

Index

a

accessible surface area (ASA) 15–16, 23, 28

adaptive biasing force (ABF) method 151, 153, 155

adaptive umbrella sampling 151

allosteric modulators 328–329

alternative splicing (AS) 359

 in diseases 368–369

 and intrinsically disordered regions 362–367

 protein-protein interactions 367–368

 and protein structure 362

antimicrobial peptides (AMPs) 195, 282, 294, 307–308

antisense oligos (ASOs) 278–279

APEX 231, 273, 281, 282

asymmetric unit 107–108

ATTRACT 58, 62, 66

ATTRACT-EM 116

autism spectrum disorders 369

automated assignment

 crystallographic data 107–108

 leveraging evolutionary information 109–110

 machine-learning methods 108–109

 methods 106–107

b

barcoded nucleosome libraries 242, 245

Bennet acceptance ratio (BAR) 153, 296, 298, 301

biological assemblies 54, 106–108

Biological General Repository for Interaction Datasets (BioGRID) 78–79, 340–341, 369

biological membranes 293–308

biomolecular modeling 41

biomolecular structure 39–40, 46

biomolecules 11, 39–41, 77, 128, 129, 139, 271, 362

Boltzmann distribution 163, 167–168, 179, 182

bonded interactions 41

Brownian dynamics 57, 147–149, 152

Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm 44

c

cell machinery 127–128

cell membrane 2, 133, 139, 293, 295, 307, 384–385

cellular assays 227–231, 236–239, 245–246

cellular chromatin 227, 248

chaperones 241–248

Chapman–Kolmogorov (CK) test 174

chromatin

 biochemical composition 213

 complex composition 213

 remodelers 241–242

chromatin assembly 241–242, 245, 247

chromatin immunoprecipitation (ChIP) 203–204, 229–231, 236, 238, 245

chromatin interactomics 247–248

- clustering 84–86, 89–92, 110, 115, 134, 169, 172–173, 177, 184, 185, 297
 - Clustering with Overlapping Neighborhood Expansion (ClusterONE) 89
 - algorithm 90–92
 - definition 89–90
 - coarse graining 147–149, 174–175
 - coevolutionary analysis 40, 45, 46
 - collective variables (CVs) 145, 151, 153–156, 164
 - combinatorial docking 110
 - assisted docking 116–117
 - homology-based complex reconstruction 114–115
 - prediction from sequence 115–116
 - computational methods 15, 28, 54, 56, 131, 138, 146–157, 244–245, 334, 344
 - computational protein-protein docking
 - biological macromolecules 54–55
 - complex structures 53–54
 - conformational changes 59–61
 - definition 54
 - flexible refinement 64–65
 - integration of bioinformatics and experimental data 61–62
 - quality criteria 54–55
 - rigid body approaches 56–59
 - scoring 66–67
 - steps 55
 - template-based 62–64
 - types 56
 - contact guided protein and RNA structure prediction 45
 - “corpora non agunt nisi fixata” 316
 - couplings matrix 43
 - Critical Assessment of Predicted Interactions (CAPRI) 54, 55, 67, 111, 276, 277
 - Critical Assessment of protein Structure Prediction (CASP) 7, 139
 - cross linking mass spectrometry (CL-MS) 128, 133, 136
 - cryo electron microscopy (cryoEM) 9, 18, 39, 54, 62, 68, 102, 117, 127, 128, 131, 132, 136–138, 215, 403
 - cryo-electron tomography (cryo-ET) 128, 133, 138
 - cyclin-dependent kinases 390, 393
 - cysteine 2, 222, 275, 299, 302, 318, 381
- d**
- dative bonds 318, 324
 - deep mutation scanning (DMS) 348
 - deoxyribonucleic acid (DNA) 3, 9–11, 39, 83, 84, 101, 195–208, 231–246, 271–272, 275, 315, 320, 325, 341, 362–363, 383–385, 401–403
 - Direct Coupling Analysis (DCA) 40, 42–44
 - disordered proteins 9–10, 45, 363, 403
 - docking methods 54, 55, 60, 64, 67, 68, 110, 115
 - DockStar 113
 - Domain-Aware Cohesiveness Optimization (DACO) 92–95, 207
- e**
- eigenfunctions 167–170, 174
 - endpoint methods 149–150
 - enhanced CLIP (eCLIP) 280
 - enhanced sampling methods 146, 175
 - epidermal growth factor receptors (EGFRs) 305
 - eukaryotic genes 205, 359
 - exocytosis 133, 139
- f**
- Förster resonance energy transfer (FRET) 116, 129, 164, 181, 182, 227, 242, 245
 - fluorescence polarization (FP) assays 40, 223, 226, 241, 245
 - folded proteins 3, 10
 - free energy 143, 144, 146–156
 - Gibbs 144–145

g

gatekeeper mutation 328
 generalized Born (GB) model 66, 149, 150
 generalized/Hamiltonian
 replica-exchange methods 156
 Gibbs free energy 107, 144–145, 344
 glycoprotein A (GpA) 303–306
 G-protein coupled receptors (GPCRs)
 301, 306, 308, 317, 328
 guardian of the cell 197

h

HADDOCK 58, 62, 64, 66, 117, 132, 139, 347
 helical transmembrane proteins 6, 7
 hidden Markov models (HMM) 175, 182, 183, 338, 367
 histone posttranslational modifications 215–231
 cellular assays to characterize 227–231
 homodimeric interfaces 46
 homology modeling 39
 HUMA 338
 hybrid methods 117, 131–133
 hydrogen bonds
 allosteric effects 327–329
 anion- π interaction 325–326
 bifurcated 322–323
 cation- π interaction 325
 classification 321
 definition 319–320
 delocalized π -electron systems 325
 entropic aspects 327
 fluorine 322
 by force fields and docking simulations 326–327
 halogen bonds 323–324
 in biological systems 320–321
 nitrogen versus oxygen 322
 unusual protein-ligand contacts 326
 van der Waals interactions 324
 weak 321–322
 HyperTRIBE 280

i

immunofluorescence (IF) 227, 229
 immunoprecipitation assays 203–204, 244
 induced fit 102, 176, 177, 184, 316
 in-silico protein structure prediction 39
in situ structural biology 128
 integral transmembrane proteins 6
 Integrative Modeling Platform (IMP)
 132, 134, 139
 integrative structural biology 128, 131–133
 International Molecular Exchange (IMEx)
 consortium 79
 intrinsically disordered regions (IDR)
 147, 273, 359, 360, 362–367, 402
 Inverse Potts Model 40, 43–45
in vitro remodeling assays 245
 isothermal-isobaric ensemble 143–145

k

kinase specificity 390

l

light-based microscopy 129
 linear interaction energy (LIE) 149, 150
 linear response approximation 149
 liquid-liquid phase separation (LLPS)
 215, 273, 384
 live-cell fluorescent microscopy 129
 local statistical inference, limitations of 41–42
 long disorder regions (LDRs) 10

m

machine-learning methods 26, 28, 108–109, 338, 342–349, 401
 mammalian DNA 231
 Markov state models (MSM)
 adaptive and enhanced sampling strategies 175
 association-dissociation path ensemble 177–178
 and coarse-graining 174–175
 clustering 172–173

- Markov state models (MSM) (*contd.*)
 dimensionality reduction 171–172
 estimation, validation, and analysis
 169–178
 feature selection 170–171
 model estimation and validation
 173–174
 protein–protein encounters 176–177
 sources of errors and uncertainty
 179–180
 spectral gaps 174–175
 theory and properties 165–169
 VAC and VAMP 169–170
- mass spectrometry 61, 80, 106, 117, 128,
 223, 225, 229, 232, 242, 276, 282,
 361, 391, 397, 401
- maximum-entropy principle 42
- mCSM 347
- mean-field DCA (mfDCA) 44
- mean first passage times (MFPT) 178
- mean force, potential of 150–155
- metadynamics 153–156
- methylation-sensitive SELEX 235
- MM/GBSA 149–150
- MM/PBSA 149–150
- Molecular Complex Detection (MCODE)
 84
 algorithm 86–88
 definition 85
- molecular dynamics (MD) simulations
 60, 64–67, 146, 149, 151, 164–165,
 176–177, 181, 200, 204–205,
 296–297, 301–302, 306, 345
- M-TASSER 63, 115
- multiple sequence alignment (MSA)
 42–46, 344, 346
- multiplexed arrays of variant effect
 (MAVE) 348
- mutation
 biophysical properties 344–345
 dynamic structural annotation
 pipelines 339–342
 mechanistic effects 345–348
 phenotypic effects 343–344
 predicting effects 342–343
 structural annotation methods
 334–335
 databases 335–338
- MutDB 335
- myristoylation 299, 301, 303
- n**
- non-symmetric complexes 111
- normal-mode analysis 150
- NPT ensemble 144
- nuclear magnetic resonance (NMR)
 spectroscopy 9, 26, 39, 45, 53,
 102, 127, 131, 164, 181, 196, 208,
 241, 272, 298, 324
- nucleic acid methylation 235
- nucleosomes 199, 204, 213, 222,
 224–227, 229, 232, 241, 242,
 244–248
- o**
- obligatory interfaces 25
- oligomerization 6, 154, 156, 306
- p**
- palmitoylation 296, 303, 363
- parallel tempering 154–156
- pathway methods 150, 151, 156
- peripheral membrane proteins (PMPs)
 294–303, 307, 308, 384
- Perron Cluster-Cluster Analysis (PCCA)
 174, 175
- phosphatases 295, 299, 382, 386, 387,
 389–391, 396
- phospho-amino acids 383, 392
- phospho-proteomics 383, 393, 396
- phosphorylation 381
 in cellular signaling 386–387
 mechanisms of 390–391
 molecular switches 388–390
 phospho-switch based biological
 regulation 392–395
 structural and functional effects
 383–386
- PhyreRisk 342
- physico-chemical interactions 195, 316

- PinSnps 340, 341
- plasma membrane 7, 133, 136, 293–303, 305, 306, 308
- pleckstrin homology (PH) domains 300
- Poisson-Boltzmann (PB) equation 150
- position-specific scoring matrix (PSSM) 28, 202, 204, 344–345
- post-translational modifications (PTM) 338, 381
 - histones 215–231
 - molecular parameters 226–227
 - peptides and nucleosomal templates for 222–224
 - qualitative analysis 224–226
- potential of mean force (PMF) 145, 150–155
- Potts model 40, 42–45
- prenylation 299, 301–303, 381
- principal component analysis (PCA) 172, 397
- protein binding sites 15, 230, 236
- protein complexes
 - computational approaches 110–117
 - quaternary structure 101–102
- Protein Data Bank (PDB) 18, 21–24, 26, 102, 106–107, 150, 316, 334–336, 338, 340–341, 349, 369, 373
- protein interactions, types of 41, 401
- protein–ligand interactions 147, 317, 319, 327, 329
- protein–protein binding 183
 - free energy 24
 - interfaces 15–30
- protein–protein encounters 143–156, 163–186
- protein–protein interaction interfaces 109, 114–115
 - classification 102, 104–105
- protein–protein interaction network (PPIN) 77
 - of human 83
 - identify protein complexes 84–94
 - of model organisms 80
 - Molecular Complex Detection 84–88
 - Saccharomyces cerevisiae* 80–83
- protein–protein interactions (PPIs)
 - alternative splicing 367–373
 - public data repositories 79
- protein–protein interface
 - amino acid composition 22–23
 - biological vs. crystal interfaces 26
 - characterization 29
 - conserved residues and hot spots 28–29
 - definition 15–18
 - distance-based methods 15–16
 - gap volume 22
 - homo- and hetero-dimeric complexes 24
 - non-obligate and obligate complexes 25
 - regions 16–17
 - secondary structure 23
 - structure 16
 - surface area 21
 - 3D structures 21
 - transient and permanent complexes 25–26
 - types 27
- proteins
 - active sites 3–6
 - binding interfaces 10–11
 - composition 2–3
 - conformational dynamics 8
 - disordered 9–10
 - domains 1–2
 - evolutionary conservation 10
 - folded structure 7
 - large-scale domain motions 8–9
 - membrane 6–7
 - N-terminal and C-terminal tails 9
 - secondary structure elements 3
 - size 1
 - surface dynamics 9
 - surface loops 11
 - post-translational modifications 11
- proximity ligation assays (PLA) 229
- proximity-dependent biotinylation 282
- PROXIMO 117

pseudo-likelihood maximization direct coupling analysis (plmDCA) 44–45

q

quaternary protein structure 101

r

Ras proteins function 301

relative accessible surface area (rASA) 15–16

relative binding free energies 147, 156

replica-exchange methods 155–156

ribonucleoprotein complexes 359, 384

RNA modifications 232, 234, 236

RNA polymerases 274

RNA-protein interaction detection (RaPID) 279, 280

RNA-protein interactions 272–273, 275–276, 278–280, 282

RNA-protein interactomics

co-purification methods 280

interactomes 278–280

metabolic RNA labelling with modified nucleobases 273

proximity-dependent labelling methods 280–282

RBPome 276–278

RNA-protein crosslinking 274–276

s

SAAMBE-SEQ 345–347

Saccharomyces cerevisiae 4, 80–83, 88, 89, 92, 196, 207, 236, 238, 274, 336, 393

salt bridges 24–25, 278, 317–318, 326, 330, 333, 383, 387

short linear motifs (SLiMs) 363, 366–367, 391–392, 396

single-nucleotide variants (SNVs) 333, 341, 343

small angle X-ray scattering (SAXS) 39, 62, 116, 156, 241

in solution 62

solvent accessible surface area (SASA) 16, 150

spliceosome 359, 361, 384

statistical inference of coevolution 41–43

steered molecular dynamics (steered MD) 153

super-resolution microscopy methods 129

surface plasmon resonance (SPR) 226

symmetric complexes 105, 111, 113

t

targets of RBPs identified by editing (TRIBES) 280–282

thermodynamic ensembles 143–145

thermodynamic integration (TI) 151

3D-MOSAIC 113–114, 116

time-lagged independent component analysis (TICA) 172

transition path theory (TPT) 177–178

trans-membrane proteins (TMPs) 6–7, 294–296, 300–301, 303–308, 317

transcript variants 359–360, 368

transcription factors (TFs) 195

binding sites 201

detection 201–204

position-specific scoring matrix 204

cis-regulatory modules 205–207

dimerization 198

DNA curvature/bending 200

epigenetic modification 199

gene expression 207

modifications 200–201

molecular dynamics 204–205

sequence recognition principle 197–198

transmembrane proteins (TMPs) 6–7, 294, 300, 303–308, 317

two-component signal transduction system (TCS) 45–46

tyrosine phosphorylation 382–383,
386–387

u

UHRF1 239–241, 247

umbrella sampling (US) 151–153,
184

v

variational approach for conformational
dynamics (VAC) 169–170

variational approach for Markov processes
(VAMP) 169–171, 173, 185

van der Waals interactions 150, 305, 319,
324, 326, 330

VarQ 341–342

w

water soluble proteins 2, 6

weighted histogram analysis method
(WHAM) 151

x

X-ray crystallography 9, 26, 39, 53, 102,
106, 127, 131–132, 164, 196, 200,
208, 298, 315

