

Index

- Abelian group, 402–406, 409
- Ab initio methods, 219
- Accessible surface area (ASA), 345–346
- Acid-denatured ACBP, 58
- Active site modification, 97–98; chs. 10, 12, 28, 46
- Acyl carrier proteins (ACPs), 249–251
- Adenosine monophosphate (AMP), 105, 253–256
- α -expansion move algorithm, 318–320, 325–326
- All-atom Cartesian coordinates-based model, 191
- All-class mean, 188
- Alpha shapes, 410–411
- Alpha-synuclein, 334
- AMBER, 222
- Aminoacyl-AMP formation, 254–255
- Amprenavir, 246–247
- Anisotropic network model (ANM), 214–215
- ApoPheATE, 254
- Approximation algorithms, ch. 12
- Aryl carrier proteins (ArCPs), 249–251
- A^* search, 106–107, 234, 237, 241–242
- Assignment problems, 3, 23; chs. 8, 14, 15–18, 33, 34, 37, 38, 43
 - Autolink and, 363
 - energy function and, 315–317
 - graph cuts and, 315–321, 323
 - residual dipolar couplings (RDCs) and, 27, 116, 129, 131, 151
- ATP, 254–255
- AURELIA, 293
- Autolink
 - assignment problems and, 363
 - fragment and, 363
 - overview of, 363–365
 - pair scoring and, 365–367
 - priority list and, 363, 367, 369–370
 - random factor bias and, 369
 - repeat bias and, 369
 - spin system and, 363, 365–367, 370
- AutoStructure
 - discriminating power (DP) score and, 289–291
 - distance geometry and, 285–292
 - final structure assessment and, 289–291
 - F-measure score and, 289–290
 - heuristic distance network initialization and, 285–287
 - initial ambiguous network construction and, 285
 - initial structure construction and, 288
 - input data preparation and, 285
 - input data validation and, 285–287
 - iterative fold analysis and, 288–289
 - precision score and, 289–290
 - recall score and, 289–290
 - resonance assignment and, 287
- Backrubs, 108, 111, 241, 412
- Badoiu, M., 310
- Bailey-Kellogg, C., 61
- Baker, D., 155, 167
- Bayesian networks (BN), 228–229
- Bayesian probability, 119, 161–162, 172, 228–229
- BD algorithm, 104, 107–108, 111, 241
- Belief propagation algorithms
 - Bethe approximations and, 231
 - dead-end elimination (DEE) and, 234
 - defined, 229
 - ensemble properties and, 233
 - free energy and, 230–232
 - future directions for, 240–242
 - generalized belief propagation (GBP) and, 231–233
 - graphical models and, 229–233
 - max-marginal probability and, 236–238
 - max-product, 236–237
 - mean-field approximations and, 231
 - multiple low-energy sequences through, 237–238
 - protein design and, 236–237
 - region graphs and, 231
- Bernardo, P., 333, 337
- Bethe approximations, 231
- Betti number, 402, 404, 407, 410
- Between-class variance, 188
- Bézier systems, 403
- B-factor, 342
- BioMagResBank (BMRB), 64, 168, 295–296

- Bioterrorism, 160
 Bipartite graphs, 23–24, 122, 166–169, 261
 Biswas, P., 309
 Boltzmann constant, 9
 Boltzmann probability, 9, 221, 235, 240, 266, 268, 372
 Boltzmann's law, 230, 235
 Boltzmann solvation model, 89
 Boykov, Y., 318
 Bragg's law, 341
 Branching factor, 100
 brDEE algorithm, ch. 12
 BroMAP, 241–242
 Brownian normal mode analysis (NMA), 199
- Canny, J., 193–194
 CAP, 151
 Carbonyl carbon, 2, 78
 computational protein design and, 94
 interatomic vectors and, 47–51
 protein-ligand matching and, 261
 residual dipolar couplings (RDCs) and, 166 (*see also*
 Residual dipolar couplings)
 Carrier proteins, 249–251
 Cartesian coordinate-based root-mean-square deviation,
 351
 CAST, 411
 Catalytic machinery, 77, 80–81; ch. 28
 CBF β , 65–67
 CH bond, defined, 55, 126
 Chain groups, 402, 404–407
 Chain Tweak, 191–195
 CHARMM, 222
 Chemical shift, 10–11, 18
 CS-Rosetta and, 371–376
 JIGSAW and, 59–60, 64
 MARS and, 293–300
 NOESY data interpretation and, 285, 288
 nuclear vector replacement (NVR) and, 119–122
 protein-ligand NOE matching and, 261
 protein unfolding and, 333
 residual dipolar couplings (RDCs) and, 129, 131–132,
 135, 145–146, 150, 152, 155, 162, 166–170
 resonance assignment and, 355, 365
 rotamers and, 371, 373
 simulated annealing (SA) and, 371–373
 Chen, C.-Y., 99, 105
 Chen, L., 111
 Chernoff tail bounds, 155
 Circular dichroism spectroscopy (CD), 397
 Claus, B.I., 259–260, 262
 Clustering modulo a group, chs. 14, 17
 Cohomology operation, 411–412
 Collins, G.E., 409
 Computational protein design, chs. 11, 12, 30, 45, 46, 49
 complexity analysis and, 92 (*see also* Dead-end
 elimination)
 Dahiyat-Mayo protocol and, 87–92
 dead-end elimination (DEE) and, 88–92
 enumeration and, 89, 91
 experimental validation and, 92, 94
 global minimum energy conformation (GMEC) and,
 88–92
 inverse folding problem and, 87
 JIGSAW algorithm and, 94
 nuclear magnetic resonance (NMR) and, 92, 94
 nuclear Overhauser effect (NOE) and, 94
 protein backbone and, 87–92
 pruning and, 88–89, 92
 residual dipolar couplings (RDCs) and, 94
 rotamers and, 87–92
 self-energies and, 89
 template energies and, 89
 zinc finger and, 88, 91–92
 Computer-aided resonance assignment (CARA), 363
 Conformation sampling, 206–207, 388
 Conformation trees, 100, 106–107, 141, 162
 Continuum solvent model, 344
 Covariance matrix
 principal components analysis (PCA) and, 27–44
 residual dipolar couplings (RDCs) and, 33–44
 Critical values, 410
 Cross-rotation analysis algorithm (CRANS), 160,
 383–386
 Cross-rotation function (CRF), 383–384, 391
Cryptosporidium hominis, 160
 CS-Rosetta
 all-atom minimization and, 373
 correctness and, 372–373
 DC and, 374
 DP score and, 375
 fragment library generation and, 373
 fragment replacement and, 373
 Monte Carlo procedure and, 372
 nuclear magnetic resonance (NMR) and, 371–376
 performance analysis of, 374–376
 prediction and, 371
 Rosetta and, 371–375
 side-chain minimization and, 373
 SPARTA and, 373–374
 termination and, 374
 testing set and, 371
 CYANA, 371
 Cyclic coordinate descent (CCD) algorithm, 220–222
- DACS (divide-and-conquer splitting), ch. 12
 Dahiyat, D.I., 87
 Dead-end elimination (DEE), 84, 175, 265
 A* search and, 106–107, 234, 237–238, 241–242
 backrub and, 108, 111, 241, 412
 BD algorithm and, 104, 107–108, 111, 241
 belief propagation algorithms and, 234
 brDEE algorithm and, ch. 12
 computational protein design and, 88–92
 DACS and, ch. 12

- ensemble-based protein design and, 99–111
- flexible, 104, 107–108, 111, 241
- minimized (minDEE), 99–111, 241
- negative design and, 109–110
- peptide design and, 69–70
- protein-ligand binding and, 107
- side-chain conformations and, 106, 108
- steric filter and, 100
- Degrees of freedom (DOFs), chs. 7, 14, 15–18
 - equilibrium fluctuations and, 220, 222
 - protein flexibility and, 191–194, 206
 - rigidity theory and, 201
- Delarue, M., 388
- Delaunay triangulation, 399
- Denatured protein structure, 55–58; ch. 39
- Denavit-Hartenberg (DH) local frames, 192
- Design of proteins, peptides, small molecules, chs. 9–12, 25–28, 30, 41, 45, 46, 49
- Differential scanning calorimetry (DSC), 397
- Dihedral angles-based model, 191
- Dihydrofolate reductase (DHFR), 109
- Dimpleoid, 30
- Directed evolution, 83–85
- Directional variance, 28–33, 39–44
- Distance geometry, 2, 4–5, 92, 130, 157, 344
 - all distances but with some errors and, 3
 - all exact distances and, 3
 - angle information and, 310–311
 - AutoStructure and, 285–292
 - divide-and-conquer approach and, 274
 - graph connectivity and, 274–275
 - graph embedding and, 310–311
 - graph partitioning and, 275–276
 - interval bounds and, 3
 - molecule problem and, 273
 - NP-hard problem, 279–283, 308
 - orientation constraints and, 307, 310–313
 - residual dipolar couplings (RDCs) and, 311–313
 - rigidity theory and, 200, 274–277
 - semidefinite programming and, 307–313
 - sparse sets of distances and, 3
 - stress matrix and, 274–275
 - subgraph realization and, 276
 - time complexity and, 312
 - unique realizability conditions and, 274–275
- Distance root-mean-square deviation, 351
- Domain-swapping, 97–98
- Donald, Bruce R., 98–100, 108, 111
 - flexible backbone DEE and, 107
 - nuclear vector replacement (NVR) and, 119, 122, 124
 - residual dipolar couplings (RDCs) and, 55, 58, 127–185
- Dunathan's hypothesis, 83
- DYANA, 371
- Dynein light chain (DLC2A), 301, 303–305
- E-approximation algorithms, ch. 12
- E. coli*, 84, 94
- Edelsbrunner, Herbert, 410
- Edge-weight, 122, 283
- EEF1 solvation model, 99, 102–105
- Elastic Network Models (ENM), 389–390
- Energy functions
 - class \mathcal{F}^2 and, 332
 - class \mathcal{F}^3 and, 330–332
 - classification of, 326–327
 - ensemble-based design and, 102–104
 - expectation/maximization (EM) algorithm and, 324–326
 - feature space clustering and, 323–326
 - graph cuts and, 315–332
 - nonparametric clustering and, 325
 - parametric clustering method and, 325
 - pixel labeling problem and, 323–326
 - receptor docking and, 344
 - soft-core, 344
- Energy-weighted average, 343
- Ensembles
 - analyzing protein structures using, 351–353
 - conformational ensemble and, 337–340
 - dead-end elimination (DEE) and, 106–111
 - docking and, 343
 - K^* algorithm and, 98–106, 110–111
 - local sampling and, 333–336
 - long-range order and, 335
 - loops and, ch. 24
 - metric space and, 352–353
 - minimized, 106–111
 - nonribosomal peptide synthetase (NRPS) enzymes and, 97–98, 108, 110
 - principal components analysis (PCA) and, 27
 - residual dipolar couplings (RDCs) and, 20, 27–28, 34, 222, 226, 333–337, 343, 352
 - rotamers and, 98–108, 111
 - small-angle X-ray scattering and, 337–340
 - steric clash checking and, 333–335
- Entropy
 - amprenavir and, 246–247
 - using belief propagation, ch. 25
 - configurational, 245–248
 - conformational, 104, 232–233, 245–248
 - design implications and, 247–248
 - free energy and, 232
 - K^* and, chs. 12, 25
 - ligand binding and, 245–248
 - multiplicity and, 245
 - vibrational, 246
- Enzyme redesign
 - catalytic machinery and, 77, 80–81
 - computational, ch. 12
 - directed evolution and, 83–85
 - Dunathan's hypothesis and, 83
 - improving promiscuity and, 77, 82–84
 - interconverting homologous enzymes and, 79–80
 - minimalist active site and, 77–83
 - partitioning of reaction intermediates and, 82

- Enzyme redesign (cont.)
 product and, 77
 removal of catalytic nucleophiles and, 81
 stereo- and regiochemistry control and, 82
 substrates and, 77
 subtilisin and, 78
- Equilibrium fluctuations, 220–221, 226
- Erdmann, Michael, 412
- Euclidean geometry
 cross-rotation and, 384
 distance geometry and, 309
 ensembles and, 351
 protein flexibility and, 208
 residual dipolar couplings (RDCs) and, 134
 topology and, 400, 405–409
- Euler angle, 120–121
- Euler number, 404
- Expectation/maximization (EM) algorithm
 graph cuts and, 320–321, 324–326
 NVR and, chs. 14, 15, 17.2, 18.1.2
 protein flexibility and, 208
 residual dipolar couplings (RDCs) and, 118–119, 122, 150–152, 166–171
 resonance assignment and, 361
 X-ray crystallography and, 388
- Face, 410
- Factor graphs, 229
- Feature space, 323–326
- Feature vectors, 323–324, 326, 393
- FELIX, 293
- Finite simplicial complex, 400–401
- FIRST, 197, 200
- Flexible-backbone DEE (BD), 107–108. *See also* brDEE
- Flexible-backbone GMEC, 107
- Foldamers, 71–74
- Fourier transform, 10, 166, 341, 383, 387
- Fragment ensemble method (FEM), 220–221
- Framework rigidity optimized dynamic algorithm (FRODA), 208–210
- Free chain complex, 405
- Free energy, 247
 belief propagation algorithms and, 230–232, 236–237
 entropy and, 232
 Gaussian summation and, 346
 Gibbs, 227–228, 230–231, 245, 394
 global minimum energy conformation (GMEC) and, 69, 88–92, 104–110, 233, 235, 237, 240–242, 332, 412
 Hessian matrix of, 389
 holoPheATE catalysis and, 256–257
 K* and, ch. 12
 metadynamics and, 346–349
 predicting binding energetics and, 345–346, 349
 protein-ligand interactions and, 67
 pseudoenergy and, 295
- Free energy surface (FES), 346, 349
- Free induction decay (FID), 9–10
- Frobenius inner product, 36–38
- Fromer, M., 235, 237–238
- Fundamental group, 400–401
- Fushman, D., 51
- GAPDH, 81
- Gaussian distribution, 30–33, 61, 168, 346
- Gaussian Network Model (GNM), 167, 213–214
- Gaussian summation, 346
- GD algorithm, 131, 133, 153, 155, 161, 166, 169
- Generalized belief propagation (GBP), 231–233
- Generalized Born/Poisson-Boltzmann solvation models, 89
- Generalized Potts Model (GPM), 316
- Genetic algorithms (GA), 135, 219, 265–268, 393–395
- Geometry-weighted average, 343
- Georgiev, I., 98–100, 108, 111
- Gibbs free energy, 227–228, 230–231, 245, 394
- Gilson, M.K., 246–247
- Global minimum energy conformation (GMEC), 412;
 chs. 11, 12
 belief propagation and, 233–237, 240–242
 computational protein design and, 88–92
 ensemble-based protein design and, 104–110
 graph cuts and, 332
 peptide design and, 69
- Global orientational information from RDCs, 19
- Gorzynski, M.J., 67
- Gramicidin S, 253–258
- Graph cuts
 α -expansion move algorithm and, 318–320, 325–326
 classifying power of, 323–332
 computing visual correspondence and, 318–321
 energy function and, 315–332
 expectation/maximization (EM) algorithm and, 324–326
 feature space clustering and, 323–326
 Generalized Potts Model (GPM) and, 316
 graph construction and, 316–317
 grid topology and, 315
 Hammersley-Clifford theorem and, 315
 hard membership and, 326
 MAP estimation and, 315
 Markov Random Fields (MRFs) and, 315–316
 multiway cut algorithm and, 317
 nonparametric clustering and, 325
 NP-hardness and, 317, 329
 occlusions and, 318–321
 parametric clustering method and, 325
 pixel labeling problem and, 323–326
 residual dipolar couplings (RDCs), 333–334
 soft membership clustering and, 326
 spatial coherence and, 323–324, 326
 super pixel approach and, 326
- Graphical models
 application in protein design and, 233–242
 Bayesian networks and, 228–229
 belief propagation algorithms and, 229–233, 236–237

- factor graphs and, 229
- future directions for, 240–242
- marginal probability and, 228
- Markov Random Field and, 229–232, 315–316, 318, 325–326
- max-marginal probability and, 236–238
- probabilistic protein design and, 238–240
- protein design problem and, 235–237
- region graphs and, 231
- Graph partitioning, 275–276
- Graph theory
 - AutoStructure and, 285–292
 - bipartite graphs, chs. 14, 15–18, 43
 - distance geometry and, 273, 279
 - Froda, ch. 22
 - graph embedding, chs. 31–33, 36
 - graph cuts, chs. 37, 38
 - JIGSAW and, 59, ch. 8
 - Laman's theorem, 201
 - lower bounds and, ch. 32
 - NOE network, ch. 33
 - protein flexibility and, 200–202
 - ROCK, ch. 22
- GrsA-PheA, 97–98, 105, 108
- Hammersley-Clifford theorem, 315
- HANA
 - Chernoff tail bounds and, 155
 - Hausdorff distance and, 155
 - residual dipolar couplings (RDCs) and, 133, 155–158, 171–172
- Harmonic approximation, 197–198
- Hausdorff distance, 117–118, 155
- Hellinga, H., 104
- Hendrickson, B., 273–277
- Hessian matrix, 199, 214, 389
- Heteronuclear single-quantum coherence (HSQC), 13, 23
 - errors in structure determination and, 303
 - HX rates and, 355, 357
 - JIGSAW and, 59–60, 65
 - NMR structural biology and, 17–20
 - nuclear vector replacement (NVR) and, 119–121
 - protein-ligand NOE matching and, 259–260
 - reductive methylation and, 357–361
 - residual dipolar couplings (RDCs) and, 130, 150–152, 166, 168
 - resonance assignment and, 355–357
- Heuristic algorithms, chs. 6, 9–11, 20–26, 30, 31, 41, 42, 44, 45, 49
- HIV, 246–247, 411
- HNH^α experiment
 - JIGSAW and, 59–60, 65
 - residual dipolar couplings (RDCs) and, 135, 145
 - scalar (spin-spin) coupling and, 4–5, 15–16, 20, 59–60
- HoloPheATE, 254, 256
- Homologous enzymes, 79–80
- Homology groups
 - Abelian groups and, 402–406, 409
 - alpha shapes and, 410–411
 - Betti number and, 402, 404, 407, 410
 - chain groups and, 402, 404–407
 - cohomology operation and, 411–412
 - complexity and, 403
 - computable types and, 402–404
 - computing, 404–410
 - critical values and, 410
 - face term and, 410
 - free chain complex and, 405
 - kernels and, 402, 405
 - mathematical/biological terms for, 399
 - matrices and, 401–409
 - reduction algorithm and, 406–409
 - simplicial complexes and, 400–404
 - Smith normal form and, 402–403, 409
 - topology and, 399–412
 - torsion coefficients and, 402, 407
- Homology modeling
 - arginine side-chain and, 306
 - Bézier systems and, 403
 - catalytic machinery and, 77
 - cross-rotation function and, 384
 - domain-swapping and, 84
 - dynein light chain and, 301–302
 - equilibrium fluctuations and, 219–220
 - free energy and, 227
 - improving promiscuity and, 82–83
 - interconverting enzymes and, 79–80
 - MARS and, 298
 - mathematical/biological terms for, 399
 - molecular replacement (MR) and, 387, 391
 - noncrystallographic symmetry (NCS) and, 383, 391
 - nuclear vector replacement (NVR) and, 119, 123–124
 - peptide design and, 68
 - protein flexibility and, 195
 - protein-ligand NOE matching and, 259
 - residual dipolar couplings (RDCs) and, 117–118, 130–133, 149–155, 161, 167, 169–170
 - Rosetta and, 372
- Homonuclear correlation spectroscopy (COSY), 15
- Homo-oligomers, 158–159
- Hopcroft-Karp algorithm, 24
- Hungarian algorithm, 259–263, 361
- Hus, J.-C., 151
- HX rates, 355, 357
- Hydrophobic interactions, 207–208
- Iliopoulos, C.S., 409
- In-class variance, 188
- Inference problems, 235
- Initiation module PheATE (GrsA), 253–258
- Inner product, 29
- Interatomic vectors, 47–51
- Intramutation pruning, 100–102

- Intrinsically disordered proteins, chs. 7, 39, 42
 Intrinsic dimensionality, 188
 Inverse folding problem, 87, 234
 Inverse kinematics, 117, 164, 175, 191–195, 219, 312
 Ion conductance studies and, 158
 Isotopic labeling, 132; chs. 2, 3, 8, 14, 15–18, 29, 43
 Isupov, M.N., 391
- J coupling, 13, 15, 20, 59, 65
 J-doubling, 135, 145
 JIGSAW
 automated NMR peak assignment and, 59–66
 chemical shift and, 59–60, 64
 computational protein design and, 94
 graph representation of atom interactions and, 60–61
 HNH α and, 59–60, 65
 HSQC and, 59–60, 65
 local distance patterns and, 61
 match score and, 61, 64–65
 NOESY and, 60–64
 nuclear Overhauser effect (NOE) and, 65
 overview of, 59
 pattern discovery and, 61
 protein backbone and, 59
 secondary structures and, 59, 61–64
 side-chain fingerprint alignment and, 64–66
 TOCSY and, 60, 66
 Jung, Y.S., 293
- K** algorithm
 basics of, 98–102
 branching factor and, 100
 EEF1 solvation model and, 99, 102–105
 energy functions and, 102–104
 ensemble-based protein design and, 98–106, 110–111
 graphical models and, 234
 intramutation pruning and, 100–102
 mutation space filters and, 99
 negative design and, 109–110
 peptides and, 67
 protein backbone and, 99, 102, 104–105
 protein-ligand binding and, 98–99
 redesigning enzymes with, 105–106
 side-chain conformations and, 99, 102, 104
 single-structure scoring and, 99
 stop-threshold and, 100
 synthetic biology problems and, 110
 Kamisetty, H., 228, 231–232, 240
 Karplus curve, 4, 15, 102
 Kay, L.E., 145
 Kearsley, S.K., 74
 Keller, R., 367
 Kernel methods, 381, 402, 405
 Kinematic loop closure problem (KLCP), 162
 Kinematics, 1, 3n1, 54, 220; chs. 1, 2, 9, 13, 15–18, 20–24, 30
 Chain Tweak and, 191–195
 Denavit-Hartenberg (DH) local frame and, 192
 ensembles and, 108
 initiation module PheATE (GrsA) and, 253–258
 inverse, 117, 164, 175, 191–195, 219, 312
 loop closure problem and, 191–195
 Probik and, 191–195
 protein flexibility and, 191–194
 Ramachandran plots and, 194
 residual dipolar couplings (RDCs) and, 117, 123–124, 127, 132, 138, 143–144, 160–164, 173, 175
 ROCK and, 205–208, 210
 Klein bottle, 403–404
 Knegtel, R.M., 344
 Kuhn-Munkres matching algorithm, 116
 Kullback-Leibler distance, 117–118, 121, 230
- Laman's theorem, 201
 Langmead, C.J., 122, 155, 167, 169
 Lazaridis-Karplus EEF1 pairwise solvation energy, 102
 Lebedev, A.A., 391
 Liang, J., 410
 Lie group, 123, 151, 383
 Ligand configurational entropy, 245–248
 Ligand-protein binding, 213–217
 Ligand-protein docking
 accessible surface area (ASA) and, 345–346
*K** and, ch. 12
 metadynamics and, 346–349
 Minkowski sum and, 345
 predicting binding energetics and, 345–346
 protein flexibility and, 345–349
 Lilien, R.H., 98–99, 111
 Linear discriminant analysis (LDA), 188
 Linker computation, 57
 Local distance information, 19
 Lock-and-key model, 213
 Looger, L., 104
 Loop closure problem, 191–195
 Loop libraries, 219
 L-phenylalanine (L-Phe), 105, 253–257
- Manocha, D., 193–194
 Markov Random Field (MRF), 229–232, 315–316, 318, 325–326
 MARS
 backbone assignment and, 293–300
 chemical shift and, 293–300
 methodology of, 293–296
 pseudoresidues (PRs) and, 293–295, 298
 residual dipolar couplings (RDCs) and, 296–300
 sequential connectivity and, 294–296
 Martin, Jeffrey, 127–185
 Masse, J.E., 367

- Mass spectrometry classification algorithms (MSCAs), 187–189
- Mass spectrometry (MS), chs. 19, 43
- ¹³C-methylated proteins and, 357–361
 - HX rates and, 355, 357
 - NMR assistance and, 355–361
 - proteomic disease classification algorithm and, 187–189
 - resonance assignment and, 355–361
- Matching modulo a group, ch. 14
- Mathematical equations
- anisotropic network model (ANM), 214–215
 - Autolink, 367
 - belief propagation, 230
 - Bragg's law, 341
 - chemical shift, 11
 - dead-end elimination (DEE), 69, 89, 91
 - dipolar coupling constant, 19
 - Elastic Network Model (ENM), 389–390
 - energy function, standard form, 327
 - energy function with occlusion, 318
 - ensemble representation, 351–352
 - expectation/maximization NVR algorithm, 122
 - Gaussian Network Model (GNM), 213
 - Generalized Potts Model, 316
 - generalized quality factor, 48–49
 - generalized sampling parameter, 48
 - graph cuts, 315–316, 318, 324–325, 327–330
 - Hammersley-Clifford, 315
 - K^* algorithm, 98–99
 - ligand configurational entropy, 245–246
 - max-marginal probability, 236
 - quantum energy difference, 8
 - quartic RDC, 53–54, ch. 15–18
 - nonparametric clustering, 325
 - normal mode analysis (NMA), 197–198, 213, 215
 - nuclear Overhauser effect (NOE), 18
 - orientational sampling, 48–49
 - parametric clustering, 325
 - posterior energy function, 316
 - principal components analysis (PCA), 28–44
 - protein design problem, 235–237
 - protein ensemble method (PEM), 222
 - pseudoenergy, 295
 - residual dipolar couplings (RDCs), 19, 34–44, 115–117, 137, 162–165, 173
 - resonance assignment, 121–122
 - ROCK algorithm, 206
 - rotamer probability term, 104
 - sampling tensor, 48
 - semidefinite programming, 308–309
 - sequential connectivity, 295
 - tensor estimation, 120
 - topology, 402–403, 405–409
 - Wang-Donald, 146, 163–164
 - Yan-Donald tensor notation, 137
 - Z-score, 394
- Matrices
- Hessian, 214, 389
 - Kullback-Leibler distance and, 117, 121, 230
 - positive semidefinite (PSD), 307n1
 - principal order frame (POF) and, 19, 31–32, 53–54, 57, 116–117, 120, 151–152, 162–163, 166
 - probability, 121–122
 - quaternions and, 69, 121, 383–384, 391
 - Saupe, 19, 27–28, 34–37, 43, 55, 57, 115, 137, 162–163, 334
 - semidefinite programming and, 307–310
 - special linear group SL(3) and, 121
 - special orthogonal group SO(3) and, 121, 384
 - stress, 274–275
 - topology and, 401–409
 - within-class scatter, 188
- Matthews correlation coefficient (MCC), 378
- Maximum a posteriori (MAP) estimation, 241, 315
- Maximum bipartite matching (MBM), 23–24, 116
- Max-marginal probability, 236–238
- Mayo, S.L., 87
- Mean-field approximations, 231
- Meccano algorithm, 340
- Membrane proteins
- computational topology and, 411
 - peptide design and, 70
 - residual dipolar couplings (RDCs) and, 129–133, 158, 172–173
- Metadynamics, 346–349
- Methionine residues, 104
- Metric space, 352–353
- Metropolis sampling, 216, 265, 268, 373
- MinBounds, 237n1
- minDEE (minimized dead-end elimination) algorithm, ch. 12
- Minkowski sum, 345
- Missing loops, 219
- Molecular dynamics (MD), 27, 346
- metadynamics and, 346–349
 - protein flexibility and, 191
 - simulated annealing (SA) and, 129, 132, 141, 145, 173, 279
- Molecular replacement (MR)
- cross-rotation and, 383
 - nuclear vector replacement (NVR) and, 119
 - residual dipolar couplings (RDCs) and, 127, 130–131, 150, 160, 166–167, 173
 - X-ray crystallography and, 387–391
- Moment (nuclear spin angular momentum), 7–12
- Monte Carlo methods, chs. 6, 20, 30, 34, 39, 41, 44, 45, 49
- CS-Rosetta and, 372
 - equilibrium fluctuations and, 219
 - peptide design and, 69–70
 - protein-ligand binding and, 344
 - residual dipolar couplings (RDCs) and, 129, 135, 145
 - side-chain/backbone flexibility and, 265–266, 268

- Multiple receptor conformations (MRCs), 215–217
Multiple-ring closure, 206–207
- NADPH-dependent family, 79–80
Native protein structure, 53–55
Newtonian physics, 134, 144
NH bond vectors, 50–51, 53, 124, 140, 334, 343
Nilges, M., 161–162, 343
NMRView, 293
Noncrystallographic symmetry (NCS), 160
 cross-rotation analysis algorithm and, 383–384
 Elastic Network Model (ENM) and, 389–390
 molecular replacement and, 387–391
 oligomeric models and, 390–391
 self-rotation function and, 383–384, 390–391
 X-ray crystallography and, 387–391
Nondeterministic Polynomial-time (NP) hard problem
 3SAT reduction and, 280–282
 adding dimensions and, 282
 annealing schedule and, 372–373
 computational protein design and, 92
 CS-Rosetta and, 372
 distance geometry and, 279–283, 308
 graph cuts and, 317, 329
 k-embeddability and, 282–283
 1-embeddability and, 280–282
 partition reduction and, 280
 residual dipolar couplings (RDCs) and, 145
 rotamer assignment and, 234
 subgraph realization and, 276
Nonparametric clustering method, 325
Nonribosomal peptide synthetase (NRPS), 83–85;
 chs. 10, 12, 27, 28, 46
 active site modification and, 97–98
 carrier protein and, 249–251
 computational redesign of, ch. 12
 domain-swapping and, 97–98
 ensemble-based protein design and, 97–98, 108, 110
 kinetic studies and, 253
 peptide biosynthesis and, 249–251
 phenylalanine (Phe) adenylation domain and, 97–98
 signature sequence and, 97–98
 site-directed mutagenesis and, 97–98
 support vector machine (SVM) and, 377
Normal mode analysis (NMA)
 anisotropic network model (ANM) and, 214–215
 Brownian, 199
 coarse-grained, 199
 different modes for, 199
 Elastic Network Model (ENM) and, 389–390
 essential dynamics, 199
 FIRST and, 197, 200
 Gaussian Network Model (GNM) and, 213–214
 harmonic approximation and, 197–198
 Laman's theorem and, 201
 ligand-protein binding and, 213–217
 lock-and-key model and, 213
 molecular replacement (MR) and, 387–391
 multiple receptor conformations (MRCs) and, 215–217
 noncrystallographic symmetry (NCS) and, 387–390
 Pebble Game and, 200–202
 preexisting equilibrium model and, 213
 protein flexibility and, 213–217
 protein-protein binding and, 213–217
 receptor ensemble docking (RED) and, 215–217
 relevant normal model and, 215–217
 residual dipolar couplings (RDCs) and, 167, 169
 rigidity theory and, 197–202
 side-chain optimization and, 216
 vibrational, 199
 virtual screening and, 216–217
 X-ray crystallography and, 387–390
Nuclear magnetic resonance (NMR), chs. 1–8, 13–18, 29,
 31, 33–35, 39, 43–45
 ¹³C-methylated proteins and, 357–361
 Autolink and, 363–370
 automated assignment and, 59, 119, 122–123, 129–133,
 145, 149–158, 162, 166 (*see also* Resonance
 assignment)
 basic principles of, 7–13
 chemical shift and, 10–11, 371–376
 combinatorial precision and, 146
 computational protein design and, 92, 94
 CS-Rosetta and, 371–376
 distance geometry and, 2–5, 92, 130, 157, 200, 273–283,
 307–313, 344
 dynein light chain (DLC2A) and, 301, 303–305
 electromagnetic properties and, 7–9
 ensemble averages and, 20, 27–28, 34, 222, 226,
 333–337, 343, 352
 expectation/maximization (EM) algorithm and, 118–119,
 122, 150–151, 166, 208, 324
 free induction decay (FID) and, 9–10
 frequency signals and, 2
 heteronuclear single-quantum coherence (HSQC) and,
 13, 17–20, 23, 59–60, 65, 119–121, 130, 150–152, 166,
 168, 259–260, 303, 355–357
 high-throughput, 129–130
 homonuclear correlation spectroscopy (COSY), 15
 isotope labeling scheme and, 132
 J coupling and, 13, 15, 20, 59, 65
 JIGSAW and, 59–66
 mass spectrometry and, 355–361
 methodological development of, 127–128
 moment (nuclear spin angular momentum) and, 7–12
 multidimensional, 13, 20
 NOE, 3, 15 (*see also* Nuclear Overhauser effect (NOE))
 physical basis of, 8–10
 protein backbone and, 17–20, 304–306, 371–376
 protein-ligand binding and, 341–344
 RDCs, 19–20, 34, 115–118, 127 (*see also* Residual
 dipolar couplings (RDCs))
 scalar coupling, 4–5
 spectrometer, 11–12

- structure-activity relation (SAR) by, 21, 131, 411
- structure determination errors and, 301–306
- SYMBRANE and, 132–134, 158–160, 172–177
- symmetric homo-oligomers and, 158–159, 172–176
- three-dimensional HNH^α, 4–5, 15, 20, 59
- total correlation spectroscopy (TOCSY), 17–18, 20, 59–60, 64–65, 172
- unassigned data, 129–133, 151, 161
- uncertainty of data, 2–3
- X-ray crystallography and, 7
- Nuclear Overhauser effect (NOE)
 - Autolink and, 365
 - automated assignment, 3, 155–158, 171–173
 - AutoStructure and, 285–292
 - computational protein design and, 92–94
 - CRANS NCS method, 160
 - CS-Rosetta and, 371–376
 - distance geometry and, 279–283, 312–313 (*see also* Distance geometry)
 - errors and, 303
 - filtering and, 156
 - graph cuts and, 321
 - graph representation of atom interactions and, 60–61
 - homology computation and, 400
 - interaction intensity and, 18–19
 - JIGSAW and, 59–65
 - limitations of, 25
 - MARS and, 294
 - NMR structural biology and, 15, 18–20
 - nuclear vector replacement (NVR) and, 119–121
 - pattern discovery and, 61
 - protein-ligand matching and, 259–263
 - redundant, 175
 - residual dipolar couplings (RDCs) and, 55, 115, 129–135, 142–162, 166, 169–176
 - resonance assignment and, 121
 - rotamers and, 155–158
 - secondary structure and, 61–64
 - side-chain conformations and, 133, 155–157, 171–172
 - spectrum matching and, 261–262
 - symmetric oligomers and, 172–177
 - topology and, 285–292, 400
- Nuclear vector replacement (NVR), 26
 - chemical shift and, 119–122
 - expectation/maximization algorithm, 122
 - experimental input, 119
 - fold recognition and, 166–171
 - heteronuclear single-quantum coherence (HSQC) and, 119–121
 - as matching problem, 123–124
 - nuclear Overhauser effect (NOE) and, 119–121
 - principal order frame (POF) and, 120
 - protein backbone and, 119–120
 - quaternions and, 121
 - residual dipolar couplings (RDCs) and, 119–124, 131–133, 149–153, 160–161, 166–171
 - resonance assignment and, 119–123
 - structure refinement and, 119
 - tensor estimation and, 119–121
 - three-dimensional structural homology detection and, 123
 - time complexity and, 121–123
 - unassigned NMR data and, 123–124
- Oligomeric number, 176–177
- Oligomeric state, 176–177
- Outer product, 29
- Pande, V.S., 352–353
- Paramagnetic quenching and, 168
- Paramagnetic relaxation effect (PRE) and, 55–56
- Paramagnetic relaxation enhancement (PRE), 55–56, 162
- Parametric clustering method, 325
- Parametric family, 167
- Partition function, 12, 25
- Pebble Game, 200–202
- Peptides
 - backbone and, 2, 67–73
 - biosynthesis and, 249–251
 - carrier protein structure and, 249–251
 - CBF β and, 65–67
 - dead-end elimination and, 69–70
 - design algorithm for, 69–71
 - design of, 67–74
 - foldamers and, 71–74
 - force field and, 69
 - global energy minimum conformation (GMEC) and, 69
 - initiation module PheATE (GrsA) and, 253–258
 - K^* algorithm and, 67
 - membrane proteins and, 70
 - monomer frameworks and, 72
 - nonribosomal peptide synthetase (NRPS), 83–85, 97–98, 108, 110, 249, 251, 253, 377
 - protein-protein interactions (PPIs) and, 67
 - short amino acid sequences and, 67
 - structure generation and, 69
 - targeting transmembrane helices, 70–71
- Peptidyl carrier proteins (PCPs), 249–251
- PheATE (GrsA), 253–258
- Phenylalanine (Phe) adenylation domain, 97–98, 105, 108
- Pixel labeling problem, 323–326
- Polyketide synthase (PKS), 249–251
- Polymerase chain reaction (PCR), 84
- Potluri, S., 158, 175
- Preexisting equilibrium model, 213
- Principal components analysis (PCA), 23, 188
 - calculating by SVD, 25
 - covariance matrix and, 27–44
 - description of, 25
 - directional variance and, 28–33, 39–44
 - efficient computation of, 44–45
 - ensemble averages and, 27
 - Gaussian distribution and, 30–33

- Principal components analysis (PCA) (cont.)
 inner product and, 29
 introduction to, 28–33
 inverse kinematics and, 193
 outer product and, 29
 pseudo covariance matrix and, 29–30, 36, 38–39, 42–43
 residual dipolar couplings (RDCs) and, 38–44
 Saupe matrix and, 27–28, 34–37, 43
 set point variance and, 28
- Principal order frame (POF)
 nuclear vector replacement (NVR) and, 120
 principal components analysis (PCA) and, 31–32, 38–44
 residual dipolar couplings (RDCs) and, 19, 53–54, 57,
 116–117, 151–152, 162–163, 166
- Probabilistic classifier, 188–189
- Probik, 191–195
- Proline, 1–2, 18, 58, 78, 80, 87
- Promiscuity design, 77, 82–84
- Protein backbone, 2, 3n1. *See also* Protein flexibility
 alignment tensors and, 55
 all-atom Cartesian coordinates-based model and, 191
 BD algorithm and, 104, 107–108, 111, 241
 brDEE algorithm, ch. 12
 Chain Tweak and, 191–195
 closure-constrained refinement and, 221
 computational protein design and, 87–92
 computational topology and, 412
 continuous flexibility DEE and, 107
 Denavit-Hartenberg (DH) local frame and, 192
 dihedral angles-based model and, 191
 errors in structure determination and, 304–306
 flexibility in DEE and, 107–108, 111
 generalized belief propagation (GBP) and, 231–241
 genetic algorithm (GA) and, 265–268
 JIGSAW algorithm and, 59
 K^* algorithms and, 99, 102, 104–105
 ligand-protein docking and, 346
 MARS algorithm and, 293–300
 modeling equilibrium fluctuations and, 220–221, 226
 native, 54–55
 NMR structural biology and, 17–18, 20
 nuclear magnetic resonance (NMR) and, 304–306,
 371–376
 nuclear vector replacement and, 119–120
 Probik and, 191–195
 protein modeling and, 265–266
 residual dipolar couplings (RDCs) and, 53–58, 116,
 130–153, 156–166, 169–171, 175, 333, 337–340
 secondary structure elements and, 55, ch. 8
 semidefinite programming and, 307
 side-chain flexibility and, 265–271
 SoftROC and, 265–271
 structure-based protein-ligand binding and, 342–343
- Protein Data Bank (PDB), 301, 303, 410
- Protein design, chs. 11, 12, 30, 45, 46, 49
 belief propagation algorithms and, 236–237
 entropy and, 247–248
 future directions for, 240–242
 global minimum energy conformation (GMEC) and, 69,
 88–92, 104–110, 233, 235, 237, 240–242, 332, 412
 graphical models and, 227–242
 as inference problem, 235
 negative design and, 109–110
 probabilistic, 238–240
 rotamers and, 231–241, 265–270
 side-chain flexibility and, 265–271
 support vector machine (SVM) and, 377–381
 synthetic biology problems and, 110
- Protein ensemble method (PEM), 221–226
- Protein flexibility, chs. 5, 9–12, 20–26, 30, 41, 42, 47
 ab initio methods and, 219
 Chain Tweak and, 191–195
 coiled-coil core motifs and, 265–266
 cyclic coordinate descent (CCD) algorithm and,
 220–222
 database methods and, 219
 Denavit-Hartenberg (DH) local frame and, 192
 docking and, 208
 ensemble docking and, 343
 equilibrium fluctuation models and, 219–226
 fragment ensemble method (FEM) and, 220–221
 genetic algorithm (GA) and, 265–268
 graph theory and, 200–202
 hydrophobic interactions and, 207–208
 inverse kinematics and, 191–195
 ligand-protein docking and, 345–349
 loop closure problem and, 191–195
 metadynamics and, 346–349
 missing loops and, 219
 Monte Carlo methods and, 219
 mutation and, 267–268
 normal mode analysis (NMA) and, 213–217
 protein ensemble method (PEM) and, 221–226
 protein-ligand binding and, 343–344
 receptor ensemble docking (RED) and, 215–217
 ROCK algorithm and, 205–208, 210
 side-chains and, 221, 265–271
 SoftROC and, 265–268, 265–271
 stochastic selection and, 266
- Protein folds, 21, 303
 AutoStructure and, 288–289
 backrubbs and, 108
 computational protein design and, 87–88
 cross-rotation and, 384
 CS-Rosetta and, 372
 ensembles and, 333–336, 351, 353
 errors in published, 301
 free energy and, 227, 233
 initial fold analysis and, 285
 inverse folding problem and, 87, 234
 K^* algorithm and, 104, 109
 ligand configurational entropy and, 247
 local sampling and, 333–336
 long-range order and, 335–336

- low energy and, 237
- MARS and, 296
- NOESY data interpretation and, 285, 288, 290
- nuclear vector replacement (NVR) and, 123
- orientational sampling and, 47
- peptide biosynthesis and, 249
- peptide design and, 67–68
- PheATE and, 255
- protein core design and, 266
- protein interface and, 77–78, 83–84
- protein unfolding and, 333–340
- residual dipolar couplings (RDCs) and, 129–134, 139, 142–143, 147, 151, 153–155, 161, 164, 166–175, 333–340
- simple-ring closure and, 206
- small-angle X-ray scattering and, 333–334, 337–340
- surface charge-charge interactions and, 394–397
- 434 cro protein, 269–270
- Protein interface
 - catalytic machinery and, 77, 80–81
 - directed evolution and, 83–85
 - Dunathan's hypothesis and, 83
 - interconverting homologous enzymes and, 79–80
 - minimalist active site redesign and, 77–83
 - partitioning of reaction intermediates and, 82
 - product and, 77
 - promiscuity improvement and, 77, 82–84
 - removal of catalytic nucleophiles and, 81
 - stereo- and regiochemistry control and, 82
 - substrates and, 77
 - subtilisin and, 78
- Protein kinematics, chs. 1, 2, 9, 13, 15–18, 20–24, 30
- Protein-ligand binding, 345, chs. 9, 10, 12, 25–29, 40, 41, 46
 - continuum solvent model and, 344
 - dead-end elimination (DEE) and, 107
 - density mapping and, 341
 - energy-weighted average and, 343
 - ensemble docking and, 343
 - geometry-weighted average and, 343
 - K^* algorithm and, 98–99
 - nuclear Overhauser effect (NOE) matching and, 259–263
 - protein dynamics and, 342–343
 - protein flexibility representation and, 343–344
 - receptor docking and, 344
 - residual dipolar couplings (RDCs) and, 131
 - ROCK and, 208
 - soft-core energy function and, 344
 - structure-based, 341–344
 - temperature factor and, 342
 - uncertainty in NMR structures and, 342
 - uncertainty in X-ray structures and, 341–342
- Protein-ligand matching, 259–263
- Protein loops, chs. 20, 22, 24
- Protein-protein interactions (PPIs), 67, 251, 254–255;
 - chs. 9, 10, 23, 27, 28
 - lock-and-key model and, 213
 - normal mode analysis (NMA) and, 213–217
 - preexisting equilibrium model and, 213
- Protein structure, chs. 1–18, 21–25, 27, 29–36, 39–45, 47–50
 - alpha shapes and, 410–411
 - carrier, 249–251
 - Cartesian coordinate-based root-mean-square deviation and, 351
 - denatured, 55–58
 - distance geometry and, 310–313 (*see also* Distance geometry)
 - distance root-mean-square deviation and, 351
 - ensemble representation and, 351–353
 - foldamers and, 71–74
 - glycine and, 1–2
 - inverse folding problem and, 87, 23
 - native, 53–55
 - metric space and, 352–353
 - NMR determination errors and, 301–306 (*see also* Nuclear magnetic resonance (NMR))
 - orientational sampling of interatomic vectors and, 47–51
 - predicting binding energetics from, 345–346
 - proline and, 1–2, 18, 58, 78, 80, 87
 - RDCs and, 47–51, 127–177 (*see also* Residual dipolar couplings (RDCs))
 - topology and, 399–412
 - triangulation and, 399–403, 410–411
- Proteomic disease classification algorithm, 187–189
- Pseudo covariance matrix
 - principal components analysis (PCA) and, 29–30, 36, 38–39, 42–43
 - residual dipolar couplings (RDCs) and, 36, 38–39, 42–43
- Pseudoenergy, 295
- Pseudoresidues (PRs), 293–295, 298
- PSIPRED, 295
- P. syringae syringomycin*, 84
- Pyridoxal phosphate (PLP), 83

- Q5 algorithm, 188–189
- Quaternions, 69, 121, 383–384, 391

- Ramachandran plots, 55
 - hydrophobic interactions and, 207–208
 - inverse kinematics and, 194
 - protein flexibility and, 205, 207–210
 - residual dipolar couplings (RDCs) and, 141–142, 161–162, 165, 194
 - ROCK and, 207–208
- RDC-EXACT
 - protein unfolding and, 340
 - residual dipolar couplings (RDCs) and, 130, 132–133, 145–149, 156–158, 161–162, 165, 171–172
- Receptor ensemble docking (RED), 215–217

- Reduction algorithm, 406–409
- Reductive methylation, 357–361
- Region-based approximations, 231
- Region graphs, 231
- Residual dipolar couplings (RDCs), 23
 - algorithms and, 160–167
 - alignment tensor and, 19, 55, 57, 116–123, 131, 136, 142, 144, 150, 152, 162, 167–168, 171
 - assignment problems and, 27, 116, 129, 131, 151
 - CAP and, 151
 - CH bond and, 136–137
 - chemical shift and, 129, 131–132, 135, 145–146, 150, 152, 155, 162, 166–170
 - combinatorial precision and, 146
 - computational protein design and, 94
 - conformational ensemble and, 337
 - covariance matrix and, 33–44
 - CRANS NCS method and, 160
 - defined, 25
 - denatured proteins and, 55–58
 - dihedral angles and, 127–130, 136, 140–146, 161–166
 - dipolar coupling constant and, 19
 - distance geometry and, 311–313
 - EM algorithm and, 122, 150–152, 167–171
 - ensemble averages and, 20, 27–28, 34, 222, 226, 333–337, 343, 352
 - error and, 128, 143–144, 155, 171, 175–176
 - estimation of alignment tensor and, 117
 - exact solutions for, 161–166
 - expectation graph and, 169
 - fragment computation and, 56–57
 - future development and, 43–44, 160–167
 - GD algorithm and, 131, 133, 153, 155, 161, 166, 169
 - global orientation and, 19, 150–151
 - HANA and, 133, 155–158, 171–172
 - high-throughput methods and, 129–130
 - inner product and, 36–38
 - isotope labeling scheme and, 132
 - kinematics and, 123–124, 127, 132, 138, 143–144, 160–164, 173, 175
 - Kuhn-Munkres matching algorithm and, 116
 - Kullback-Leibler distance and, 117–118
 - limitations and, 144
 - linker computation and, 57
 - local sampling and, 333–336
 - long-range order and, 335–336
 - MARS and, 296–300
 - maximum bipartite matching and, 116
 - mean tensor consistency and, 167
 - membrane proteins and, 129–133, 158, 172–173
 - methodological development of, 127–128
 - molecular coordinate frame and, 35
 - molecular replacement (MR) and, 127, 130–131, 150, 160, 166–167, 173
 - native proteins and, 53–55
 - NMR structural biology and, 115–118
 - nuclear Overhauser effect (NOE) and, 55, 115, 129–135, 142–162, 166, 169–176
 - nuclear spin and, 115
 - nuclear vector replacement (NVR) and, 119–124, 131–133, 149–153, 160–161, 166–171
 - optimization and, 139, 142–144, 144, 160, 162, 167, 171, 175
 - orientational sampling of interatomic vectors and, 47–51
 - physics of, 27
 - power of exact solutions and, 134–144
 - principal components analysis (PCA) and, 38–44
 - principal order frame (POF) and, 19, 53–54, 57, 116–117, 151–152, 162–163, 166
 - protein backbone and, 53–58, 116, 130–153, 156–166, 169–171, 175, 333, 337–340
 - protein folds and, 129–134, 139, 142–143, 147, 151, 153–155, 161, 164, 166–175, 333–340
 - protein-ligand binding and, 131
 - protein unfolding and, 333–340
 - pseudo covariance matrix and, 36, 38–39, 42–43
 - Ramachandran plots and, 141–142, 161–162, 165, 194
 - RDC-EXACT and, 130, 132–133, 145–149, 156–158, 161–162, 165, 171–172
 - resonance assignment and, 145, 150–152, 155–156, 161, 167–168, 172
 - rigidity theory and, 167
 - RNA and, 151, 158
 - rotamers and, 155–158, 160, 165, 171–172, 175
 - Saupe matrix and, 19, 34–37, 43, 55, 57, 115, 137, 162–163, 334
 - side-chain conformations and, 133, 146, 155–157, 165, 171–172, 175
 - simulated annealing (SA) and, 129–132, 135, 146, 149, 165, 173
 - small-angle X-ray scattering and, 333–334, 337–340
 - sparse data and, 127, 129–130, 133, 145–149, 152–153, 173; chs. 15–18
 - steric clash checking and, 333–335
 - structural biochemistry and, 34–38
 - structural homology detection and, 117–118
 - structure-based assignment (SBA) and, 130–131, 133, 149, 167, 332
 - structure determination problem and, 116–117
 - SYMBRANE and, 132–134, 158–160, 172–177
 - symmetric oligomers, homo-oligomers and, 158–159, 172–177
 - symmetry configuration space (SCS) and, 158–159, 173–175
 - tensor computation and, 57
 - time complexity and, 151–152
 - theorems for, 163–165, 171–172
 - TRIANGLE and, 155–158, 171–172
 - unassigned, 150–155, 166–171
 - well-packed satisfying (WPS) structures and, 159
 - X-ray crystallography and, 130–132, 143–148, 151, 153, 157, 160, 167, 173, 333, 337

- Resonance assignment, 3, 15, chs. 8, 14, 15–18, 33, 34, 37, 38, 43
Autolink and, 363–370
AutoStructure and, 287
chemical shift and, 355, 365
graph cuts and, 332
JIGSAW and, 59, 66
MARS and, 293–300
mass spectrometry and, 355–361
nuclear vector replacement (NVR) and, 119–123
protein-ligand matching and, 259–262
residual dipolar couplings (RDCs) and, 145, 150–152, 155–156, 161, 167–168, 172
topology and, 285, 287, 290
- Rigidity Optimized Conformational Kinetics (ROCK), 205–208, 210
- Rigidity theory
distance geometry and, 200, 274–277
FRODA and, 208–210
Laman's theorem and, 201
normal mode analysis (NMA) and, 197–202
Pebble Game and, 200–202
percolation and, 200
residual dipolar couplings (RDCs) and, 167
ROCK and, 205–208, 210
- RNA, 55
CAP and, 151
foldamers and, 71–72
residual dipolar couplings (RDCs) and, 151, 158
- Romeo, Fabio, 372–373
- Root mean square deviation (RMSD)
atomic coordinates, ch. 42
backbone, ch. 42
CS-Rosetta and, 374–375
ensembles and, 351–353
equilibrium fluctuations and, 226
ligand-protein docking and, 349
peptide design and, 69–70
protein flexibility and, 194, 268, 270
protein-ligand matching and, 262, 344
residual dipolar couplings (RDCs) and, 57, 141–142, 147, 153, 157, 169; chs. 15–18
X-ray crystallography and, 390
- Rosetta, 371–375
- Rotamers, 70
automated NOE assignment and, 155–158
chemical shift and, 371, 373
computational protein design and, 87–92
energy minimization and, 332
ensembles and, 98–108, 111
ligand configurational entropy and, 246
pairwise, 89
protein design and, 98–108, 111, 231–241, 265–270
protein-ligand binding and, 344
sparse residual dipolar couplings and, 155–158, 160, 165, 171–172, 175
structure determination errors and, 304–308
template energies and, 89
ROTCHK, 305n1
- Sampling tensor, 47–51
Sangiovanni-Vincentelli, Alberto, 372–373
Saupe alignment tensor, 27–28, 35, 47, 51
distance geometry and, 311–312
equilibrium fluctuations and, 227
MARS and, 298
orientational sampling and, 47, 51
principal components analysis (PCA) and, 27–28, 35–38, 43
protein unfolding and, 334, 337, 339
residual dipolar couplings (RDCs) and, 19, 34–37, 43, 55, 57, 115, 116–123, 131, 136, 137, 142, 144, 150, 152, 162–163, 167–168, 171, 334
- Saxe, J.B., 3, 279–280, 282–283
- Scalar couplings, 4–5
residual dipolar couplings (RDCs) and, 115, 131, 135, 145
spin-spin coupling constant $^3J_{\text{HNH}\alpha}$, 4–5, 15–16, 20, 59–60
- Schwartz, J., 402
- Selenosubtilisin, 78
- Self-energies, 89
- Self-rotation function, 383–384, 390–391
- Semidefinite programming, 307–310
- Sensor network localization problem, 309–310
- Sequential connectivity, 295–296
- Sharir, M., 402
- Shen, Y., 373
- Side-chain conformations, 17–18, 231, 242
backbone flexibility and, 265–271
computational protein design and, 87–88
CS-Rosetta and, 373–375
dead-end elimination (DEE) and, 106, 108
entropy estimation and, 232–233
equilibrium fluctuations and, 221
exploration for fixed backbone and, 221
fingerprint alignment and, 64–66
graph cuts and, 332
 K^* algorithm and, 99, 102, 104
minimization and, 373–374
nuclear Overhauser effect (NOE) and, 133, 155–157, 171–172
optimization of, 215–216
positioning problem and, 307–308
protein-ligand binding and, 343
residual dipolar couplings (RDCs) and, 133, 146, 155–157, 165, 171–172, 175
- Side specificity problem, 380
- Signature sequence, 97–98
- Simplicial complexes, 48, 400–404
- Simulated annealing (SA)
chemical shift data and, 371–373
computational protein design and, 92, 94

- Simulated annealing (SA) (cont.)
 distance geometry and, 279
 equilibrium fluctuations and, 219
 molecular dynamics (MD) and, 129, 132, 141, 145, 173, 279
 peptide design and, 70
 protein backbone and, 265–266
 protein-ligand NOE matching and, 261
 residual dipolar couplings (RDCs) and, 129–132, 135, 146, 149, 165, 173
 Rosetta and, 371–373
 side-chain/backbone flexibility and, 265–266
 SoftROC and, 266
 Single-ring closure, 206
 Singular value decomposition (SVD)
 basics of, 23–26
 distance geometry and, 312
 MARS and, 298
 residual dipolar couplings (RDCs) and, 55, 57, 116, 144, 152
 Saupe matrix and, 19
 Site-directed mutagenesis, 97–98
 SLIDE, 208
 SLIDEWIN, 194
 Small-angle X-ray scattering (SAXS), 333–334, 337–340
 Smith normal form, 402–403, 409
 Smith, H.J.S., 409
 Soft repacking of cores (SoftROC)
 energy calculations and, 268–269
 experiments on 434 cro and, 269–270
 experiments on T4 lysozyme and, 270
 fitness evaluation and, 266
 genetic algorithm (GA) optimization and, 266–268
 Monte Carlo sampling and, 268
 mutation and, 267–268
 protein flexibility and, 265–271
 ROC settings and, 269
 stochastic selection and, 266
 SPARKY, 293
 SPARTA, 373–374
 Spatial coherence, 323–324, 326
 Spectrum matching, 261–262
 Spin-spin coupling, 4–5, 15–16, 20, 59–60
 Spin system, 363, 365–367, 370
Staphylococcus aureus, 109
 Steady-state assays, 254–255
 Steric clash checking, 333–335
 Steric filter, 100
 Stochastic algorithms, chs. 6, 9, 10, 11, 20, 22–26, 30, 31, 41, 43, 44, 45, 49
 Stochastic selection, 266
 Stop-threshold, 100
 Structure-activity relation (SAR) by NMR, 21, 131, 411
 Structure-based assignment (SBA), 130–131, 133, 149, 167, 332
 Structure refinement, 119, 151, 290
 STYX protein, 80
 Subgraph realization, 276
 Subtilisin, 78
 Support vector machine (SVM), 377–381
 Surface charge-charge interactions, 393–397
 Surface residues, 89, 393, 395, 397
 SYMBRANE, 132–134, 158–160, 172–177
 Symmetric oligomers, 172–177
 Symmetry configuration space (SCS), 158–159, 173–175
 T4 lysozyme, 270
 TALOS, 162
 Temperature factor, 342
 Tensors
 3- α -HSD, 79–80
 alignment, 19, 27–28 (*see also* Alignment tensor)
 linker computation and, 57
 mean consistency and, 167
 merging of, 57
 nuclear vector replacement (NVR) and, 119–121
 orientational sampling and, 47–51
 updating of, 57
 Yan-Donald notation for, 137
 Thin-layer chromatography (TLC), 255–256
 Thiolsubtilisin, 78
 Three-dimensional HNH α ($^3J_{\text{HNH}\alpha}$), 4–5, 15–16, 20, 59–60
 Time complexity
 Autolink and, 367
 computational protein design and, 92
 cross-rotation analysis algorithm and, 384
 distance geometry and, 280, 312
 Laman's theorem and, 201
 multiple receptor conformations (MRCs) and, 215
 nuclear vector replacement (NVR) and, 121–123
 residual dipolar couplings (RDCs) and, 151–152
 surface charge-charge interactions and, 395
 TK-SA model, 1178
 Topology
 Abelian groups and, 402–406, 409
 alpha shapes and, 410–411
 AutoStructure and, 285–292
 Betti number and, 402, 404, 407, 410
 Bézier systems and, 403
 chain groups and, 402, 404–407
 cohomology operation and, 411–412
 complexity and, 403
 computational, 399–412
 critical values and, 410
 distance geometry and, 273
 edges and, 399
 Euler number and, 404
 faces and, 410
 free chain complex and, 405
 fundamental group and, 400–401
 Gaussian network model and, 213–214
 grid, 315, 321

- homology modeling and, 399–412
- image and, 402
- invariance and, 399
- kernels of map, 402, 405
- Klein bottle and, 403–404
- matrices and, 401–409
- meaning of term, 399
- NOESY data interpretation and, 285–292
- orientation and, 399, 401, 404, 407
- protein folds and, 175 (*see also* Protein folds)
- reduction algorithm and, 406–409
- simplicial complexes and, 400–404
- Smith normal form and, 402–403, 409
- torsion coefficients and, 402, 407
- triangulation and, 399–411
- Torsion coefficients, 402, 407
- Total correlation spectroscopy (TOCSY)
 - JIGSAW and, 59–60, 64–66
 - NMR structural biology and, 17–18, 20
 - residual dipolar couplings (RDCs) and, 172
- Transductive support vector machine (TSVM), 378, 381
- Trees, 4
 - belief propagation and, 236, 238, 241
 - conformation, 100, 106–107, 141, 162
 - ensembles and, 100, 106–107
 - residual dipolar couplings (RDCs) and, 141–142, 162
- TRIANGLE algorithm, 155–158, 171–172
- Triangle inequality, 352
- Triangle relationship, 156
- Triangles
 - orientational sampling and, 48, 50
 - simplex, 48
 - topology and, 399–401
- Triangulation
 - complexity and, 403
 - definition of, 400–401
 - Delaunay, 399
 - edges and, 399
 - as finite simplicial complex, 400–401
 - fundamental group and, 400–401
 - matrices and, 401–409
 - orientation and, 399, 401, 404, 407
 - simplicial complexes and, 400–402
 - topology and, 399–411
- Trichloroacetic acid (TCA), 255–256

- Unfolding, 58; chs. 7, 39
 - ensemble computation and, 333–336
 - local sampling and, 333–336
 - long-range order and, 335–336
 - residual dipolar couplings (RDCs) and, 333–340
 - small-angle X-ray scattering and, 333–334, 337–340
- Urea-denatured eglin C, 58

- van der Waals force, 88–89, 207, 227, 410

- Wang, Lincong
 - nuclear vector replacement (NVR) and, 119, 124
 - residual dipolar couplings (RDCs) and, 55–56, 58, 130, 133, 138, 145–146, 157, 162–165, 172
- Well-packed satisfying (WPS) structures, 159
- Wiggle loops, 193
- Within-class scatter matrix, 188

- XEASY, 293
- X-ray crystallography, 7
 - analysis and use of crystal structures, chs. 9–12, 14, 20–30, 40, 41, 49, 50
 - Bragg's law and, 341
 - computational protein design and, 87
 - cross-rotation function and, 160, 383
 - data resolution and, 341
 - diffraction and, 160, 341, 383
 - ensembles and, 110
 - equilibrium fluctuations and, 219
 - errors and, 301
 - Fourier moduli and, 387
 - ligand configurational entropy and, 245
 - methodology, chs. 40, 47, 48
 - molecular replacement and, 387–391
 - noncrystallographic symmetry (NCS) and, 387–391
 - normal-mode analysis (NMA) and, 387–390
 - nuclear vector replacement (NVR) and, 119
 - phase problem and, 387
 - principal components analysis and, 43
 - protein flexibility and, 7
 - protein-ligand binding and, 341–342
 - protein-ligand NOE matching and, 259
 - residual dipolar couplings (RDCs) and, 130–132, 143–148, 151, 153, 157, 160, 167, 173, 333, 337

- Yan, Anthony K., 27–45
- Yanover, C., 235, 237–238

- Zabih, R., chs. 37–38
- Zargovic, B., 352–353
- Zeng, J., 147, 155
- Zinc finger, 88, 91–92
- Z-scores, 305–306, 394–395
- Zweckstetter, M., 152, 293