

Index

a

α helix, see alpha helix
 ab-initio prediction methods 70
 acquired immunodeficiency syndrome,
 see AIDS
 activation function, see transfer function
 active site 3, 12–13, 27, 47–48, 112, 169, 172
 AIDS 171, 174
 alignment 37, 41, 46–47, 54, 75–78, 82, 86–
 87, 97–99, 108, 123, 125, 182
 – global 79–80
 – local 79, 94
 – multiple 89, 91–96, 98, 110–112, 117,
 121, 124, 131, 138, 151, 153, 156, 177–
 180
 – score 80–81, 83–85, 92, 122
 alpha helix 5–6, 8, 11–13, 16, 22, 43–44, 46,
 64–65, 100, 131, 137, 140, 151, 160, 164, 166,
 169
 Anfinsen 24, 26
 antibody, see immunoglobulin
 antigen 3, 103
 antigen binding loops, see antigen binding site
 antigen binding site 102–104
 architecture 11, 22–23, 35, 81, 118
 ASTRAL 21
 atomic solvation 121–122
 automatic learning methods 137, 142, 153

b

β barrel, see beta barrel
 β bulge, see beta bulge
 β sheet, see beta sheet
 β strand, see beta strand
 back-propagation 144
 background distribution 64, 83–86, 120
 Bayes theorem 131
 Bayesian logics 131, 134
 benchmarking 50–51

beta barrel 10, 161, 166–167
 beta bulge 44
 beta sheet 6–8, 11–13, 43–44, 101, 169
 beta strand 6–8, 22, 44, 46, 99, 101, 119, 131,
 137, 140–141, 151, 160, 166
 BLAST 75, 83–84, 86–87, 89–90, 96–97
 – masking 85
 BLOSUM 81–83, 154
 Boltzmann distribution 63, 68
 Boltzmann's law 63
 Born-Oppenheimer approximation 56

c

canonical structure 103–104
 CASP 52–54, 62, 70–71, 107, 109–110, 112–
 113, 121, 125, 129, 186–187
 catalytic site, see active site
 CATH 23
 Chou and Fasman algorithm, see Chou and
 Fasman method
 Chou and Fasman method 140, 142
 circular dichroism spectroscopy 14
 CLUSTAL 75, 93–94, 97
 CNG 181–183
 comparative modelling 53, 73 ff, 80, 84, 98–
 99, 108–109, 112–113, 117, 125–126, 170
 core 17, 19, 36, 38–39, 47, 75–76, 81, 97–99,
 107, 113, 119
 correlated mutations 154, 156
 Coulomb's law 58–60
 covalent interactions 56
 cumulative matrix 78, 80
 cyclic nucleotide-gated channels, see CNG
 cytokine 175

d

Darwin 29, 31
 database search 75, 83–84, 90, 97, 151
 – E-value 86–89, 91, 124

- p-value 86–87, 91
- score 86–87, 91, 96
- Dayhoff, Margaret 81
- dead end elimination 107
- deletions 31, 35–37, 76–81, 94, 96–98, 108, 120, 122, 151
- dielectric constant 58, 60–61
- dihedral angle 1, 4–5, 7, 49, 57–58, 99, 106, 108, 135
- dipole 60
- dipole moment 61
- distance geometry 138
- distance map 22–23
- distance matrix, *see* distance map
- disulfide bond 12, 24–26
- disulfide bridge 182
 - *see also* disulfide bond
- domain 10–11, 23, 27, 36, 42, 53, 90–91, 98, 102, 112, 169, 173, 182
- domain boundary 90
- double bond 1–2, 56, 159
- DSSP 44, 151

e

- eigenpair 139
- eigenvalue 138–139
- eigenvector 138–139
- electron density 15–17, 20–21, 56, 102
- electrostatic interactions 58–59
- energy landscape 35, 65
- energy minimization 65–66, 110
- enthalpy 24, 27
- entropy 12, 24, 26–27
- error space 146
- error surface 145
- EVA 51, 153
- exon 36
- extreme value distribution 86

f

- false negatives, *see* FN
- false positives, *see* FP
- FASTA 75, 83–86, 89, 97
- fibrinopeptides 32
- fitness function 68–69, 107–108, 119
- FN 87, 89
- fold 10, 22, 27, 36, 73, 123–124, 126–127
- fold recognition 53, 74, 108, 127, 131, 167, 175, 177
 - analogous 73, 117ff, 121, 125
 - homologous 73, 121, 125
- Folding@home 70

- FP 88–89, 91, 96
- Fragfold 130–131, 135
- fragment-based methods 74, 127ff
- free energy 24, 26, 34–35, 55, 62
- free energy of transfer 162–163
- frozen approximation 122–123
- FSSP 23

g

- G-protein coupled receptors, *see* GPCR
- gap penalty 83, 91, 94, 98, 120
- Garnier-Osguthorpe Robson method, *see* GOR method
- Gaussian distribution 85
- gay related immunodeficiency disease, *see* AIDS
- GB viruses 176–177
- GBV-A, *see* GB viruses
- GBV-B, *see* GB viruses
- GBV-C, *see* GB viruses
- GDT-TS 48–49, 54, 109, 186
- genetic algorithms 67–69
- genetic code 29–30
- glycerol 160
- glycoprotein 172, 176, 178
- GOR method 142
- GPCR 185

h

- HCV 176–181, 183–184
- helical wheel 165
- hemoglobin 32–34, 117
- hepatitis C virus, *see* HCV
- HGV 177
- hidden Markov models 89, 94–97, 156, 166–167
- High-throughput screening 183
- HIV 171–172, 174
- HMM, *see* hidden Markov models
- homologous proteins 27, 35, 39, 50, 84, 87, 97, 107–108, 117, 124, 147, 151
- homology 23, 34
- homology modelling, *see* comparative modelling
- Hooke's law 56
- hydrogen bond 5ff, 12–13, 17, 24, 43–44, 59, 66, 99–101, 103–104, 134, 160–161, 163
- hydrophobic moment 165–166
- hydrophobicity 28
 - hydrophobicity plot 164–166
 - hydrophobicity scale 163

i

immunoglobulin 3–4, 99, 101–106, 117
 insertions 31–32, 35–37, 76–81, 94, 96–98,
 108, 120, 122, 151
 intron 36

j

Jack knife 50

k

kernel function 148–149
 knowledge based potentials, see statistical
 mechanics potentials

l

lactalbumin 171
 lead compound 183
 leptin 174–176
 leptin receptor 175–176
 Levinthal paradox 26–27
 Livebench 51
 long-range contacts 153, 156
 loop 8, 12–13, 44, 46, 48, 98–107, 110, 141,
 167
 lysozyme 169–170

m

meiosis 31–32
 membrane 23, 117, 159–163
 membrane proteins 159–161, 164, 166–167,
 172, 185
 Mendel 29
 mitosis 31
 model accuracy 41, 47, 76–77, 109, 126, 130,
 180–181, 186
 model optimization 107, 124
 Modeller 108
 module 11
 molecular dynamics 66–67, 70, 110, 167
 molecular replacement (in X-ray crystallo-
 graphy) 15, 186–187
 Monte Carlo 67–68, 122, 156
 moonlight proteins 28
 motif 10, 27, 128
 mutation 31–32, 34–36, 68, 76, 79, 81, 151,
 154, 176, 180
 – missense 31
 – nonsense 31
 myoglobin 33–34, 169–170

n

natively unfolded proteins 28
 Needleman and Wunsch algorithm 78–80

neural network 142–144, 147–148, 150–151,
 153, 162, 167
 neuron 143–144, 147, 152
 new fold prediction 73
 Newton equation 66
 Newton's law 55
 NMR 14, 17–20, 22, 52, 138, 184
 node (in neural networks), see neuron
 nuclear magnetic resonance, see NMR
 nucleation site 141

o

orthology 34, 84

p

pair interaction potential, see statistical
 mechanics potential
 pair potential, see statistical mechanics
 potential
 PAM 81–83, 154–155
 Paracelsus challenge 36
 paralogy 34, 84
 partition coefficient 162
 Pauli exclusion principle 61
 PDB 10–11, 20–22, 25, 35, 42, 51, 74–75,
 97
 pepsin 2–3, 172
 peptide bond 1–4, 13, 63, 104
 permittivity 58
 PFAM 97
 pharmacophore 183
 PHD 151–153
 ϕ angle 6–8, 12–13, 43–44, 69
 phi angle, see ϕ angle
 phospholipids 159–160
 phylogeny 32
 potential energy 62, 65–67, 135
 prediction server 51, 110, 135
 primary structure 4, 100
 prior probability 131, 133–134
 profile 89, 94, 96, 121, 124, 132, 151, 153
 profile-based methods 119–120
 profile–profile methods 124
 protease 3, 12–14, 27–28, 170–174, 176, 178–
 180
 Protein Data Bank, see PDB
 protein design 135
 protein kinase 13, 184
 pseudo-counts 94
 ψ angle 6–8, 12–13, 43–44, 69
 psi angle, see ψ angle
 PSI-BLAST 75, 83, 96–97, 124, 151, 153
 PSIPRED 153

q

- Q index (in secondary structure prediction) 44
- Q3 index (in secondary structure prediction) 44
- quaternary structure 4, 27, 178

r

- radius of gyration 134
- Ramachandran plot 7–8
- random coil 24, 28
- random distribution, see background distribution
- Receiving Operator Curve, see ROC
- replacement (of amino acids), see mutation
- retrovirus 171
- rhodopsin 164, 185
- rmsd 19, 37–39, 47–48, 170
- RNase A 25–26
- ROC 89
- root mean square deviation, see rmsd
- Rosetta 130, 132, 135–136
- Rossmann fold 9–10, 117
- rotamer libraries 106–107

s

- sampling 145
- Schrödinger equation 55–56, 58
- SCOP 23
- SCWRL 107
- SDR, see structurally divergent regions
- secondary structure 4, 7, 11, 17, 20, 22–23, 36, 38, 42, 46, 51, 59, 64, 81, 98–99, 102, 117–119, 135, 137, 141, 160, 166–167
 - prediction 41, 43–44, 120, 124, 138, 140, 144, 147, 150–153, 161, 176–177
- Segment Overlap measure, see SOV
- Semliki Forest virus, see SFV
- sensitivity 87–89
- sequence alignment, see alignment
- SFV 178–179
- side chain 6–7, 12–14, 17, 25, 27, 49, 75–77, 99–100, 103, 106–108, 110, 113, 121, 154, 186
- similarity matrices, see substitution matrices
- simulated annealing 68, 108, 135
- single bond 1, 56
- sliding window 151–152, 163–164
- Smith and Waterman algorithm 80
- solvation energy 121
- solvation free energy 121
- solvation potential 121
- solvent accessibility 51, 119–121, 134

SOV 44–46

sparse coding 150

specificity 87–89

spin 18

statistical mechanics potentials 62–65, 107–108, 121–122, 131, 135

stems 101–102

STRIDE 44, 151

structural genomics 27, 29

structural superposition 23, 37, 46–47, 51, 96, 98

structurally divergent regions 75, 98–99, 135

substitution, see mutation

substitution matrices 81–83, 86, 98

sum of pairs score 93

superfolds 117, 119

superposition, see structural superposition

supersecondary structure 9–10, 127, 131

supervised learning, see supervised training

supervised training 144

support vector machines 148–150, 162

SVM, see support vector machines

SwissModel 98

t

T-COFFEE 75, 93–94, 97

target function, see fitness function

Taylor power series 57

TBEV 178–179

template 77, 84, 91, 96–99, 101, 108–112, 121–124, 126–127, 131, 170, 187

tertiary structure 4, 8, 26, 46–47, 49

test set 50, 147–148, 151

threading 119, 121–123, 175

Tick Borne encephalitis virus, see TBEV

TIM barrel 11, 117

time step (in molecular dynamics) 66–67

TMHMM 166

TN 88–89

topology 11, 23, 27, 36, 118, 167

TP 87, 89

training set 50–51, 142, 146–147, 151

transfer function 143–144, 147

transmembrane proteins, see membrane proteins

tree 32, 34, 94

triangle inequalities 139

true negatives, see TN

true positives, see TP

trypsin 2–3, 21

turn, see loop

u

unfolded conformation, see random coil
unfolded protein, see random coil
unsupervised learning, see unsupervised training
unsupervised training 144

v

validation set 146–148, 151
van der Waals potential 24
van der Waals radius 61
virtual screening 183–186

w

water 24–25, 59, 121

x

X-ray crystallography 14–17, 19, 52, 61, 101, 110, 169–170, 181, 184, 186
– B factor 17–18, 21–22, 49
– Debye-Waller factor, see B factor
– occupancy 17, 21
– R factor 17, 20, 22
– resolution 17, 22
– Rfree factor 17, 20
– temperature factor, see B factor

z

Z-score 123

