting that the complexity and the chiral purity of modern RNA create challenges for thinking about it arising de novo. On the other hand, the recent finding that activated pyrimidine ribonucleotides can be synthesized under plausible prebiotic conditions (Powner et al. 2009) means that it is premature to dismiss the RNA-first scenarios. Yarus (this collection), an unabashed enthusiast for an RNA world, argues for a closely related replicative precursor. In vitro evolution studies directed towards an RNA replicase ribozyme continue apace and are of great importance in establishing the biochemical plausibility of RNA-catalyzed RNA replication (Johnston et al. 2001; Zaher and Unrau 2007; Lincoln and Joyce 2009; Shechmer et al. 2009).

What might the first ribo-organism have looked like? Schrum et al. (this collection) describe progress in achieving replication of simple nucleic acid-like polymers within lipid envelopes, thereby constituting “protocells.” These liposomes can grow and upon agitation can divide to give daughter protocells, carrying newly replicated nucleic acids. Whether by lipids or other means, some form of encapsulation must have been a key early step in life. Encapsulation can protect the genome from degradation and

---

Figure 1. An RNA world model for the successive appearance of RNA, proteins, and DNA during the evolution of life on Earth. Many isolated mixtures of complex organic molecules failed to achieve self-replication, and therefore died out (indicated by the arrows leading to extinction.) The pathway that led to self-replicating RNA has been preserved in its modern descendants. Multiple arrows to the left of self-replicating RNA cover the likely self-replicating systems that preceded RNA. Proteins large enough to self-fold and have useful activities came about only after RNA was available to catalyze peptide ligation or amino acid polymerization, although amino acids and short peptides were present in the mixtures at far left. DNA took over the role of genome more recently, although still >1 billion years ago. LUCA (Last Universal Common Ancestor) already had a DNA genome and carried out biocatalysis using protein enzymes as well as RNP enzymes (such as the ribosome) and ribozymes.
predation, allows useful small molecules to be concentrated for the cell's use, and enables natural selection by ensuring that the benefit of newly derived functions accrues to the organism that stumbled across them.

2 THE CONTEMPORARY RNA WORLD

Today, RNA is the central molecule in gene expression in all extant life, serving as the messenger. It is also central to bio-catalysis, seen dramatically in the ribosome but also in ribozymes and RNPzymes such as telomerase and the signal recognition particle. More recently, its diverse roles in regulation of (DNA) gene expression have been discovered. It is useful to organize the discussion of contemporary RNA activities as a spectrum, going from those activities that are so RNA-centered that one could conceive of them having operated in a primordial RNA world very much as they do today, to those that rely more and more on collaboration with proteins, to those RNAs that work on DNA (Fig. 1).

What can RNA do by itself? It can bind small metabolites (such as guanine, S-adenosylmethionine, and lysine) with exquisite specificity, and then use this binding energy to switch from one RNA structure to another. These riboswitches are common regulators of gene expression in Gram-positive bacteria, and are also found in other organisms including plants (Breaker [this collection]; Garst et al. [this collection]). Furthermore, even very small RNAs can act as ribozymes, accomplishing sequence-specific self-cleavage (Ferre´-D’Amaré and Scott [this collection]). These self-cleavers can be easily re-engineered into multiple-turnover RNA-cleaving enzymes, so it is straightforward to imagine that they could have served such a function in a primordial RNA world. Larger ribozymes can accomplish sophisticated RNA splicing reactions, as described for group II introns by Lambowitz and Zimmerly (this collection). There are a number of similarities, both mechanistic and structural, between group II intron self-splicing and spliceosomal splicing of mRNA introns, providing a plausible continuum from the RNA world to post-protein contemporary biology.

Although RNA can perform many activities by itself, in modern cells RNA more often works in concert with proteins. The ribosome uses both RNA and protein to catalyze message-encoded protein synthesis. Yet the heart of the peptidyl transferase center is a ribozyme, and other fundamental activities such as mRNA start-site selection, codon–anticodon interaction, and decoding involve direct RNA–RNA interactions, so the RNA world ancestry of the ribosome is apparent (Moore and Steitz, Noller, Ramakrishnan [all in this collection]). The same can be said of the spliceosome (Will and Lührmann [this collection]). Although a detectable level of catalysis of an isolated step of RNA splicing can be achieved with pure snRNAs (Valadkhan et al. 2009), the efficient and regulated splicing of an entire genome’s collection of primary transcripts requires the collaboration of almost 200 proteins and five snRNAs in the modern spliceosome. Telomerase represents another paradigm, as it includes a canonical protein enzyme (TERT) that operates in intimate collaboration with RNA (Blackburn and Collins [this collection])—so it appears to derive from more recent evolution, after protein enzymes and DNA chromosomes were well established.

It seems likely that the most recently evolved functions of RNA involve regulation of DNA—because there would have been no DNA to regulate in a primordial RNA world! Nevertheless, similar principles could have been active in an RNA world. Gottesman and Storz (this collection) describe RNA regulation in bacteria, which occurs through a range of mechanisms ranging from the simple “antisense RNA” principle of inhibition by complementary base-pairing to RNA–protein interactions. In eukaryotes, several classes of noncoding RNAs perform diverse functions in the regulation of gene expression. Small double-stranded RNAs (for example, the 21-bp small-interfering RNAs and the microRNAs) regulate the stability or the translatability of mRNAs (Joshua-Tor and Hannon [this collection]). Here the RNA provides such a simple function—recognition of complementary sequences on the mRNA target—that the authors choose to organize their discussion according to subfamilies of the Argonaute proteins that bind the small RNAs. The RNA interference (RNAi) pathway is involved not only in mRNA-level events, but also in the regulation of chromatin structure as described by Volpe and Martienssen (this collection). Maintenance of the highly condensed heterochromatin found at chromosome centromeres depends on this RNAi activity. Finally, long noncoding RNAs usually acting in cis (on the chromosome or the local region where they were synthesized) can turn off gene expression by attracting proteins that modify chromatin structure. The effect can spread to an entire chromosome, in the case of the Xist RNA that condenses one of the two X chromosomes in female mammals and thereby gives gene dosage compensation (Lee [this collection]). In other cases, the effect is more local, affecting transcription of a single gene or a group of genes (Wang et al. [this collection]). These recently discovered activities of RNA show that the RNA world never stopped (and has not stopped) evolving.

Diverse viral encoded mRNAs are used as weapons either to circumvent host defense or otherwise manipulate host cellular machinery for their own purposes (Steitz et al. [this collection]). Although several of the classes of viral
ncRNAs are counterparts of cellular equivalents, some are distinctive. Bacteria have evolved the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat) defense system to protect themselves from alien DNA such as that injected by bacteriophages (Wang et al. [this collection]). Here, the information identifying the invading genome is stored in the form of DNA, but it is subsequently converted to small guide RNAs that recognize and interfere with subsequent invaders. Although there is a clear analogy between CRISPR and eukaryotic RNAi, the two systems appear to have evolved completely independently.

3 THE WORLD OF RNA TECHNOLOGY AND MEDICAL APPLICATIONS

I oversimplified when I said that there were two RNA worlds. There is in fact a third—the world of RNA research and development. This third RNA world should be of special interest to students, because this RNA world offers opportunities for gainful employment!

RNA function depends on its structure—it is the seemingly limitless variety of structures that allows so many diverse functions. We can now predict RNA secondary structure quite well (Mathews et al. [this collection]) and see much progress on predicting 3D structure (Westhof et al. [this collection]). Remarkably, we can now watch molecules of RNA fold and unfold and switch from one state to another in “single-molecule experiments” (Tinoco et al. [this collection]). We can use double-stranded RNAs and the intrinsic RNAi machinery present in organisms to do genome-wide knock-downs of gene function (Perrimon et al. [this collection]). Finally, RNA science is poised to make an impact on medicine. For example, aptamers can monitor the concentrations of many of the proteins in human serum, which has diagnostic applications because the presence of many proteins is correlated with health and disease (Gold et al. [this collection]). In addition, both microRNAs and antisense nucleic acids that inhibit miRNAs have pharmaceutical potential, which is under development in numerous biotechnology and pharmaceutical companies.

Thus, the authors of this collection take us on a fascinating journey through three RNA worlds. The primordial RNA world (ca. four billion years ago) relied on the dual ability of RNA to serve as both informational molecule and biocatalyst, providing a self-replicating system. Coupled with other ribozymes that carried out complex metabolism and encapsulated in some sort of envelope, self-replicating RNA constituted an early life form that was the ancestor of contemporary biology. The second RNA world is that of contemporary biology, where RNA occasionally acts by itself (ribozymes and riboswitches) but more often acts in concert with proteins. The ribosome and the spliceosome still “remember” their ribozyme heritage, whereas telomerase and the signal recognition particle have moved on to incorporate canonical protein enzymes. The RNA interference system and CRISPR have gone further, reducing the role of the RNA to that of a simple guide sequence. Finally, the third RNA world—that of RNA technology and medical applications—is a baby compared to even the second RNA world, because it arose only in the past half-century. Although this last RNA world is only perhaps one millionth of one per cent as old as the primordial RNA world, it is a vibrant community, and my co-editors John Atkins and Ray Gesteland and I feel privileged to be part of it.

REFERENCES


