

## AUTHOR INDEX

- Adachi, J., 27  
Adey, N. B., 6, 8, 61, 62  
Aharoni, A., 226  
Akashi, H., 100  
Allard, M.W., 53  
Allemann, R. K., 52  
Anfinsen, C. B., 134  
Arai, K. I., 86  
Ardelt, W., 47  
Arnason, U., 9  
Ashburner, M., 105  
Axe, D. D., 114  
Ayala, F. J., 107
- Babbitt, P. C., 193  
Baker, M. E., 106  
Barabasi, A. L., 149  
Barnard, E. A., 38, 46  
Barrett, P. M., 105  
Bartel, D. P., 222  
Bartlett, G. J., 209–210  
Baudin, A., 105  
Beintema, J. J., 37, 48, 53, 54  
Benner, S. A., 1, 5, 6, 9, 12, 16, 27, 36,  
37, 47, 50, 51, 52, 67, 77, 92, 96, 100  
Benton, M. J., 107  
Berbee, M. L., 100  
Bielawski, J. P., 5  
Blackburn, P., 41  
Bollback, J. P., 31, 32  
Bonaventura, J., 37
- Boulton, B., 94  
Bourne, P. E., 182  
Brand, A. H., 72  
Breukelman, H. J., 37, 52  
Brochier, C., 87, 89  
Buck, M., 3  
Buckley, T. R., 57  
Bull, J. J., 107  
Busslinger, M., 69
- Caetano-Anollés, D., 178  
Caetano-Anollés, G., 178  
Cai, W., 27  
Cao, Y., 76  
Carroll, R. L., 52  
Carsana, A., 52  
Cavalier-Smith, T., 87  
Cavender, T. M., 85  
Chalepakis, G., 69  
Chambon, P., 106  
Chandrasekharan, U. M., 6, 66, 68  
Chang, B. S. W., 6, 75, 76, 77, 78  
Chinen, A., 6, 82, 83  
Churchill, G. A., 31  
Ciglic, M. I., 58  
Collier, L. S., 64, 65  
Collinson, M. E., 100  
Conery, J. S., 51, 54, 100  
Copley, S. D., 213  
Corbin, C. J., 10  
Critchley, H. O., 56

---

*Advances in Enzymology and Related Areas of Molecular Biology, Volume 75:  
Protein Evolution* Edited by Eric J. Toone  
Copyright © 2007 John Wiley & Sons, Inc.

- Cunningham, C. W., 31  
 Czerny, T., 69, 70, 72  
  
 Dahl, E., 69  
 D'Alessio, G., 47, 52  
 Dean, A. M., 112, 113  
 Deeds, Eric J., 133  
 Denkwalter, R. G., 36  
 DeSalle, R., 5  
 Di Cosmo, A., 106, 109  
 Dietmann, S., 5  
 Dixon, B., 85  
 Doggrell, S. A., 65  
 Dokholyan, N. V., 143, 144, 145, 147  
 Domachowske, J. B., 49  
 Donoghue, M. J., 31, 76  
 Douzery, E. J., 53  
 Dubois, J. Y. F., 48  
 Dupuy, A. J., 64  
 Dutilh, B. E., 177  
  
 Eernisse, D. J., 106  
 Ellington, A. D., 51, 96, 223  
 Emmens, M., 37  
 Engelkamp, D., 69  
 England, J. L., 168, 170  
 Escriva, H., 106  
  
 Faculty of 1000, 11  
 Feder, M. E., 7  
 Felsenstein, J., 27  
 Fernandez-Espinar, M. T., 100  
 Fersht, A. R., 96  
 Finkelstein, A. V., 139  
 Fleet, G. H., 94  
 Foote, M., 9  
 Fortin, A. S., 73  
 Franzen, J. L., 9  
  
 Gastra, W., 37  
 Galtier, N., 27, 86  
 Ganzhorn, A. J., 99  
  
 Gaschen, B., 90  
 Gaucher, E. A., 5, 6, 9, 86, 87, 89,  
 92, 93  
 Gerloff, D. L., 5  
 Gerlt, J. A., 193  
 Gerstein, M., 178  
 Giguere, V., 106  
 Glasner, M. E., 193  
 Glenner, H., 3  
 Golding, G. B., 113, 120  
 Goldstein, R. A., 31, 157, 158, 171, 172  
 Gonnet, G. H., 27, 107  
 Gould, S. J., 3  
 Gouy, M., 27  
 Graur, D., 43  
 Green, S., 106  
 Groen, G., 37  
 Gromiha, M. M., 86  
 Gruber, M., 37  
 Guerrero, R., 59  
  
 Halder, G., 72  
 Harder, J., 47  
 Hargrave, T. B., 56  
 Hasegawa, M., 27  
 Hassanin, A., 53  
 Heard, G. M., 94  
 Hedges, S. B., 9  
 Hermanson, S., 64  
 Hernandez Fernandez, M., 53  
 Hesse, M., 4  
 Hey, J., 8  
 Hillis, D. M., 107  
 Hilvert, D., 241  
 Hirschmann, R., 36  
 Hirst, J. D., 153  
 Holm, L., 5, 137, 138  
 Hooker, J. J., 100  
 Horowitz, N. H., 196, 198  
 Huelsenbeck, J. P., 27, 29, 31, 32, 76  
 Hugenholtz, P., 86, 87  
 Hurley, J. H., 112

- Imada, K., 112  
Ivics, Z., 6, 62, 64  
Iwabata, H., 6, 90, 91
- James, K., 56  
Janis, C. M., 46  
Jekel, P. A., 37  
Jenkins, S. R., 36  
Jensen, R. A., 196  
Jermann, T. M., 6, 7, 40, 43, 46, 52, 53, 54  
Jukes, T. H., 52  
Jurgens, C., 209  
Justice, M. J., 64
- Kelemen, B. R., 56  
Kellis, M., 99, 104  
Kelly, R. W., 56  
Kelmanson, I. V., 110  
Kim, J. S., 58  
Kimura, M., 52  
Kleineidam, R. G., 53  
Knauth, L. P., 86, 93  
Knight, P. A., 65  
Koch, A. L., 228  
Koonin, E. V., 141, 162, 178  
Koshi, J. M., 31  
Kreitman, M., 100  
Krishnan, N. M., 31  
Kruiswijk, C. P., 85  
Kumar, S., 9  
Kuper, H., 37
- Lang, K., 42, 43  
Lee, J. E., 58  
Lenstra, J. A., 37  
Lin, J., 178  
Logue, A. D., 36  
Lorenz, E. N., 8  
Lynch, M., 51, 54, 100
- Maddison, D. R., 27  
Maddison, W. P., 27
- Malcolm, B. A., 6, 59, 60  
Margoliash, E., 37  
Margulis, L., 59  
Marshall, C. R., 52  
Martena, B., 37  
Martins, E. P., 31  
Matousek, J., 47  
Matsumura, I., 223  
Matz, M. V., 110  
McClellan-Green, P., 109  
McGovern, P. E., 94  
Menon, S. T., 75  
Miller, H. R., 65  
Mirny, L. A., 158  
Miskey, C., 63  
Mitehell-Olds, T., 7  
Miyazaki, J., 6, 90  
Moers, A. O., 31  
Moore, S., 36, 41  
Murzin, A. G., 137, 138  
Muskiet, F. A. J., 37
- Nambiar, K. P., 5, 36, 40  
Navidi, W. C., 76  
Nei, M., 25, 27, 70  
Neidhardt, F. C., 63, 113  
Nielsen, R., 25  
Nobeli, I., 205  
Nock, S., 86
- Oberdorster, E., 109  
O'Harra, C. C., 9  
Ohno, S., 51, 85  
Okabe, Y., 47  
Omland, K. E., 31  
Onorato, J., 49  
Opitz, J. G., 43, 52, 53
- Pagel, M., 25, 31, 32  
Pauling, L., 4  
Perrimon, N., 72  
Peterson, K. J., 106

- Philippe, H., 87, 89  
Posada, D., 57  
Presnell, S. R., 5  
Pretorius, I. S., 101, 105  
Preuss, K. D., 52  
Pupko, T., 27
- Raillard, S. A., 53  
Raines, R. T., 58  
Ree, R. H., 31  
Richards, F. M., 36  
Riggs, A., 37  
Rosen, M. K., 3  
Rosenberg, H. F., 6, 49, 50  
Rost, B., 5  
Runnegar, B., 86
- Sanangelantoni, A. M., 86  
Sander, C., 137, 138  
Sassi, S., 1, 54, 56, 57  
Saunders, M., 36  
Schaaff, I., 101  
Schluter, D., 31  
Schmidt, D. M. Z., 203  
Schmidt, F. X., 42, 43  
Schreiber, D., 37  
Schroder, J. M., 47  
Schroder, W., 56  
Schultes, E. A., 222  
Schultz, T. R., 31  
Shagin, D. A., 109  
Shakhnovich, B. E., 170, 172  
Shakhnovich, E. I., 133, 143, 144,  
158, 168, 170  
Shi, Y., 6, 78, 80, 81, 82  
Shields, D. C., 104  
Singhania, N. A., 48  
Soucek, J., 47, 52  
Stackhouse, J., 5, 6, 7, 16, 24, 38, 40  
Stein, W. H., 36  
Strachan, R. G., 36
- Stroband, H. W. J., 85  
Strydom, D. J., 47  
Stryer, L., 96  
Sun, G., 100  
Sun, H. M., 6, 69, 70, 71, 72, 73, 74  
Swofford, D. L., 27, 96
- Tauer, A., 5  
Taverna, D. M., 157, 158, 171, 172  
Taylor, J. W., 100  
Thompson, J. D., 112  
Thomson, J. M., 6, 29, 96, 98  
Thornton, J. M., 136, 137  
Thornton, J. W., 5, 6, 27, 105, 106, 107,  
108, 109  
Tiana, G., 170, 172, 174  
Trabesinger-Ruef, N., 40  
Trautwein, K., 53
- Ugalde, J. A., 6, 109, 110, 111  
Underhill, D. A., 69, 73
- Vamcava, K., 241  
Vandenberg, A., 37  
Vandijk, H., 37  
van Heyningen, V., 69  
Veber, D. F., 36  
Venter, J. C., 3, 93  
Vos, J. P., 52  
Vrba, E. S., 53
- Wade, N., 7  
Wanstall, J. C., 65  
Weiser, K. C., 64  
Welling, G. W., 37  
West, O., 59  
Wilke, S. E., 81  
Willis, K. J., 105  
Wills, C., 94  
Wilson, A. C., 59  
Wingreen, N., 151

- Woese, C. R., 86, 176  
Wolfe, K. H., 104  
Woycechowsky, K. J., 241
- Xu, H. E., 73
- Yang, Z., 25, 27, 31, 76, 78, 87, 96,  
107, 110  
Yang, Z. H., 5
- Yant, S. R., 64  
Yasutake, Y., 112  
Yokoyama, S., 6, 78, 80, 82
- Zhang, J., 6, 25, 27, 47, 49, 50, 51, 70  
Zhao, W., 48  
Zhu, G., 6, 112  
Zilliacus, J., 108  
Zuckerkindl, E., 4



## SUBJECT INDEX

Page references followed by t indicate material in tables.

- Absorption spectrum, amino acids  
causing a difference in, 84
- Acetate, bacterial adaptation to growth  
on, 113
- N*-Acetylneuraminylase (NAL),  
247–248
- Acid–base mechanisms, 208
- Acid–sugar dehydratase families,  
216–220
- Active-site architecture-constrained  
evolution, in suprafamilies, 209–214
- Active-site architecture-constrained  
suprafamily, 198
- Active site mutations, as activity  
switches, 247–248
- Active-site residue transplantation,  
enzyme redesign by, 253–254
- Active sites  
changing the chemical properties of,  
263–268  
molding new, 255–256
- Activity swapping, in enzyme  
superfamilies, 243–249
- Activity switches, active site mutations  
as, 247–248
- N*-Acylamino-acid racemase  
(NAAAR), 220–221
- Adaptive evolution, 51, 56–57
- Adh1/2* duplication, 100. *See also*  
Alcohol dehydrogenase entries
- Ala293Asp* mutation, 267–268
- L*-*Ala-D/L*-Glu epimerase (AEE), 245–247
- Alcohol dehydrogenase, 93–105  
timing of expression of, 94–96
- Alcohol dehydrogenase 1 (*Adh1*), 94, 96  
ethanol range for, 99  
kinetic properties of, 98t
- Alcohol dehydrogenase 2 (*Adh2*), 96  
kinetic properties of, 98t
- Alcohol dehydrogenase ancestor  
(*ADH<sub>A</sub>*), 96, 98t, 99  
assignment of a primitive function to,  
100
- Aldolase antibodies, 272–273
- Aldolases, designer, 249–250
- Alignment scoring matrix, 13–14
- Alkaline phosphatase, 220
- Ambiguity in “deep time”  
paleogenetics studies, 93  
in elongation factor trees, 87  
management of, 31–35, 58, 114
- Amidohydrolase superfamily, 198
- Amino acid changes, in ANI and ANII  
DNA-binding properties, 71–72
- Amino acid replacements, 17–20  
distinguishing, 69–70  
impact of, 34  
models of, 79–80  
parsimony analysis inference  
of, 24–25

---

*Advances in Enzymology and Related Areas of Molecular Biology, Volume 75:  
Protein Evolution* Edited by Eric J. Toone  
Copyright © 2007 John Wiley & Sons, Inc.

- Amino acids  
  difficulty in reconstructing, 30  
  discovery of, 80–81
- Amino acid sequence evolution,  
  patterns of, 17
- 4-Aminophenol oxidation, generating a  
  novel metal-dependent catalyst of,  
  286–287
- Aminotransferases, 265–266
- Ancestor I (AnI) binding, 70–72
- Ancestor II (AnII) binding, 72
- Ancestors, gene duplication of, 73–74
- Ancestral coral fluorescent proteins,  
  109–111
- Ancestral *Pax* sequences, 70–71
- Ancestral promoters, resurrected, 62
- Ancestral protease, 67–68
- Ancestral protein sequence modeling,  
  model hierarchy for, 17–26
- Ancestral residues, inferred, 25
- Ancestral rhodopsins, models for  
  reconstructing, 76–78
- Ancestral ribonucleases, resurrecting  
  from artiodactyls, 37–46
- Ancestral RNase analysis, evolutionary  
  tree used in, 45
- Ancestral seminal proteins,  
  reconstruction and resurrection of,  
  54–56
- “Ancestral Sequence Reconstruction”  
  conference, 31
- Ancestral sequences  
  computer programs for inferring, 27  
  constructing, 92  
  methods for inferring, 25–26  
  most probabilistic, 87  
  sources of ambiguity in, 28  
  topologies used to reconstruct, 87–89
- Ancestral states, drawing inferences  
  about, 17–20, 28
- Ancestral steroid receptor (AncSR1),  
  106, 107–108  
  planetary context of, 109  
  resurrecting, 105–109
- Ancestral SWS2 pigment, 83–84
- Ancestral transposable elements, 60–61
- Ancient RNases, laboratory  
  reconstructions of, 52
- AncSR1-DBD, 108
- AncSR1-LBD, 108
- Antibodies, catalytic, 269–275
- Antibody-catalyzed aldol condensation  
  products, selectivity of, 274–275
- Antibody esterases, 272
- Aplysia californica*, estrogen receptor  
  sequence from, 106–107
- Aquificaceae, in the bacterial  
  lineage, 87
- Archae, origins of, 90–93
- Archaeomeryx*, 39, 41–43, 45, 46
- Archosaurs, rhodopsins from, 75–78
- aromatase* genes, 9–11
- AroQ CM (chorismate mutase), 257,  
  259, 260
- Artiodactyls, 46–47  
  evolutionary tree ambiguity of, 43  
  resurrecting ancestral ribonucleases  
  from, 37–46
- Aspartate aminotransferase (AspAT)  
  substrate specificity, changing,  
  265–268
- ATB17 variant, 266–267
- Automated design  
  of catalytic sites, 277–281  
  of ligand binding sites, 275–277
- Bacteria  
  adaptation to growth on acetate, 113  
  thermophily in, 85–86
- Bacterial genomes, large-scale  
  sequencing of, 93
- Bacterial lineage, elongation factor  
  sequences in, 86–87
- ( $\beta/\alpha$ )<sub>8</sub>-barrel proteins, 214, 215

- Barnard hypothesis, 38, 46
- Bayesian analysis, 32
- $\beta$ -glucuronidase, directed evolution of, 223–225
- Big bang (BB) model, for protein structural evolution, 147–149, 157, 165
- Binding motifs, incorporating into template scaffolds, 214
- Binding pockets, for hydrophobic ligands, 153
- Biological activities, ribonuclease homologs involved in, 47–48
- Biological data, meaning of, 7
- Biology, roots in natural history, 3
- Biomolecular behavior change, in adaptive evolution, 57
- Biomolecular function, protein structure space and, 51–52
- Biomolecular resurrections, 7  
in evolutionary narratives, 11–12
- Biomolecular sequences, divergent evolution of, 12–13
- Biomolecular systems exploration of, 5–7  
narratives describing, 115
- Biophysical methods, 3
- Biosynthesis enzymes, histidine and tryptophan, 206
- Birth, death, and innovation models (BDIMs), 141–144
- Blue opsins, 83–84
- Bovine seminal fluid, ribonuclease homologs in, 51–58
- Bulk ethanol, emergence of, 105
- Cancer genes, identifying, 64–65
- Candidate ancestral protein synthesis, ambiguity management using, 34–35
- Catalysis, as an evolutionary constraint, 196
- Catalysts. *See also* Catalytic entries  
design of, 268  
ready-made, 242
- Catalytic antibodies, 269–275  
as a source of novel protein catalysts, 287
- Catalytic capabilities, of enzyme superfamilies, 249–256
- Catalytic mechanisms, conservation of, 194–195
- Catalytic promiscuity  
activity switching resulting in, 248  
examples of, 217–219t  
in Tyr265Ala Alr, 252, 253
- Catalytic proteins, *de novo* design of, 285–287
- Catalytic residues, in active-site architecture-constrained suprafamilies, 209–213
- Catalytic sites, automated design of, 277–281
- CATH (Class–Architecture–Topology–Homology)  
database, 137  
wheel, 136
- Charge complementarity, 269–272
- 4-Chlorobenzoyl-CoA (4-CBA-CoA) dehalogenase, 243–245
- Chorismate mutases (CMs), 256–257, 259–263
- Chymase–angiotensin converting enzyme, 65–68
- Chymases  
divergence of, 68  
substrate specificity of, 65–66
- Cladistic analysis, 9
- Class III ligase ribozyme, 222
- Cluster-size distribution, in the protein domain universe graph, 146–147
- Clusters of orthologous groups (COGs), 176, 177

- Coenzyme specificities, amino acids responsible for, 112
- Cofactor specificity, divergence of, 112
- Combinatorial mutagenesis, in restricting amino acid usage, 259
- Computational redesign, 276–281  
outcomes of experiments in, 288
- Computational screening, 284
- Computer analysis, of protein sequences, 5
- Computer programs, for multiple sequence alignments, 27
- Conformational diversity, 262
- Consensus reconstructions, 64
- Consensus sequences, 62
- Consensus tools, 28
- Conservation of catalysis, in protein design, 203–204
- Conservatism of conservatism (CoC) entropies, 144
- Conservative enzyme redesign, 243–256
- Contact density (CD), 172–174
- Convergent–divergent debate, 134–135, 139–140, 144, 151–152, 167–168
- Coral fluorescent proteins  
ancestral, 109–111  
planetary biology of, 111
- Cretaceous, changing ecosystem in, 93–105
- Crotonase activity, 243–245
- Crotonase superfamily, 198
- Cryptic genes, 215  
in evolution of new protein functions, 227–228
- Customized enzymes, on-demand, 288
- Cytochrome C oxidases, 37
- Cytochrome maturation proteins (CMPs), 213
- DALI structural comparison algorithm, 138
- DALI Z-score, 145, 146
- Darwinian strategy, 3
- “Deep time” paleogenetics studies, 93
- Degree distribution  
of a protein domain universe graph, 147–149  
random-graph-like, 154–155
- Dehydratase families, 216–220
- Dehydrogenases, isopropylmalate and isocitrate, 90–93
- De novo* design  
of catalytic proteins, 285–287  
of protein structures, 283
- Designability, 135, 139, 152  
influence on sequence and structure dynamics, 166–167  
influence on structural evolution, 170–173  
protein evolution and, 168–175
- Designability measure, 169–170  
on the basis of structural information, 174
- Designer enzymes, 242
- Designer proteins, from scratch, 281–287
- Dihydrolipicolinate synthase (DHDPS), 248
- Diketone haptens, 272
- Directed evolution, 223–227, 256–268, 287–288  
for AspAT variants, 266–267  
to improve TIM activity, 280  
novel enzyme activities and, 263–264
- Direct sequence comparisons, 176–177
- Disordered proteins, 262
- DNA binding domain (DBD), activity of, 106
- DNA regulatory elements, 68–69
- Dollo constraint, 180
- Dollo parsimony, 180–181
- Domain “death” mechanisms, 141–143

- Domain evolution, theoretical models of, 179–181
- Domain families, duplication of, 141–143
- Early life, role of thermophily in, 85–86. *See also* Life on Earth
- ECyP cyclophilin, redesign of, 253
- Edman degradation, 53–54
- Effective molarity (EM), 270–271
- EF proteins, thermostability test for, 91
- Elongation factors (EFs), 86–90  
inferences drawn from, 90
- Engineered enzymes, promiscuity of, 222–227
- England–Shakhnovich (ES)  
designability measure, 170, 171–172
- Enolase superfamily, 198, 200–203, 245–247  
active sites of, 248  
engineering of, 225  
reactions catalyzed by, 201, 202, 247  
subgroups of, 202–203
- 2-Enoyl-CoA hydratase/isomerase superfamily, 244–245
- Enzyme design, template selection for, 195–196. *See also* Enzyme redesign
- Enzyme engineering, difficulty of, 203
- Enzyme reactivity, altering, 288
- Enzyme redesign. *See also* Existing enzymes by active-site residue transplantation, 253–254  
rational, 243–244
- Enzymes. *See also* Existing enzymes changing the physical properties of, 257–263  
conservative redesign of, 243–256  
promiscuous, 216–227  
robustness to mutation, 226–227
- Enzyme simplification, 257–259
- Enzyme superfamilies. *See also* Superfamilies  
activity swapping in, 243–249  
extending catalytic capabilities of, 249–256
- Eosinophil cationic protein (ECP), 49–51
- Eosinophil-derived neurotoxin (EDN), 49–51
- Eosinophil RNase homologs,  
paleochemistry with, 49–51
- Eotragus*, RNase sequences in, 39
- Equilibrium convergence models, 165
- Ester hydrolysis catalysts, generating, 286
- Estrogen activation, specificity of, 108
- Estrogen signaling, origin of, 105–109
- Eukaryotes, cryptic gene activation in, 227–228
- Eukaryotic domains, designability of, 171
- Evolution  
adaptive, 56–57  
directed, 256–268  
functional diversification during, 71  
simple theory of, 17–18
- Evolutionary analysis, 7–11
- Evolutionary bioinformatics, 116
- Evolutionary history, effect on family size, 172–173
- Evolutionary hypotheses, resistance to experimental testing, 8
- Evolutionary models, terminology related to, 196–198
- Evolutionary narratives, biomolecular resurrections in, 11–12
- Evolutionary pathways, intermediates in, 215–230
- Evolutionary relatedness, calculating, 175–176
- Evolutionary trees, protein, 14–15

- Existing enzymes, redesign of, 242–268
- Experimental paleobiochemistry, 12–28  
for exploring cofactor specificity divergence, 112
- Experimental paleogenetics, 5, 11  
ambiguity in, 29–31  
evolutionary theories and tools in, 36–113
- Extinct organisms, molecular systems from, 7
- eyeless* gene, ectopic expression of, 72–73
- Families. *See also* FSSP (Families of Structurally Similar Proteins) database  
dates of divergence of, 85  
defined, 197
- Family DNA shuffling, 196
- Family size, designability as a predictor of, 172
- Fish, green opsin from, 82–83
- Folding cutoff, 160
- Fold size  
distribution of, 142  
heterogeneity in, 141
- Frog Prince (FP) transposon system, 63–64
- F*-scores, 158–160
- FSSP (Families of Structurally Similar Proteins) database, 137
- F-type sequences, 61–62
- Functional promiscuity, 195
- Gene content comparisons, 176
- Gene conversion, 229
- Gene duplication, 216
- Gene therapy, 64
- Genetic selection, 256
- Genomic sequences, history conveyed by, 4–5
- Giant component (GC), of the protein domain  
universe graph, 145–146
- Global biotherapy, 115
- Glu323Gly MLE II, 264
- Gluconate/galactonate dehydratase, 220
- Glu<sup>H50</sup> activation, 271
- Glutathione transferase A1-1 (GSTA1-1), redesign of, 254
- Goldfish  
blue opsins in, 83–84  
visual pigments in, 82
- Graph-theoretic approaches, implications of, 160–161
- Green fluorescent protein (GFP), 109
- Green fluorescent protein-like proteins  
convergent evolution of, 111  
phylogeny of, 110
- Green opsin, 82–83
- Green (RH2) opsin genes, 82–83
- Haldane analysis, 99–100
- Haloacid dehalogenase superfamily, 198
- Hamming distance, 181
- Hapten design strategy, 269
- Haptenic charge, strategic use of, 270
- Hemoglobins, 37
- Hepatitis delta virus (HDV) ribozyme, 222
- HEX + Ala293Asp variant, 267–268
- HEX variant, 267–268
- Hierarchical Bayesian approaches, 32
- Hierarchical models of inference, ambiguity management using, 31–32
- HisA substrate specificity, changing, 264–265
- HisA/TrpF enzyme, 216. *See also* TrpF
- Histidine biosynthesis, 205–209  
enzymes in, 206–207
- Historical biology, 3–4
- Historical models, ambiguity in, 28–35

- Homologous sequences, 12, 13  
 Homologs, sequence, 137–138  
 Homology modeling, 267–268  
 Homoplasy, 29  
 Horizontal gene transfer (HGT), 176  
 Hsp90 function, 228  
 Human maleylacetoacetate isomerase (MAAI), 220  
 3 $\alpha$ -Hydroxysteroid dehydrogenase (3 $\alpha$ -HSD), 245  
 Hyperthermophilic hypothesis, for last universal common ancestor, 93  
  
 Immune response, defensive, 56  
 Immunization, reactive, 272–275  
 Inert scaffolds, conferring activities to, 268–281  
 Inference, hierarchical models of, 31–32  
 Intermediates, in evolutionary pathways, 215–230  
 Inverse folding solutions, 282–285  
 Isocitrate dehydrogenases (IDH), 90–93  
   NADP-dependent, 112–113  
 Isocitrate lyase (ICL), 113  
 Isopropylmalate dehydrogenases, 90–93, 112–113  
  
 “Junk DNA,” 51–52  
 “Just so” stories, 8–9, 46  
   converting into scientific narratives, 59, 82, 96  
  
 KGPDC suprafamily, 210, 213  
  
 Large-scale bacterial genome sequencing, 93  
 Last universal common ancestor (LUCA), hyperthermophilic hypothesis for, 93  
 Lateral gene transfer (LGT), 176–177, 178  
   prevalence of, 179  
   protein structural domains and, 182  
  
 Lateral structural domain transfer (LSDT), 179–180  
 Lattice polymers  
   as models of protein structure and behavior, 149–153  
   sequence evolution of, 151  
 Lattice polymer structures, designability of, 168  
 Lattices, structural evolution on, 153–161  
 Lattice spaces, 154  
 Lattice structure graph (LSG), 154–161  
   results from analysis of, 155  
 Lattice structures, choosing, 155–157  
 Lattice systems  
   modeling higher levels of functionality in, 153  
   protein evolution in, 166–167  
 Least universal common ancestor (LUCA), 171, 183  
 Life on Earth, coherent narrative for, 94.  
   *See also* Early life  
 Ligand-binding domain (LBD), 106  
   functionality of, 107–108  
 Ligand-binding sites, automated design of, 275–277  
 Ligand substructures, conservation of, 205  
 Long interspersed nuclear elements (LINES), 61  
 Long interspersed repetitive elements of type 1 (LINE-1), 61–62  
 L-threonine aldolase (TA), comparison with alanine racemase, 251  
 Lysozymes, neutrality and parallel evolution in, 59–60  
  
 MacClade program, 27  
 “Majority rule” consensus strategy, 62–63, 77

- Make–accumulate–consume yeast strategy, 100, 104–105
- Mammal reconstructions, 33
- Mammals, ribonucleases from, 36–59
- Maximum likelihood (ML) analysis, 22, 24  
programs for, 27  
versus maximum parsimony methods, 25–26
- Maximum likelihood evolutionary tree, 96, 97
- Maximum likelihood tools, 29  
value of, 58–59
- Maximum parsimony (MP) methods, 180  
versus maximum likelihood analysis, 25–26
- Maximum parsimony theory, 20
- Mechanistically diverse superfamilies, 197, 198–200
- Medicine, incorporating natural history into, 115–116
- Metazoan endocrine system, 109
- Metazoan phyla, 33
- Microbial world, biodiversity in, 33
- Minimum change concept, 20
- mMjCM (chorismate mutase), 261–263
- Models  
“nonergodic” characteristics of, 160  
statistical tests on, 29  
using topological observables to compare, 157
- Molecular biology, role for history in, 3–7
- Molecular characterization, failure of, 4
- Molecular paleoscience, potential of, 59
- Molecular records, correlating with paleontological records, 15–17
- Molecular resurrections, 6t
- Molecular techniques, for calculating evolutionary relatedness, 175–176
- Molten globules, 261–262, 263
- Montastraea cavernosa*, color diversification in, 110–111
- Most probabilistic ancestral sequence (MPAS), 87
- Most probabilistic ancestral state, 96
- Mouse RNase genes, 48
- Muconate lactonizing enzyme II (MLE II), 263–264  
directed evolution of, 225
- Multiple candidate ancestral protein synthesis, ambiguity management using, 34–35
- Mutagenesis experiments, 53
- Mutation(s)  
accumulation of, 228  
in cryptic genes, 227  
robustness to, 226–227
- N*-acetylneuraminate lyase (NAL), 247–248
- N*-acylamino-acid racemase (NAAAR), 220–221
- NAD-dependent isopropylmalate dehydrogenase (NAD-IMDH), 112
- NADP-dependent isocitrate dehydrogenases (IDH), 112–113
- Natural enzymes, promiscuity of, 216–221
- Natural selection, 3
- Natural world, structural diversity in, 134
- Neighbor joining (NJ) method, 181, 182
- Neutral drift, 35
- Neutralist–selectionist debate, 8
- Neutral network concept, 150–151
- Neutral paths, enzyme evolution along, 222
- p*-Nitrophenylacetate, hydrolysis of, 277
- Novel enzymes, design of, 277–278, 287
- Novel reactions, redesigning enzymes to catalyze, 195
- Novo-TIMs, 280–281

- Null convergent model of structural evolution, 161–165  
 testing of, 163
- Octarellins, 284
- OMPDC suprafamily, 210–213
- Oncogenes, identifying, 65
- Open reading frames (ORFs), 48, 63
- Opsins  
 blue, 83–84  
 green, 82–83  
 planetary biology of, 84–85  
 vertebrate, 74
- Organismal PDUGs (oPDUGs),  
 constructing, 162–165. *See also*  
 Protein domain universe graph  
 (PDUG)
- o*-succinylbenzoate synthase (OSBS),  
 245–247  
 activity of, 220–221, 263–264
- o*-succinylbenzoate synthase enzymes,  
 203–204
- o*-succinylbenzoate synthase reaction,  
 catalysis of, 225
- Outgroups, 14–15, 23
- Oxaloacetate decarboxylation, peptide  
 capable of promoting, 285–286
- 4-Oxalocrotonate tautomerase (4-OT),  
 redesign of, 254–255
- Pachyportax*, 16–17  
 RNase sequences in, 38
- Pairwise sequence alignment, 12–14
- Paleobiochemical experiments, 11  
 RNases in, 37
- Paleobiochemistry, 5  
 with eosinophil RNase homologs,  
 49–51  
 family management via, 81  
 with ribonuclease homologs in bovine  
 seminal fluid, 51–58
- Paleobioscience, 71
- Paleogenetic experiments, inferences  
 from, 58
- Paleogenetic paradigm, ambiguity's  
 defeat of, 35
- Paleogenetics studies, “deep time,” 93
- Paleomolecular biology, 5
- Paleomolecular resurrections, 94  
 biological understanding via,  
 108–109  
 protein behaviors identified by, 115
- Paleontological records, correlating  
 with molecular records, 15–17
- Paleosystems biology, 5
- Parallel evolution, in lysozymes, 59–60
- Paralogs  
 assigning primitive behavior to,  
 99–100  
 duplication of, 100–104
- Parsimony, 20
- Parsimony analysis, 23–24. *See also*  
 Maximum parsimony entries;  
 PAUP\* program  
 of serine proteases, 66–68
- Partial chemical reactions (PCRs),  
 194, 198
- PAUP\* program, 27
- Pax-6* homologs, ectopic expression of,  
 72–73
- Pax* domains, evolution of, 73–74
- Pax* genes, sequence analysis of, 69
- Pax* system, 68–74
- “Perfect enzyme,” 280
- Periplasmic binding proteins,  
 computational redesign of, 276
- Per-monomer free energy ( $f$ ),  
 expression of, 168–169
- Peroxiredoxin suprafamily, 212,  
 213–214
- Phenylserine aldolase activity, 250
- Phosphate-binding pocket, 208
- Phosphotriesterase, 221
- Phylogenetic categorization, 176

- Phylogenetic information, protein  
  structure distributions as a source  
  of, 177–178
- Phylogenetic reconstructions, 76–78
- Phylogenies, 175–176  
  based on protein structural domains,  
  179
- Planetary biology  
  of coral fluorescent proteins, 111  
  of opsins, 84–85
- Polypeptides, structural space for,  
  153–154
- Population genetics, 151
- Posterior probabilities, 96
- Preferential attachment models, 149
- PRETTY program, 62
- Primate evolution, EDN and ECP  
  mutations in, 49–50
- Primitive behavior, assignment to  
  paralogs, 99–100
- Primordial protein synthesis, 257
- Probabilistic ancestral sequences, 20  
  differences in, 29
- Probabilistic model, 22
- Processed pseudogenes, 229
- Prokaryotes, metabolism in, 229
- Promiscuity  
  catalytic, 217–219  
  degrees of, 216  
  increase in, 226  
  in protein design, 229–230
- Promiscuous enzymes, 216–227  
  engineered, 222–227  
  natural, 216–221  
  recently evolved, 221–222
- Promiscuous intermediates, 222,  
  223–227  
  in evolutionary pathways, 215–230
- Promiscuous proteins  
  constraints on, 215  
  design of, 229–230  
  evolutionary plasticity of, 230–231
- Proteases  
  diversity of, 68  
  specificity of, 65–68
- Protein activities, genetic sources for,  
  215
- Protein catalysts, 256  
  novel, 269
- Protein Data Bank, 275
- Protein design, 135, 139  
  conservation of catalysis in,  
  203–204  
  exploiting promiscuity in,  
  229–230  
  exploiting suprafamilies in,  
  214–215
- Protein domain universe graph (PDUG),  
  144–149. *See also* Organismal  
  PDUGs (oPDUGs)  
  random subgraphs of, 163  
  topological features of, 147
- Protein domains, 137–138
- Protein engineering, 50, 194–195  
  novel enzyme generation in, 242  
  success of, 230  
  template selection for, 203–204,  
  230
- Protein evolution, 143, 193–239  
  big bang model for, 147–149  
  conservation in, 195–215  
  convergence versus divergence  
  paradigms for, 134–135, 139–140,  
  144, 151–152  
  degree distributions from sequence-  
  based models of, 159  
  designability and, 168–175  
  dominance of active-site architecture  
  in, 210–213  
  structure-centric view of, 133–135
- Protein family evolution model  
  building, 12–17
- Protein folding, 276, 282–285  
  lattice polymer systems and, 150

- Protein folds  
 distribution of, 139–140  
 estimating the number of, 140–141
- Protein–protein interactions,  
 276–277
- Proteins. *See also* Enzyme entries  
 designability increase in, 171  
 disordered, 262  
 percent sequence identity between  
 ancestral and modern, 89t  
 unnatural-fold, 284–285
- Protein sequence divergence,  
 sophisticated theories of, 20–22
- Protein sequence evolution, theories  
 of, 17
- Protein sequences, computer analysis  
 of, 5
- Protein structural domains, lateral gene  
 transfer and, 182
- Protein structure(s)  
 classification of, 135–138  
*de novo* designed, 283  
 lattice polymers as models of,  
 149–153
- Protein structure space, 51–52
- Pseudogenes, 215, 228–229
- P*-values, random subgraph, 164
- Quaternary enzyme assembly,  
 simplifying, 259–260
- Rana pipiens* genome, 63
- Random design, intersection with  
 rational design, 269
- Random-graph-like degree distribution,  
 in the lattice structure graph,  
 154–155
- Random subgraph *P*-values, 164
- Random subgraphs, 163–164
- Rate matrix, 14
- Rational design, intersection with  
 random design, 269
- Rational enzyme redesign, 243–244
- Rat RNase genes, 48
- Reactive immunization, 272–275
- Reactive immunogens, design of,  
 273–274
- Recently evolved enzymes, promiscuity  
 of, 221–222
- Reconstructed ancestral ribonucleases,  
 kinetic properties of, 41t
- Reconstructed ancient ribonucleases  
 catalytic activity of, 43–46  
 sequence changes in, 44t  
 thermal transition temperatures for,  
 42t
- “Reconstructing Ancestral Character  
 States” symposium, 31
- Reconstructions, sources of ambiguity  
 in, 28–31
- Red fluorescent proteins, 111
- Red/green color diversification,  
 in coral fluorescent proteins,  
 109–110
- Regulatory systems, resurrection of,  
 68–74
- Resurrected proteins, behaviors of, 50
- Resurrected sequences, information  
 provided by, 114
- Resurrection, of regulatory systems,  
 68–74
- Retrograde evolution hypothesis, 205,  
 209
- Retrotransposons, 60, 61
- RH2 pigments, ancestral, 82
- Rhodopsin genes, multiple sequence  
 alignment for, 75
- Rhodopsins, from archosaurs, 75–78
- RIBAses, 47
- Ribonuclease homologs  
 in biological activities, 47–48  
 in bovine seminal fluid, 51–58
- Ribonuclease resurrections, lessons  
 learned from, 58–59

- Ribonucleases (RNases). *See also*  
RNase evolutionary models  
ancestral, 37–46  
cell-based activities of, 58  
composition of, 36  
as digestive enzymes, 38  
functions of, 36  
from mammals, 36–59  
multiple sequence alignment of,  
14–15  
in vertebrates, 47–48
- Ribose-binding protein (RBP),  
conversion into triose phosphate  
isomerase, 278–281
- Ribozyme-design experiment,  
222–223
- RNase A family, 37
- RNase evolutionary models, 54–56.  
*See also* Ribonucleases (RNases)
- Ruminant artiodactyls, 46–47
- Ruminant digestion, origin of,  
46–47
- Ruminant digestive physiology, 38
- S94A mutation, 83, 84
- S295C mutation, 84
- Saccharomyces cerevisiae* genome,  
duplication in, 102–103t
- Scaffolds, inert, 268–281
- “Scale-free” fold-size distribution, 141
- Scale-free networks, 147, 149  
evolving, 156–157  
sequence dynamics and, 157
- SCOP (Structural Classification of  
Proteins)  
database, 137
- Scoring functions, 13–14
- Scylatone dehydratase activity, rational  
design of, 255–256
- Seminal genes, expression of, 53–54
- Seminal ribonucleases, divergence  
of, 55
- Seminal RNases, 52–58  
immunosuppressivity of, 57–58  
properties of, 56
- Sequence alignment, 12–14
- Sequence ambiguity, managing, 31–35
- Sequence-based model, of structural  
evolution, 166
- Sequence collection, ambiguity  
management using, 32–33
- Sequence comparisons, 176–177
- Sequence data, trees inferred from, 26
- Sequence database, ambiguity in, 28
- Sequence homologs, 137–138
- Sequences, *F*-scores of, 158–160
- Sequence “similarity,” scoring  
function and, 13
- Sequence–structure pockets, 143–144
- Serine proteases, 65
- Short interspersed nuclear elements  
(SINES), 60–61
- Short wavelength–sensitive type 1  
(SWS1) visual pigments, 78–82
- Signaling, estrogen, 105–109
- Site selection, ambiguity management  
using, 33–34
- Sleeping Beauty (SB) transposon, 62–65  
compared to Frog Prince transposons,  
63–64
- Small subunit rRNA (SSU rRNA)  
relationships, analysis of, 176
- Speciation mechanism, in the big bang  
model, 165
- Species tree, structural domain natural  
history built on, 184
- Specificity-diverse superfamily, 197
- SRHEPT heptavariant, 268
- S*-score, 154
- Stability studies, 43
- Steroid receptors  
analyses of, 107  
ancestral, 105–109  
in metazoans, 106–107

- Stochastic models, of protein evolution, 167
- Structural biology, 9–10
- Structural classification databases, 137
- Structural distances, automated method for defining, 138
- Structural evolution “free-energy landscape” picture of, 149  
influence of designability on, 170–173
- Structural evolution models  
big bang, 165  
convergent, 161–165  
failings of, 166  
sequence-based, 157–158
- Structural features, progression of, 174–175
- Structural genomics, 165–166  
projects, 183–184
- Structural phylogeny (phylogenies), 175–184  
building, 178–183  
unrealized promise of, 183–184
- Structural proteomes, 161–165  
importance of, 162
- Structural universe, random subgraphs of, 163–164
- Structure–Function Linkage Database (SFLD), 197, 231
- Substrate binding, as an evolutionary constraint, 196
- Substrate-constrained evolution, in suprafamilies, 204–209
- Substrate-constrained suprafamily, 197–198
- Substrate specificity, changing, 264–268
- o*-Succinylbenzoate synthase (OSBS), 245–247  
activity of, 220–221, 263–264
- o*-Succinylbenzoate synthase enzymes, 203–204
- o*-Succinylbenzoate synthase reaction, catalysis of, 225
- Superfamilies. *See also* Enzyme superfamilies  
analysis of, 198  
defined, 197
- Suprafamilies active-site architecture-constrained evolution in, 209–214  
defined, 197–198  
exploiting in protein design, 214–215  
substrate-constrained evolution in, 204–209
- Suprafamily relationships, research into, 214
- SWS1 pigments. *See* Short wavelength-sensitive type 1 (SWS1) visual pigments
- Systems biology, historical perspective on, 7
- Template enzymes, screening, 195
- Templates, promiscuous, 230
- Template selection, for protein engineering, 203–204
- Tetrachlorohydroquinone dehalogenase (TCHQD), 221–222
- Thermophilic organisms, 85–86
- Thermophilic proteins, designability of, 170
- Thermostability  
assays, 92–93  
profiles, 90  
similarities in, 89–90
- Thioredoxin, esterase activity in, 277–278
- Thioredoxin suprafamily, 211, 212, 214
- 34E4 antibody, 270–271
- 3 $\alpha$ -hydroxysteroid dehydrogenase (3 $\alpha$ -HSD), 245
- Three-dimensional lattice simulation, 170
- L-Threonine aldolase (TA), comparison with alanine racemase, 251

- Transitional form conundrum, 69
- Transition redundant exchange (TREx) clock, 100
- Transition-state mimicry, combining reactive immunization with, 273–274
- Transposable elements, ancestral, 60–61. *See also* Transposons
- Transposons, 60  
biomedical applications of, 64–65
- Tree topologies, differing, 80
- Trifurcated evolutionary tree, 23–25
- Triose phosphate isomerase (TIM), conversion of ribose-binding protein into, 278–281
- TrpF, 264–265. *See also* HisA/TrpF enzyme
- Tryptophan biosynthesis, 205–209
- Tryptophan biosynthesis enzymes, 206–208. *See also* TrpF
- 20× 20 matrix, 13–14
- Two-dimensional lattice polymers, sequence evolution of, 151
- Tyr265Ala mutation, 250–253
- Tyrosine aminotransferase (TyrAT), 267
- Ubiquitin, redesign of, 284
- Ultraviolet sensitivity studies, 81
- Ultraviolet vision, 78, 80  
planetary biology questions related to, 81–82
- Unconstrained parsimony, 181
- Vertebrate opsins, 74
- Vertebrate-type steroids, 106
- Virtual libraries, computational screening of, 275–281
- Virtual sequence libraries, automated screening of, 288
- Vision, evolution of, 75
- Visual pigments, 74–85  
evolution of, 78  
short wavelength-sensitive type 1, 78–82
- Yeasts  
domestication of, 100  
ethanol accumulation by, 94
- Zebra finch, ultraviolet sensitivity in, 81
- Zebrafish  
blue opsins in, 83–84  
visual pigments in, 82–83