

Index

a

- ab initio* methods 268, 303, 304, 320, 327–333, 335–337
- ab initio* molecular dynamics (MD) 303, 304
- absorption band 168
- acidity 128, 130, 336
- activity cliffs 458, 468, 470
- addition reaction 141
- adjacency matrix 18–21, 367
- ADME (absorption, distribution, metabolism and excretion) 384, 466
- advanced interatomic potentials 290–291
- adverse drug reactions (ADRs) 479
- agglomerative HCA 420
- ALADDIN 362
- algorithms
 - backtracking 83, 236–239, 251, 258
 - brute force 237
 - complexity 232
 - deep learning 453
 - depth-first search 238
 - dynamic programming 256–258
 - efficiency 232
 - hash 49
 - heuristic 415
 - leapfrog 304
 - learning 380, 453, 458, 529
- alignment
 - global 256
 - local 258
 - pair-wise 254
 - target-template 516
- allosteric 315, 382
- AMBER 280, 281, 283, 284, 286, 288, 293–295, 303, 304, 311, 312, 315, 316
- AM1 326, 327
- analytical chemistry 214, 399, 423, 434
- angle bending 283–284, 289, 295
- antimalarial 479, 480
- antioxidation 369
- applicability domain (AD) 158, 346, 434, 450, 466, 468, 475–478
- application programming interface (API) 71, 222
- Arabidopsis Information Resource 207
- arithmetic mean 288, 402, 414, 433
- aromaticity 29, 48, 79, 109, 110, 144, 193
- aromatization 469, 484, 486
- artificial intelligence (AI) 166, 168, 453
- artificial neural network (ANN) 345, 346, 438, 442–451, 467
- asphericity 378
- association coefficients 247, 248
- atom-by-atom 163, 164, 237, 243
- atomic mean force potentials 517
- atomic orbitals (AOs) 321–325, 327–329
- atom pairs (AP) 249, 283, 287, 288, 303, 304, 352, 359–362
- atom surface assignments (ASA) 381
- atropisomers 32
- attenuation models 271–277, 384
- augmented annealing 445
- augmented atoms (AA) 249

- Augmented Connectivity Molecular Formula (ACMF) 162
- autocorrelation vectors 249, 370, 378
- automorphism searching 74, 241–242
- autoscaling 408, 470
- average linkage 420
- average predictive ability 414, 415
- b**
- backtracking algorithm 83, 236–239, 251, 258
- bagging 432
- Balaban index 234, 364
- Ball-and Stick Model 101–102
- barrier of rotation 284
- basicity 336
- basis set 157, 321, 324, 327–329
- Bayes
- information criterion 400, 414
 - rule 457
- Bayesian
- methods 456
 - probability 359
 - regularization 445, 457
- Bayesian neural network (BNN) 459–461
- Beilstein Registry Number (BRN) 234
- B*-factor 314
- bias 186, 413, 444, 455
- bibliographic databases 213, 214
- full text 216, 217
 - patent 214, 215
 - searching 216
- Big Data 159, 453, 466, 526
- binary QSAR 467
- Bingo 222
- binning 250, 359
- bioactivity 192, 199, 209, 481, 486, 487
- biochemical reactions 123, 136, 148, 360
- bioinformatics 479–518
- Bioinformatics Links Directory 207, 208
- BioJava 109
- Biological Abstracts (BIOSIS) 214
- biological activity 4, 33, 35, 58, 91, 155, 169, 174, 192, 202, 358, 359, 362, 366, 375, 381, 384, 385, 456, 458, 466, 471, 475, 478
- biological information 192, 206, 207, 501
- biological macromolecules 44, 102–103, 307
- biomolecular systems 294, 308, 316
- biopolymers 195–198, 207, 208, 293
- BIOVIA 71, 134, 220, 222
- bit collisions 240
- bitmap 160
- bit string 160, 239, 366
- bit vectors 160, 163, 239–241, 246, 247, 249, 250, 252, 356
- biuniqueness 11, 12, 25
- BLAST 195–197, 258, 259, 503, 505–507, 509, 516
- blastn 509
- blastp 503, 505
- blastx 509
- blood-brain barrier 366
- BLOSUM matrices 256, 503–509
- Boltzmann constant 304
- bond counts 359, 364, 377
- bond–electron (BE) matrix 21–23
- bond increments 287
- bond rotation 96, 250, 284, 302
- bond stretching 282–283, 289, 295
- Boolean array 160, 353
- bootstrap 416, 430, 431
- Born free energy of solvation 307
- Born–Oppenheimer approximation 280
- boxplot 401, 402, 433
- box scaling 170
- BRENDA (Braunschweiger Enzymdatenbank) 208
- brute force algorithm 237
- Buckingham potential 288
- Burden matrix 367
- c**
- CACTVS 58, 71, 75, 86, 222
- Cahn–Ingold–Prelog (CIP) rules 34, 57, 86–88
- calcium channel 360

- calibration 6, 175, 294, 404, 406, 411–414, 423, 429
- Cambridge Crystallographic Data Centre (CCDC) 67
- Cambridge Structural Database (CSD) 91, 99, 199
- canonical
 - ensemble 309
 - numbering 77, 233
 - numeration 485
- canonicalization 25, 48, 76–78
- CAplus 189, 206, 210, 213–216, 218–220
- Capped Sticks Model 101
- cardinality 16
- Cartesian coordinates 36, 50, 51, 55, 58, 63, 94, 98, 159, 160, 280, 303, 307, 308, 379–381
- CAS numbers 485
- CASREACT 189, 210–212, 216
- CAST 87
- catalog databases 190, 200
- cell membranes 315
- cell organelles 315
- centering 6, 404, 407–408
- centroid 407, 420, 421, 429
- Certara 48, 53, 384
- charge
 - distribution 126–127, 145, 147, 271–272, 274, 275, 286, 290, 295, 333
 - net atomic 286, 334–335, 338
- charge–charge interactions 286, 306
- charged polar surface area (CPSA) 336
- CHARMM 283, 286, 288, 293–295, 303
- CHARTS 47
- CHELP 335
- ChemAxon 70, 222, 469, 470, 484
- ChEMBL 208, 234, 482, 483
- ChemBridge 479
- CHEMCATS 189, 200
- Chemical Abstracts Service (CAS)
 - registry 12, 77, 91, 162, 188, 189, 191, 194–197, 201, 202, 204, 206, 209, 215, 216, 482
- chemical–biological read across (CBRA) 471
- chemical compound representation 6, 363, 527
- Chemical Markup Language (CML) 44, 45, 68, 135, 244
- chemical nomenclature 11–12
- chemical notations 12–14
- chemical reactions 121–151
- chemical space 244, 246, 249, 251
- Chemical Structure Lookup Service (CSLS) 192
- ChemInform RX database 146
- Chemistry Development Kit (CDK) 47, 70, 109, 222, 485, 486
- CHEMLIST 189
- CHEMnetBASE 201, 204
- chemogenomics 483–485
- chemometrics 175, 399, 400, 406, 418, 423, 425, 431
- chemotypes 68, 69, 244–245, 356–358, 484
- Chemspider 192, 200, 201, 234, 485
- chirality 33–36, 86–88, 198
 - characterization 47
 - codes 376, 377, 387, 388
 - descriptors 375
- CHIRON 86
- chord 80, 83, 86
- CHORTLES 47
- chromatographic data 203, 204
- CHUCKLES 47
- CIF2PDB 70
- CIFTEX 67
- CIFTr 70
- citation databases 212–219
- city-block distance 247, 409
- classification trees (CART) 429, 432, 433
- clinical trials 208, 209
- Clustal Omega 196, 511
- Clustal X 509
- Clustal W 509
- cluster analysis 406, 412, 417, 419–420
- cluster detection 419
- clustering 145, 419, 421
- coarse graining 290

- coefficient
 - adjusted squared correlation 414
 - association 247, 248
 - correlation 247, 404, 407, 429
 - cosine 247, 355, 356
 - distance 247
 - distribution 366
 - energy 129
 - Fourier 167, 175
 - Hadamard 167
 - Hamming 356
 - Jaccard similarity 410
 - orbital 129, 130, 275
 - Pearson correlation 414, 424, 472
 - probabilistic 247
 - regression 423–425, 450
 - simple matching 248, 404
 - Tanimoto 248, 253, 254, 356, 358, 485, 486
 - wavelet transformation 175
- combinatorial chemistry 28, 272
- combinatorial libraries 48, 134, 231, 243, 377, 379, 479
- command-driven (retrieval language) systems 187, 188
- comparative molecular field analysis (CoMFA) 364, 383, 467
- comparative molecular similarity analysis (CoMSIA) 384
- comparative molecular surface analysis (CoMSA) 382
- complete linkage 420, 423
- complete neglect of differential overlap (CNDO) 325, 326
- completeness and non-redundancy 232
- complexity exponential 238
- compound databases 191–206
- computational chemistry 267–339, 345
- computer-aided drug design (CADD) 315
- CONCORD 250
- configuration interaction (CI) 330, 331
- conformation 250, 302, 361, 376
- conformational analysis 96, 302
- conformational ensemble 385
- conformational flexibility 251
- conformational space 96, 280
- conformer generation 94, 251
- connection table (CT) 23–25, 28–31, 36, 51–54, 71, 76, 93, 135, 140, 150, 156, 159, 191, 195, 232, 367
- Connolly surface 104–106
- ConQuest 199
- consensus models 459
- consensus prediction 468, 478, 479
- constant energy 308–311
- continuum solvent models 306
- coordinate system 36, 37, 50, 159
- CORINA 92, 250, 321, 361, 427
- correct classification rate (CCR) 471, 473
- correlation
 - adjusted squared 414
 - coefficient 247, 400, 429, 472
 - electron 330–332, 337
 - Kendall tau 404
 - Pearson 414, 424, 472
 - Spearman rank 404
- COSMOS DB 222
- cost function 456
- Coulomb energy 306, 383
- Coulomb's law 286, 307, 310
- Coulson analysis 334
- counter-propagation ANN 442, 445–447
- coupled cluster (CC) method 331
- covariance matrix 375, 400, 407, 410, 418, 424, 430
- CPK model 102
- cross-correlation 360
- cross entropy 432
- cross terms 281, 289, 408
- cross-validation 415–416, 470–472, 475
- CrossFire 188, 189
- cryo-electron microscopy 107, 198
- Crystallographic Information File (CIF) 45, 65, 198
- Crystallography 66, 135, 198
- Crystallography Open Database 192, 199
- C1s-ESCA shift 274

- CSRML (Chemical Subgraphs and Reactions Markup Language) 68–69, 244, 245, 356
- CTfile, 52 53
- cut-off 288, 302, 304, 308, 310, 384, 477
- cyclomatic number 81
- d**
- DAT file 171
- data
- accuracy 170, 483
 - analysis methods 6, 345, 387, 528
 - chromatographic 203, 204
 - curation 7, 346, 466, 470, 482–487
 - cycle 346, 356, 481–483, 487, 488
 - mining 68, 110, 156, 158, 163, 171, 172, 356, 480, 481, 488
 - processing 12, 44, 109, 167, 204, 481, 482
 - science 481, 487, 488
 - thermochemical 202–204, 268
 - thermodynamic 202–204, 291, 312
 - transformation 175, 176
 - types 156–170, 177, 222, 351, 363
- databases 185–223
- mining 467, 478
 - search 160, 177, 187, 191, 244–245, 254, 258, 353, 499, 507–509
 - substance 191
 - system 74, 185–191, 200, 205, 208, 219, 222, 353
- data complexity 173–174
- dataset curation 467, 469–470
- data storage 169–175, 406
- data type
- bit vector 160–164
 - chemical reactions 164–165
 - molecular spectra 165–168
 - molecular structures 159–160
 - numerical data 157–159
- Daylight Chemical Information Systems 134–135
- DDAT file 171
- deep learning 6, 346, 453–461, 528
- deep neural network (DNN) 454–459
- degree-of-fit 476
- deletion 17, 255, 506
- delocalization 29–30, 127, 193
- DelPhi 307
- dendrograms 417, 419, 420, 423
- denoising smoothing (DS) 176
- density functional theory (DFT) 6, 282, 320, 325, 332–334
- density matrix 335
- depth-first search algorithm 238
- Derwent Drug File 214–215
- Derwent Patents Citation Index (DPCI) 219
- Derwent World Patent Index (WPI) 28, 219
- descriptor 6, 36, 79, 158, 163, 170, 171, 246, 247, 250, 252, 277, 335, 349–389, 423, 427, 458, 467, 480, 485
- 0D molecular descriptors 363, 364
 - 1D molecular descriptors 363, 364
 - 2D molecular descriptors 352–361, 365–369
 - 3D molecular descriptors 361–363, 369–375, 387
 - 4D molecular descriptors 384–385
- atom pair 359, 361
- BCUT 364, 367–368
- biological 471
- CATS2D (Chemically Advanced Template Search) 360–362
- CATS3D 361–362
- CHARGE3D 361–362
- chemical structure 385–387
- chirality 375–382
- classification 351
- coding 249–250
- E-state 367
- EVA 375
- fragment 367
- molecular shape 377–378, 385
- molecular surfaces 377–378
- pharmacophores 362
- physicochemical 146–148, 360, 361, 369
- quantum chemical 126, 295, 363, 385
- RDF code 371–375

- descriptor (*contd.*)
- selection 249–250
 - surface 336
 - topological 21, 352, 364–366, 387, 388
 - TPSA 366, 382
 - WHIM 364, 375
- Design Institute for Physical Property Research (DIPPR) 202–203
- Detherm 202, 203
- DGEOM 92
- diagram method 255
- diastereoisomers 32, 87
- Dice 247, 248
- dictionary definition language (DDL) 66
- dielectric constant 27, 286, 306, 307, 383
- Diels–Alder reaction 124, 129
- digital object identifier (DOI) 217
- dipole moment 90, 274, 275, 326, 335, 364, 385
- discriminant partial least squares (D-PLS) 406–407
- dispersion 288, 294, 333, 359
- dissimilarity 355, 473
- distance
- coefficients 247
 - matrix 20–21, 87, 254, 365, 420
 - topological 20, 249, 368, 447
- distribution coefficient 366
- divide and conquer (D&C) 337
- Django 222
- DNA codon 508
- DNA Data Bank of Japan (DDBJ) 197, 499
- docking 37, 91, 199, 316, 340
- Document Type Definition (DTD) 66, 68
- Dortmund Data Bank (DDB) 202, 203
- dot matrix analysis 255
- double cross-validation (dCV) 413, 416
- Dragon 427
- Dreiding model 101
- drug design 37, 106, 171, 199, 208, 315, 356, 362, 366, 379, 453, 458, 466, 527–529
- drug discovery 2, 7, 26, 31, 44, 103, 108, 164, 316, 453, 454, 460–461, 466, 479, 480, 484, 529
- drug target 200, 208, 209, 254, 316, 503
- dynamic programming algorithm 256–258
- e**
- eccentricity 364, 378
- EC code 361
- EclipseLink 222
- effective prediction domain (EPD) 476
- eigenvalue (EVA) 322, 323, 331, 336, 367, 368, 375
- eigenvectors 322–324, 331, 418
- electrocyclic reaction 141, 149
- electron affinity 272, 325, 336, 385
- electron density 37, 103, 106, 107, 126, 280, 323, 327–329, 332–336, 362–363
- electron diffraction 91, 369–372
- electron donating 145
- electronegativity 126, 127, 272–274, 276, 277, 280, 286, 287, 339, 360
- electronic journals 187, 219–220
- electronic laboratory notebooks (ELN) 137, 138, 156
- electrophilic aromatic substitution 124, 141
- electrostatics 37, 280, 282, 294, 304
- interactions 287, 288, 305, 307, 311
 - potential 107, 108, 321
 - shielding 307
- elimination reaction 124, 125, 141
- EMBASE 214
- eMolecules 190, 200
- EMPIRE 337
- empirical cumulative probability density function (ECDF) 401, 402
- empirical formulas 9, 12–14, 25, 234
- empirical models 399, 412–417
- enantiomers 32–36, 204, 210, 375–377, 387
- enCIFer* 67

- encyclopedias 221
 energy
 correlation 330
 force field 282, 303
 kinetic 304, 309, 311, 312, 332
 minimization 294, 517
 molecular orbital (MO) 302
 potential 95, 96, 99, 279–285, 295,
 301, 303–305, 308–312
 of protonation 385
 total 282, 304, 309, 311, 312, 329
 van der Waals 104
 zero-point 302
 Entrez Gene 207
 enumeration 17, 28, 74, 75
 environmental information 210
 enzymatic reactions 338, 369
 equilibrium constant 27, 130, 131
 EROS 142
 error back-propagation (EBP) ANN
 443–445
 essential set of essential rings (ESER)
 83
 eTOX database 222
 Euclidean distance 253, 355, 360, 370,
 409, 417, 420, 423, 429, 430, 432,
 477
 Euclidean metric 251
 Euclidean space 458
 European Bioinformatics Institute (EBI)
 196
 European Nucleotide Archive (ENA)
 197, 499
E-value 507, 508
 exchange-correlation (XC) energy 332
 excited state 301, 330, 331
 ExPASy 197, 207, 208, 513
 experimental design 169, 481
 explicit solvation model 307
 extended connectivities 77, 78, 164,
 364, 366
 extended Hückel theory (EHT) 323
- f**
- false negative 415, 514
 false positive 164, 415, 514
 Fast Fourier transformation (FFT) 175
 fast sequence searching 258
 FASTA 196, 258, 259, 503
 federated database system (FDBS) 222
 file formats 5, 36, 43–70, 72, 94, 109,
 134, 171, 172
 fingerprint 160, 163, 239, 247, 249,
 350, 352, 353
 biological 471
 circular 354, 364, 366
 extended connectivity fingerprints
 (ECP) 164, 364, 366
 hashed 240, 241, 250
 MACCS 353
 Fisher discriminant analysis 430
 Fisher weight 405
 FlexX 369
 flux balance analysis 150
FlyBase 207
 force field 94, 280, 292, 302, 304
 Allinger 280
 AMBER 283, 284, 294
 CFF 284, 289
 CGenFF 294
 CHARMM 283, 294
 GLYCAM 293
 GROMOS 283, 294
 MARTINI 290
 MM2, MM3, MM4 280, 292
 MMFF 284
 OPLS 295
 protein 294
 TINKER 293
 UFF 292
 Urey–Bradley 289
 Fourier coefficients 175
 free energy 280
 Langevin dipoles 307
 perturbation 316
 of solvation 307
 Friedel–Crafts alkylation 146
 frontier molecular orbital theory (FMO)
 128
 full structure search 161, 162,
 232–235, 251
 full-text databases 219–221
 fully integrated database systems 188

- functional groups 13, 14, 28, 124, 138, 145, 146, 163, 212, 240, 292, 295, 352, 469, 486
- g**
- gap penalty 506
- gas chromatography 204, 205
- Gaussian functions 323, 326, 328, 384
- Gaussian type orbital (GTO) 328
- GenBank 197, 254, 497, 499–501
- Generic Match Algorithm (GMA) 238
- general citation databases 218
- generalized Born (GB) method 306
- generalized gradient approximation (GGA) 333
- genetic algorithms (GA) 98, 301, 427, 445, 529
- genome 207, 497–499, 511
- GenomeQuest 196
- Gibbs free energy 131
- Gini index 432
- global minimum 95, 96, 301, 415
- global optimum 449
- Glycan Structures Database 208
- GlycomeDB 208
- glycomics 207
- Google Patents 214, 221
- Google Scholar 213, 218
- graph 15, 44, 79, 80, 150, 157, 232, 251, 252, 365
- isomorphism problem 74, 251
- graph theory 16–20
- GRid-INdependent descriptors (GRIND) 384
- grid search 96
- Groningen Molecular Simulation (GROMOS) 283, 294, 303
- Grubb's tests 470
- h**
- Hadamard coefficient 167
- Hamming distance 256, 356
- Hammond's postulate 246
- harmonic potential 282–285, 289, 290
- Hartree–Fock (HF) 159, 325
- hash algorithm 49
- heat capacity 269
- hash codes 87, 144, 160–162, 234, 235
- hashed fingerprints 240, 241, 250
- hashing 161, 162, 164, 234, 235, 240, 352, 353
- heat of formation 58, 269, 385
- heuristic algorithms 415
- HHsearch 516
- Hibernate 222
- hidden layers 443, 445, 454, 455, 457, 458
- hidden-Markov models (HMMs) 514
- hierarchical cluster analysis (HCA) 406, 420, 473
- HORACE 145–146
- highest occupied molecular orbital (HOMO) 107, 128, 331, 336, 385
- high-scoring segment pair (HSP) 503
- high-throughput screening (HTS) 372, 379, 471
- Hill order 194
- histograms 99, 100, 359, 362, 401, 402
- HomoloGen 197, 207
- homologous proteins 502, 503, 505, 509, 515–517
- homology modeling 515–518
- homology searching 195, 254
- Hooke's law 268, 283, 284
- hosts 187, 188, 190, 196, 210, 213–215, 218, 383
- Hückel matrix 274, 275, 322–324
- Hückel molecular orbital (HMO) 6, 274, 320–324
- human genome 497
- hybrid functionals 333
- hybridization states 92, 270–271, 275, 280, 369
- hydrogen bond donor 58, 251, 271, 360, 362
- hydrogen bonding 27, 108, 364, 379
- hydrogen motions 305
- hydrophobicity potential 37, 108, 364, 379
- hyperconjugation effect 276, 277
- hypersurface 95, 301, 309

i

iAMOEBA 290
 image recognition 454, 529
 implicit solvation models 307
 implicitly polarized Q (IPolQ) 293
 InChI 48, 50, 135, 191
 InChIKey 14, 48–50, 75, 135, 234
 incidence matrix 19–20
 index(ing)
 branching 365
 charge-related topological index (CTI) 234
 Jaccard 248
 topological 87
 induced subgraph 17
 inductive effect 127, 128, 131, 147, 272, 274, 276
 inductive learning 3–6, 124, 132, 268, 345, 349, 527
 InfoChem 143–144, 213
 infrared spectra 166–168, 175
 Inorganic Crystal Structure Database (ICSD) 198, 199
 INPADOC 215
 interaction
 charge–charge 286, 306
 Coulomb 286, 310
 dispersion 288
 electrostatic 286–288, 303, 304, 306, 310
 intermolecular 281
 van der Waals 286–289, 303, 333
 interchangeability classes (ICs) 83–85
 intermediate neglect of differential overlap (INDO) 326
 internal coordinates 36, 37, 45, 51, 52, 95, 97, 98, 159, 160, 289
 International Nucleotide Sequence Database Collaboration (INSDC) 197, 499
 International Pharmaceutical Abstracts (IPA) 214
 International Union of Crystallography (IUCr) 66
 International Union of Pure and Applied Chemistry (IUPAC) 11, 165
 InterPro 207

InterProScan 514
 interquartile range (IQR) 400, 403
 ionization potentials 30, 31, 272, 320, 324–326, 336, 337, 385
 iScienceSearch 192, 200
 isomers 13, 25–27, 31–33, 73–74, 86, 89, 188, 234
 isomorphism 73–74, 163, 235–239, 241, 243, 244, 251, 252
 isosurfaces 107–108
 isovalue-based electron density 106

j

Jaccard similarity coefficient 410
 Jacobi rotation 418
 Java Persistence API (JPA) 222
 JavaScript Molecular Editor (JSME) 71, 72, 135, 136
 JCAMP-CS 44, 45
 JCAMP-DX 44, 171–172
 JChem 70, 222, 484
 JME 72, 135
 JOELib 70
 Journal Impact Factor 218

k

KAGGLE datasets 459
 Karlin–Altschul parameters 508
 Kennard–Stone clustering 473
 kinetic energy 304, 309, 311, 312, 332
 k-means clustering 419, 421
k-nearest neighbors (kNN) 147, 407, 430–431, 467
 Knime 100, 484
 Kohonen artificial neural network 146–148, 360, 379, 445, 447, 448
 Kohonen map 147, 148, 170, 419
 Koopmans' theorem 336, 337
 K-rings 83
 Kyoto Encyclopedia of Genes and Genomes (KEGG) 199

l

labeled data records (LDR) 171
 Langevin equation 307
 Laplacian, of electron density 335
 LCAO-approximation 323, 329, 334

- Leapfrog algorithm 304
 - learning
 - algorithms 380, 453, 458, 529
 - cost-sensitive 474
 - inductive 3–6, 124, 132, 268, 345, 349, 527
 - methods 176, 346
 - rate 445, 449
 - leave-one-out (LOO) 415, 470
 - Lee–Richards surface 106
 - Lennard–Jones potential 287, 288, 295, 310, 383
 - Lexika 221
 - LHASA 86
 - library design 467
 - LigandBox 199
 - LigandScout 362
 - LigPrep 484
 - linear combination of atomic orbitals (LCAO) 321, 323, 327, 329, 334
 - linear discriminant analysis (LDA) 33, 412, 418, 429, 430
 - linear free energy relationships (LFER) 130–132, 306
 - linear latent variables 410–412, 417, 423, 424, 430
 - linear scaling methods 337–338
 - line notation 14, 44, 48, 87, 134, 232, 233, 243, 244
 - lipophilicity 364, 366
 - liquid chromatography 377
 - loading
 - matrix 400, 418
 - plot 412, 417, 422
 - local density approximation (LDA) 333
 - local ionization potential 336–337
 - localized molecular orbital (LMO) technique 338
 - local minimum 285, 301, 449
 - local optima 450, 455
 - DIALOG 187
 - lone pairs 30, 326, 328, 329, 338
 - long-range electrostatics 291, 302, 306, 310
 - long-range forces 303, 310–311
 - Lorentz–Berthelot 288
 - lowest unoccupied molecular orbitals (LUMO) 107, 128, 331, 335
- m**
- MACCS keys 353
 - machine learning 145, 158, 434, 454, 458, 459, 461, 488
 - macromolecular Crystallographic Information File (mmCIF) 65, 67–68, 70
 - macromolecules 44, 48, 58, 62, 67, 102–103, 107, 231, 283, 284, 307
 - Mahalanobis distance 407, 410
 - Manhattan distance 247, 355, 409
 - mapping 67, 74, 134–136, 138, 210, 211, 234–238, 241, 242, 252, 380, 381, 419, 439–441, 451
 - atom-to-atom 134–136, 139, 244, 254
 - Markush structures 27, 47, 192, 193, 215, 243
 - MARPAT 28, 189, 215
 - mass spectroscopy 166, 204, 205, 404, 421–423, 432–433
 - materials design 294, 458, 461
 - mathematical methods 346, 406, 408
 - matrix 18, 20, 21, 90, 140, 170, 253, 256, 367, 400, 406, 418, 430, 503–506, 509
 - adjacency 18–19
 - bond 21
 - bond–electron (BE) 21–23, 140
 - distance 20
 - hat 476
 - incidence 19
 - permutation 89
 - matrix reaction 140
 - maximal segment pair (MSP) 503
 - maximum common substructure (MCS) 242–243, 250–254
 - Maxwell distribution 304
 - mBed algorithm 511
 - MDL 44, 52, 53, 94, 134, 143, 163, 165, 353
 - mean absolute error (MAE) 473
 - mean-centered matrix 175, 404, 407, 418, 423, 424

- mean field approximation 325, 330, 331
- mean squared error (MSE) 413, 414
- median 402, 403, 408, 427, 428
- median absolute deviation (MAD) 400, 403, 408
- medical information 208–209
- Medical Subject Heading (MeSH) 213
- medicinal chemistry 208, 231, 232, 366, 468, 483
- Medline 189, 191, 206, 213, 214, 216
- membranes 291, 292, 315
- memory complexity 232
- Merck Index* 208, 221
- Merck Molecular Activity 459
- mesomeric effect 128, 274
- mesomerism 27, 193, 274
- metabolic flux analysis 150
- metabolic information 135, 208
- Metabolic Information Input System (METIS) 135, 136
- metabolic pathways 150, 208
- metadata 44, 65–69, 167
- meteorite 421–423, 432–433
- Metropolis Monte Carlo (MC) 98
- Michael addition 124, 146
- microcanonical ensemble 308
- MIMUMBA 98
- MINDO 326
- MineSoft PatBase* 215
- Minkowski distance* 409
- MLR 440, 450
- mmCIF File Format 67–68, 70
- MMFF 287, 293
- MM2/MM3/MM4 280, 286, 292
- MNDO-ESP 335
- modelability 467–471
- modeling 68, 345, 440
 - QSPR/QSAR 345, 346
 - quantitative 453
- MODELLER 515, 519
- model validation 467, 468, 472, 478, 480, 481, 517
- Modified Neglect of Diatomic Overlap (MNDO) 326, 327, 335
- Mol2 45
- molar refraction 269
- MolConverter 470
- molecular descriptors 36, 58, 79, 163, 164, 170, 249, 345, 349–351, 359, 363, 388, 406, 423, 454, 484
- molecular diameter 378
- molecular dynamics (MD) 98, 280, 302–317, 385
- molecular electrostatic potential (MEP) 108, 335, 378, 380
- molecular graph 70, 74, 79–82, 85, 86, 100–102, 365, 369
- molecular interaction fields 384
- molecular mechanics 280–292, 339
- molecular modeling 71, 100, 199, 293, 340, 481, 488
- Molecular Operating Environment (MOE) 470
- molecular orbital (MO) theory 28, 107, 270, 320–327
 - ab initio* 320, 327–333
 - energies 302
- molecular representation 5, 10, 11, 14, 157, 364
- molecular self-assembly 316
- molecular shape analysis 385
- molecular span 378
- molecular structure 6, 10, 23, 25, 28–29, 36, 44, 48, 51, 58, 69, 101, 103, 159–161, 232, 301, 321, 345, 351, 385, 527
- molecular surface 6, 37, 103–108, 336, 337, 378–382, 388
- molecular volume 336
- molecular weights 58, 155, 234, 269, 386
- molecularity 124, 537
- molecule editor 48, 57, 70–73, 135, 525
- molecule viewer 70, 71, 73
- Molfile 44, 52–58, 134
- MOLGEO 92
- Møller-Plesset (MP) theory 331
- MOLMAP 369
- Monte Carlo (MC)
 - search 302
 - simulation 98, 280, 296
- Morgan algorithm 48, 77, 87, 233, 366

- Morgan algorithm (*contd.*)
 stereochemically extended (SEMA)
 162
- Morse function 282, 283
- Mulliken population analysis 271, 334
- Multi-level Neighborhoods of Atoms
 (MNA) 358–359
- multilinear regression 440
- multi-variate data representation
 439–441
- multigraph 16
- multiple sequence alignment (MSA)
 509, 511
- multipole expansion 286
- multipole moments 287, 335
- multivariate calibration 406, 413, 423,
 429
- multivariate classification
 binary classification 429
 classification trees 429, 432
 discriminant PLS 430
 k-nearest neighbor 430
 linear discriminant analysis 430
 multiclassification 429
 multivariate similarity 429
 pattern recognition 429
 regression methods 429
 support vector machine 429, 431
- multivariate data
 cluster analysis 406
 linear latent variables 410–412
 matrix 406
 multivariate calibration 406
 multivariate classification 407
 object distance 408–410
 preprocessing 407–408
- multivariate data analysis 346,
 399–402
- mutation 255, 311, 504, 509, 515
- n**
- natural atomic orbital-point charge
 (NAO-PC) 335
- natural population analysis (NPA) 334,
 335
- Needleman and Wunsch 256
- neglect of diatomic differential overlap
 (NDDO) 326, 327
- net atomic charges 286, 334–336, 338
- neural network 166, 167, 360, 454–457
 architecture 455
 artificial 438
 back propagation 457
 counter-propagation 445
 deep 453, 454
 error-back propagation 443
 feed-forward 457, 458
 self-organizing 146, 360, 379, 381
 shallow 453, 454
- neuron 147, 442–450, 458
- Newton equation 303
- Newtonian mechanics 292
- NIPALS 418, 425
- NIST Data Gateway 203
- NMR chemical shifts 274
- NMR spectra 58, 87, 91, 198, 204–206,
 291, 304, 315, 369, 374, 386, 526
- nmrshiftdb2* 206
- noise 169, 426, 456
- nomenclature 11–12, 34, 73, 234, 324,
 328
- non-bonded electrostatic interactions
 303
- nonlinear mapping 417, 419, 422, 441
- non-polynomial (NP)-complete
 problem 235
- normalization 6, 48, 170, 222, 355,
 408, 484
- notation 12, 14, 73, 400
- nucleic acids 58, 62, 195, 200, 207,
 254–259, 293, 294
- Nucleic Acid Database (NDB) 199,
 200, 254–259
- nucleophilic aliphatic substitutions
 124, 128, 141, 148
- o**
- octet rule 23
- OEChem 69
- OELib 70
- ONIOM superimposition 338
- Online Public Access Catalogue (OPAC)
 217, 220

- ontology 156, 208
 - Open Babel 70, 109, 470
 - OpenEye 47, 69
 - optical character recognition (OCR) 220
 - optimization 159, 232, 236, 259, 432, 442, 450, 455–458, 515
 - algorithms 467
 - backtracking algorithm 238
 - geometry 302
 - global 294
 - Optimized Potentials for Liquid Simulations (OPLS) 293, 295, 303, 425
 - orbitals 30
 - coefficients 129, 275
 - electron densities 385
 - electronegativity 272
 - energies 337
 - OrChem 222
 - ORCID 216
 - ordinary least-squares (OLS) regression 407, 418, 423–425, 430
 - organometallic 31, 92, 212, 484
 - orthologs 502
 - outliers 174–175, 404, 470, 477
 - out-of-plane bending 285
 - overfitting 176, 412, 431, 432, 445, 450, 455, 457, 467, 475
 - oversampling 474
 - overtraining 176, 450, 451, 455
- p**
- parachor 269
 - PARAFAC 419
 - paralogs 502
 - parameterization 280–282, 284, 288, 291–295, 305, 327, 335
 - partial atomic charges 127, 271, 280, 292, 356, 362, 366, 368, 374, 376, 382
 - partial equalization of orbital electronegativities (PEOE) 127, 272–275
 - partial least-squares (PLS) 383, 385, 406, 423
 - discriminant analysis 430
 - regression (PLSR) 383, 406, 407, 425–426
 - particle mesh Ewald (PME) method 310, 311
 - partitioning 239, 385, 419
 - patents 196, 214, 220, 221
 - databases 28, 189, 214, 215, 219
 - Patent Chemistry Database (PCD) 189
 - path 17, 240
 - counts 250, 253
 - number 365
 - patterns 512
 - matching 75, 83, 85
 - recognition 345, 346, 351, 429
 - PatternHunter 259
 - Pauli principle 288
 - PDB file format 44, 58–65, 94
 - PDB2CIF 70
 - Pearson correlation coefficient 404, 407, 414, 424
 - Pearson product-moment correlation coefficient (PPMCC) 247, 404
 - percentiles* 403
 - Perception Reaction Center 138–139
 - periodic boundary conditions 308, 310, 311
 - permutation 35, 74, 88–90, 149
 - perturbation theory 130, 331
 - PFAM 207, 514
 - pharmaceutical and medical information 208–209
 - pharmaceutical excipients 209
 - pharmacophore 160, 251, 360, 362, 367, 385
 - screening 75
 - searching 251, 252
 - physical properties 35, 37, 187, 201–202, 209, 232, 308, 335, 386, 502, 509
 - physicochemical effect 126–128, 145, 368
 - physicochemical properties 68, 159, 185, 244, 249, 356, 368, 369, 376, 384, 387
 - π -charge distribution 274
 - PLS2 425

- PM3 327
 PM5-PM7 328
 PMML 171–172
 point accepted mutations (PAM) 256, 504, 505
 point charges 286, 307
 Poisson–Boltzmann equation 307
 Poisson equation 307
 polar surface area (PSA) 336, 366, 382
 polarity 126, 504
 polarizability 108, 128, 271, 275–277, 290, 335, 336, 360, 363, 368, 369
 polarizability effect 127, 128, 147, 271, 275–277
 polarization 274, 290, 291, 294, 308, 329, 337–339
 POL5-TZ 290
 polycyclic aromatic compounds (PAC) 427–429
 polymers 47–49, 191, 194, 204, 324, 385, 386, 527
 Pople diagram 329, 330, 332
 Pople-Pariser-Parr (PPP) 325
 population analysis 271, 334
 PostgreSQL 222
 potential energy 95, 96, 279, 282–284, 295, 301, 304, 308, 309, 311
 potential energy function (PEF) 99, 281, 303–305, 310, 311
 prediction 423, 440
 chemical reactivity 128
 data 176
 properties 345, 347
 Prediction of Activity Spectra of Substances (PASS) 358, 359
 Predictive Model Markup Language (PMML) 172–173
 predictive power 451, 459, 461, 467, 474, 478, 513, 514
 predictivity 450, 455–460, 473, 485
 pre-processing 6, 236, 239, 407–412, 438
 primary literature 61, 186–189, 191, 196, 206, 212
 principal component analysis (PCA) 375, 406, 407, 412, 417–419, 422–425, 430
 principal component regression (PCR) 423–425
 principal moments of inertia 378
 3D printing 103
 probabilistic coefficients 247
 probability 37, 98, 161, 323, 359, 405, 457, 504, 514
 probability density function (PDF) 401, 402
 problem-related transformation 408
 PROCHECK 517
 3D profiles 517
 projection 33, 103, 159, 379, 411, 412, 417, 419, 440, 442, 447
 ProQuest DIALOG 187, 208, 214
 PROSITE 207, 512–514
 Protein characteristics 509–514
 Protein Data Bank (PDB) 44, 58, 199, 200, 311, 313, 501
 protein families 7, 207, 499, 508, 509, 512, 514, 518
 protein folding 303, 316
 protein information resource (PIR) 196, 501
 protein-ligand docking 316
 protomers 470
 protonation 275, 276, 281, 385
 protonation states 281
 prototropy 26
 pruning 432, 457
 pseudo-isomerism 195
 PubChem 75, 191, 192, 208, 234, 353, 482, 483, 485, 486
 PubMed 191, 192, 213, 220, 500
 Python 69, 161, 222
- q**
quantiles 403
 quantitative structure-activity relationships (QSAR) 335, 345–347, 384, 406, 458, 465–488
 quantitative structure-property relationships (QSPR) 334, 345–347, 349, 406, 427, 465–488
 quantum chemistry 126, 159

- quantum mechanics 51, 107, 271, 280, 281, 286, 306, 320–340, 362, 527
- query sequence 505, 507–509
- r**
- racemates 36
- racemization 149, 305
- radial basis function artificial neural network (RBF ANN) 440
- radial distribution function (RDF) 134, 371–376, 387
- radicals 30, 108, 145, 194
- radius of gyration 378
- Randic connectivity index 234, 365–366
- random forests 429, 432, 458, 473, 475, 478
- randomization of response 475
- random weight drop out 457, 458
- RDfiles 52
- RDKit 109, 222, 484, 485
- reactions 68, 121–151
- biochemical 123, 148
 - center 125, 135, 138, 139
 - chemical 164
 - classification 139, 141, 143
 - editor 135
 - equation 122, 133
 - mechanism 125
 - networks 149
 - organic 90, 141, 142, 146, 212
 - oxidation 142
 - parallel 123, 133, 165, 537
 - path 149
 - photochemical 146, 149, 369
 - prediction 142
 - rate 131
 - representation 121
 - reduction 142
 - search 138, 244
 - site 138
 - SMILES 134
 - type 123, 145
- reaction database 133–138, 210–212
- reactive force field (ReaxFF) 290
- reactivity chemical 126–132
- reagent 122, 126, 128, 133, 164, 165, 191, 210–212
- real world data 173–175
- Reaxys 165, 187–191, 193, 194, 198, 200–206, 208–218, 220
- receptor 106, 251, 315, 362, 371, 374, 381–384, 468, 474, 528
- regioselectivity 130
- regression 424, 455, 476
- coefficients 423, 450, 455
- regularization 445, 455–458, 460
- relational database management system (RDBMS) 185, 186, 222
- relaxation 77, 78, 239, 241, 309, 310, 315
- relevance vector machines 455, 458
- relevant cycles 83, 84
- Representational State Transfer (REST) 71, 222
- repulsion 104, 288, 322, 325, 326, 330
- electron-electron 323
- resampling 415, 416
- residual 127, 274, 276, 413, 414, 423, 424, 476
- residual standard deviation (RSD) 476
- residue-based fluctuation analysis 313
- resonance 27, 29–30, 128, 244, 274
- effect 127, 132, 145
- RESP 335
- REST API 222
- retention indices 382, 427–428
- RGfiles 52
- ribosomes 107, 315
- RInChI (International Chemical Identifier for Reactions) 49, 135
- ring perception 16, 17, 79–86
- RNAs 197, 200
- robustness 450, 455
- Römpf Online* 221
- root mean square (RMS) 97, 100, 251, 310, 311, 313, 314
- root mean squared error (RMSE) 414, 448, 449, 473, 499
- ROSDAL (Representation of Organic Structures Description Arranged Linearly) 14, 233
- rotamers 316, 516

- rotatable bond 94–96, 99, 100, 363
- ROTATE 97, 98, 100
- rotational barrier 95, 302
- Remote Procedure Call (RPC) 71
- RXNfiles 52
- RXN format 134

- S**
- saccharomyces genome database 207
- safety information 206–207, 209–210
- Sammon's nonlinear mapping (NLM) 417, 419, 422, 423
- scalars 107, 400
- scaling methods 337, 407
- scatter plots 417, 418, 433
- Science Citation Index (SCI) 187, 213, 216, 218–220
- Science of Synthesis (SoS) 212, 219
- Scientific Technical Network International (STN) 188–190, 195, 196, 200, 208, 210, 215, 218, 221
- scientometrics 218
- SciFinder* 165, 188–189–191, 193, 194, 198, 201, 206, 211, 216, 217
- Scopus* 190, 213, 214, 216, 218, 220
- scoring 411
 - distance 410
 - function 517
 - matrix 255, 257, 400, 418, 503
 - plot 412
- screening 91, 162, 163, 239–241, 250, 272, 379, 467, 483, 487, 517, 518
- scree plot 418
- SDF Toolkit 70
- SDfiles 52, 53, 57–58, 69, 70, 94
- SEAL algorithm 384
- Search(ing) 231–260
 - 3D similarity 252
 - 3D structure 250
 - atom-by-atom 237
 - automorphism 241
 - bibliographic databases 216
 - chemical structures 231
 - depth-first 238
 - full structure 161, 232
 - fuzzy-substructure 244
 - homology 254
 - maximum common substructure 242
 - pharmacophore 251, 252
 - reaction 244
 - sequence 254, 258, 502
 - similarity 162, 245, 252–254, 350, 352
 - substructure 162, 243
 - superstructure 241
- secondary literature 187, 212
- secondary structure 58, 83, 103, 107, 314
- self-consistent-field (SCF) 325, 327, 329–331, 333
- self-organizing map (SOM) 147, 360, 361, 380, 419, 451
- self-organizing network 369, 379
- semiempirical 6, 51, 107, 130, 320, 324–327, 331–339
- sensitivity 130, 168, 258, 259, 415, 432, 471, 473, 503, 514
- sequence 195, 254, 497
 - alignment 254, 502
 - biopolymers 195–198
 - DNA/RNA 499
 - homology 502
 - protein 499
 - search 255–259
 - similarity 255, 502
- sequence database search 499–508
- sequence signature 512
- set of all rings 81–82
- SHAKE 305
- shallow neural network (BNN) 454–459
- Shannon entropy 174, 359
- Shannon equation 174
- shape indices 359
- similarity 148, 232, 245, 354, 369, 377, 408, 430, 468, 477, 485
 - coefficients 248
 - indices 363
 - matrix 254
 - measures 247–248, 250, 355
 - search 232, 245–250, 252–254, 350, 470

- Simple Object Access Protocol (SOAP) 71
- SIMPLS 425
- simulated annealing 98
- singlet 30
- singular value decomposition (SVD) 418
- Slater-type orbitals (STOs) 327, 328, 330
- smallest set of smallest rings (SSSR) 82–84
- SMARTS 47–48, 79, 82, 83, 134, 243, 244, 514
- SMILES 14, 44–48, 58, 71, 134–135, 233, 243, 485
- SMIRKS 47, 134–135, 244
- Smith and Waterman 258, 259
- solvation shell 291
- solvent-accessible surface 104–106, 307, 336
- solvent effect 276, 290, 305–308
- solvent-excluded surface 104
- space-filling model 102, 104
- spatial distance 370
- SPC 290, 308
- specificity 353, 415, 471, 473, 474
- SpecInfo 205
- spectra 137, 165, 170, 204, 386, 420, 439
 - calculation 332
 - experimental 170, 206
 - simulation 74
- Spectral Database for Organic Compounds (SDBS) 206
- spectral library 203, 247
- spectroscopic database systems 205
- sphere exclusion 472, 473
- spin density 108
- spin state 280
- SPRESI 211
- standard deviation 400, 402, 403, 405, 407, 408, 413, 418, 476, 477, 483
- standard error of calibration (SEC) 414, 424
- standard error of prediction (SEP) 412–414, 427, 428, 459
- standardizer 469, 484
- STAR (Self-defining Text Archive and Retrieval) file 65–67
- statistical data evaluation 402–405
- statistical mechanics 309
- statistics 218, 336, 346, 400, 407, 424, 450, 455, 467, 472, 475
- stereoaxes 31
- stereocenters 31, 33, 86–88, 485
- Stereochemical Extension of Morgan Algorithm (SEMA) 87, 162
- stereochemistry 14, 19, 31, 33–35, 47, 49, 56, 57, 87, 89–90, 148–149, 162, 485
- stereodescriptors 57, 88, 90
- stereoelectronic effect 127, 128, 538
- stereoisomerism 32, 47, 470
- stereoisomers 25, 33, 87, 88, 90, 192, 375, 470, 485
- stereoplanes 31
- steric effect 127, 128, 131
- Stevens–Johnson Syndrome (SJS) 479
- stoichiometry 13, 133, 139, 528
- STRAPS 47
- structural biology 315
- structural fingerprints 164, 232, 239
- structural function 305
- structural key 162–163, 239, 350, 352, 353, 359
- structural transition 316
- structure checker 469
- structure diagram 10, 15, 24, 27, 33, 70, 72, 100, 352, 358
- 3D structures 45, 51, 57, 60, 72, 73, 86–103, 198–200, 250–252, 361, 369–370, 375, 376, 386, 503, 515, 526
- Structured Query Language (SQL) 158, 196, 222
- structure elucidation 4, 44, 73, 91, 232, 243
- structure generator 55, 72, 73, 91, 92, 98, 361, 370
- structure isomers 73–74
- structure keys 160, 163, 164, 250, 356
- structure-searchable database 222, 223
- subgraphs 17, 18, 68, 244, 356

- substitution 124, 128, 141, 148, 195, 255, 256, 504, 505, 509
- substructure 6, 54, 68, 94, 162–164, 200, 201, 212, 232, 352
- substructure search 211, 235–245
- supergraph 17, 18
- supermolecule 303–307
- superstructure 6, 93, 94, 156, 241
- supervised method 429, 439, 440
- support vector machine (SVM) 429, 431–432, 455, 458, 467
- SureChEMBL 214
- Surfaces 106, 337
- SWISS-MODEL 515, 516
- Swiss-Prot 196, 501, 513
- SYBYL line notation (SLN) 48, 243
- symmetry group 74
- synthesis design 3, 44, 142, 147, 272
- synthetic methodology databases 212
- synthetic polymers 191, 194, 204
- System Development Corporation (SDC) 187

- t**
- Tanimoto coefficient 247, 248, 250, 253, 356, 410, 420, 485
- tautomerism 26–27, 43, 48, 74–76, 109, 188, 193, 291, 469, 470, 484
- tblastn 509
- T-COFFEE program 511
- temperature factor 63, 311, 314, 374
- tertiary literature 187, 219, 221
- test set 412, 461, 467, 468, 472, 473
- text mining 525
- thermodynamic and thermochemical data 202–204
- thermodynamical properties 283
- threshold moving 474
- Tikhonov regularization 456
- TINKER 293
- TIP3P 290, 295, 305, 308
- TIP4P-FQ 290
- toolkits 109–110, 222, 485
- topography 10, 25
- topology 10, 25, 74, 145, 162, 198, 232
 - autocorrelation vectors 368
 - index 87, 234
 - torsions 249
- torsion angles 100, 251, 285, 286
- total electrostatic free energy 306
- toxicity 69, 191, 209–210, 244, 466, 478, 487, 528
- toxicity, environmental, safety information 209, 210
- toxicology 2, 209, 356, 471, 480, 487, 529
- TOXNET 209
- ToxPrint 69, 244, 245, 356
- training set 412–416, 423, 430, 448, 461, 467, 468, 472, 475, 477, 479
- trajectory 106, 279, 309, 312, 313
- transformation 63, 134, 159, 168, 170, 244, 346, 351, 404, 408, 424, 438
- transition metal 324, 327, 332, 333
- transition state 126, 139, 289, 301
- TrEMBL 196, 501
- triplet 30, 251
- Tripos 48, 53, 384
- trivial name 10–12, 192, 234
- true positive 415
- t*-statistic 405

- u**
- Ugi's Scheme 140–143
- unambiguous 11, 25, 33, 47, 76–79, 162, 186, 235, 373, 502
- underfitting 412, 455
- undersampling 474
- UniCarb-KB* 208
- UniGene* 207
- UniParc 501
- UniProt (Universal Protein Resource) 196, 201, 499, 501
- UniProt Knowledgebase (UniProtKB) 501, 503, 507, 515
- Unique Ring Families (URFs) 85
- UniRef 501
- unit cell 61, 63, 200, 308
- united-atom force field
 - parameterizations 305
- universal approximation theorem 458–459
- universal force field (UFF) 292–293

- unsupervised learning 146, 147, 381, 417, 439, 440
- Urey-Bradley force fields 289
- user-defined weighting factors 253
- USMILES 47
- V**
- valence electron 21–23, 141, 246
- Valence Shell Electron Pair Repulsion (VSEPR) 93
- valence shell ionization energy (VSIE) 324
- valence states 143, 272
- validation 156, 415, 444, 449, 461, 471, 475, 481
- van der Waals 282, 304
 - interactions 288, 290, 304, 334
 - radii 37, 102, 104, 288, 306
 - surface 104
 - volume 104, 369
- variable selection 405, 424, 426–427, 467, 471, 550
- variance 158, 401, 403, 404, 407, 417, 418, 456
- vectors 400
 - Boolean 353
 - vector field 107, 108
- Verlet algorithm 304
- vertices 15–18, 20, 21, 74, 80–83
- VESPA 335
- vibration 63, 87, 289, 292, 303, 375
- vibrational circular dichroism (VCD) 87
- vibrational frequencies 292, 375
- vibrational spectra 289
- viewers 33, 70, 71, 73, 103
- virtual libraries 28, 110, 479
- virtual orbital 321, 331
- virtual reality 103
- virtual screening 91, 199, 315, 360, 467, 468, 479–482, 487, 518
- visualization 37, 43, 60, 100–103, 107, 108, 159, 199, 200, 218, 381, 401, 412, 418
- voice recognition 453, 454
- VolSurf 384
- volumetric properties 107, 108
- W**
- WABA 259
- walk 17, 236, 243
- Ward linkage 420
- water models 291, 306
- wavefunction 37, 51, 108, 280, 330–332
- wavelet transformation 175–176
- Web of Science* 190, 218, 219
- WHATCHECK 517
- Wiener index 365
- wire frame model 101
- Wiswesser line notation (WLN) 14
- Wittig–Horner reaction 145
- WOMBAT 481
- Worldwide Protein Data Bank (wwPDB) 58, 200
- WormBase* 207
- X**
- XDfiles 52
- XML 44, 52, 65, 66, 68, 70, 71, 135, 172, 244, 245, 356
- X-ray 198, 283
 - crystal diffraction 63
 - crystallography 304, 518
 - diffraction 58
- XSMILES 47
- XYZ-Format 50
- Y**
- Y-randomization 475
- Z**
- zero-point energy 302
- ZINC 198, 466
- Z-Matrix 37, 45, 51–52
- z-transformation* 408

Substance Index

a

acetylcholine 381
N-acetyl-*p*-aminophenol (APAP) 167
 adamantane 35
 α -conotoxin PNI1 polypeptide 60
 aniline 270, 274

b

benzene 27, 29, 95, 132, 163, 320, 331
 benzodiazepine 368
 benzoic acid ester 130
 2*R*-benzylsuccinate 95
 binaphthalene 87
 binaphthyl 35
 Boranes 31
 bovine pancreatic trypsin inhibitor
 (BPTI) 310
 1,3-butadiene 29, 129, 149, 322–324

c

cannabinoid receptor 460
 Carbohydrates 208, 291, 293
 (2*R*,3*E*)-4-chlorobut-3-en-2-ol 53
 chloromethane 36
 corticosteroid-binding globulin (CBG)
 receptor 371, 379
 cubane 82–85
 cyanoanilines 382
 cyclobutene 149

d

dipeptidyl peptidase 460
 dopamine 368

f

Ferrocene 31

g

GGG{2-10}R{5-10}R{5-10}RGAGA
 195
 glyceraldehyde 34

h

helicenes 32
 herbicides 375
 2-hydroxy-propionic acid 33, 34
 hypochlorous acid (ClOH) 76

i

isoleucine 305, 505

l

lactic acid 33–35
 lidocaine 192, 193
 lysozyme 102, 104

m

membranes 291, 292, 315
 metallocenes 195
 2-methoxy-1,3-butadiene 129
 4-methylcyclohexene 233
 methyl ester 130, 131
N-methylpyrrole 159, 168
 muscarinic receptor 381, 382

n

nicotinic receptor 381

nitroanilines 382
nitrophenol 123
nucleic acids 58, 62, 195, 200, 207,
254–259, 293, 294

o

orexin receptor 460

p

paracetamol 166, 190
paracyclophane 35

r

RNAs 197, 200

s

sulfamide 58
sulfuric diamide 58

t

tetracycline repressor 311, 313
thrombin 360, 460
triazine 375

