Index

"1.13.12.4" 112, 115

a

ab initio methods 76
ab initio modeling 55
abiotic 447
absorption 5, 135, 178, 240, 333, 359
acceptor 16, 68, 118, 178, 181, 190, 249, 263, 264, 269, 270, 279, 299, 301, 302, 318, 322, 347, 428, 453
  see also hydrogen-bond acceptors
accessibility 101, 182, 210, 285, 292, 367, 516
accessible surface area 294
ACD/MS Fragmenter 151, 375
ACD software 143, 156
ACD structure elucidator 154, 155
acetylcholinesterase inhibitors 418
Achilles project 198, 199
acid dissociation constant (p*K*a) 54, 73–76, 349, 456
acidity 11, 53, 456
activation energy 121
active analog approach 263
active pharmaceutical ingredients (APIs) 453, 573
active sites 179, 283–292, 298–304, 315, 317, 367
activity cliffs 320
ADABoost 26
adaptive neuro-fuzzy inference system (ANFIS) 350, 352
adaptive soft sensors 576–578, 582
additivity schemes 56–58, 62f, 70
adenosine monophosphate (AMP) 121–123
adipose 240
ADME (absorption, distribution, metabolism and excretion) 272, 333, 359, 448
ADME-Tox 166, 178ff, 272, 350–353, 544
adrenaline 261
ADRIANA-Code 338–347
adverse drug reactions (ADRs) 398
aem-thiolate proteins 366
Aerococcus viridans, 128
aflatoxin 445
Aggregated Computational Toxicology Resource (ACToR) 431
agonists 167, 225, 361, 375
agricultural research 417ff
agrochemical industry 6, 313
air-water contact angle 553
alanine 95, 128
albumin 349, 350
alcohol and aldehyde dehydrogenases 363
aldehydes 175, 363, 378, 511
aldose reductase-2 (ALR-2) 240
algorithmic complexity 268
alignment 24
  multiple sequence 285, 291, 298
  structure 178
alignment-independent 3D QSAR methods 24
aliphatic carboxylic acids 73, 75

Applied Chemoinformatics: Achievements and Future Opportunities, First Edition.
Edited by Thomas Engel and Johann Gasteiger.

Allen scheme 62
allosteric sites 284
allosterism 413
Almond 24
ALOGP 57, 65
ALOGPS 65f
alpha-hydroxy-isocaproate 126
alpha shapes 287
ALR2 inhibitors 240, 242
alternating least squares (ALS) 492
AM1, 66, 367
AMBER 319
Ames test 432
p-aminobenzenesulfonamides 260
β-amino-phenylpropanoicmimic 242
amorphous forms 55, 69–72
α-amylase 275
α-amylase inhibitor 276
anaerobic gut bacteria 362
analytical chemistry 7, 469–493
animal studies 188, 429, 441
anisotropy 144, 149
annotated genomes 123, 125
annotations 198
  active site 293
  binding site 284
antagonism 260
antagonist 167, 260, 275, 361, 369, 409, 413
anti-cancer effects 224
ant colony optimization (ACO) 316, 492
anti-inflammation 535, 538
anti-inflammatory compounds 221, 224, 227, 240, 373
anti-influenza 226
antibacterial treatment 125
antibiotics 170
antibody markers 565
anticonvulsant effect 514–516
antidepressant 517
antipsychotic agent 374
antitarget 220
anti-T2D compound database (ADB) 239
anti-T2D drug target 238
antiviral agents 272
anti-wrinkles/anti-ageing 538
apoprotein 428
applicability domain (AD) 32, 41–43, 46, 66, 334, 365
application programming interface (API) 449
APROPOS 287
AQUASOL 342
aqueous solubility (log *S*) *see* water solubility (log S)
archazolide A (ArcA) 224
ARChem-Route Designer, 99
Arctander atlas 512
arctigenin 226
area under the curve (AUC) 46, 326, 483
arene oxide formation 381
arginine (Arg98) 424
aromatic systems 57, 63, 136, 145, 184, 263, 265–266, 318, 374, 381, 419, 507
aromaticity 55, 252, 460
Artemisinin 115f
artificial intelligence 5, 134, 413
artificial neural network (ANN) 17f, 28f, 135, 138, 154, 227, 335, 338, 478, 492, 555
artificial neural network ensembles (ANNE) 350, 352
ASN.1 format 250
Aspirin 254
assay 209, 247
assay ID (AID) identifier 247
assay-to-lead attrition 178
ASSEMBLE 153
assertional metadata 397
assertion re-generation 397
associative neural networks (ASNN) 145
Asteraceae, 227f
atom-based contribution method 57, 60
atom-centered code 136
atom counts 250, 507–509
atomic electronegativity distance vector (VAED) 139
atomic properties 56, 135, 338

atomic resolution 110, 112
atom-to-atom matching 110, 252
attribute selection 124
attrition 166, 178, 183
autocorrelation 16, 173, 177, 186, 346
  autocorrelation vectors 16, 338
AutoDock 221, 315f
automated interpretation 135
automated synthesis design 102
automated text mining 398
automatic information extraction 77
automatic knowledge extraction 151
automatic recognition 150
autoscaling 472
average absolute error (AAE) 70, 337
Ayurveda 171

b

back-propagation 29
  error 492
  neural network (NN) 65f, 154
backward-elimination regression 21
bagged decision tree (BDT) 346, 347, 349, 480
bagging 26, 346, 480
baicalein 221
base learners 25f
baseline toxicity 185
basis set 476
Bayes' rule 577
Bayes' theorem 28
Bayesian methods 552, 554
Bayesian regularized artificial neural network (BRANN) 29
Beilstein Handbook 96
benchmarking 66, 210
Benson scheme 61
benzodiazepine agonists (BDA) 173, 174
benzoic acid 456, 457
benzyloxybenzene 418
beta-secretase 408
big data 77, 413, 495, 549
binary classification 45, 347, 397
bindability 292, 294
binding affinity 69, 181, 210, 265, 317, 319, 320
binding motif 323
binding pocket 179–181, 189, 221, 225, 226, 428
binding pose 209, 230, 316, 425, 428
Bingo 449
binned nearest neighbors (BNN) 483
bioactivation 361f
bioactive polymers 561–564
bioactivity 13, 198f, 211, 226, 361, 395f
bioassays 240
bioavailability 63, 209, 274, 293, 333, 336, 353
biochemical assay 418, 431
biochemical databases 375
biochemical on-chip assay 408
biochemical pathways 106–115, 566
biochemical reactions 85, 106, 110, 125, 171
biochemical synthesis 375
biocompatibility 551
BioCyc 111
biodegradation 533
biodistribution 413
biodiversity 208
bioinformatics 2, 108, 123, 125, 166, 168, 170, 195, 407, 527
biological activities 10, 11, 17, 18, 190, 408, 442–44
biological affinity 361
biological analogs 448
biological data 14, 20, 25, 166, 229, 553
biological testing of matching molecules 272
biology files 386
biology-oriented synthesis (BIOS) 209, 412
biomarker 128, 195, 199
biomolecules databases 505
BioPath.Database 109–128
BioSM 375
biotechnology 175, 246, 313
biotic 447
biotransformation 359, 361, 362, 368
BIOVIA Direct 449
BitterDB 505, 509
bitvectors 251f

BLAST 297f
BLASTP 126
bleaching 418
BLOCKS 298
blood–brain barrier (BBB) 334, 342–346
Boehm's function 320
Boltzmann's equation 419
bond additivity 62
bond angle strain 63
bond-type E-state descriptors 65
Boolean array 15
boosted trees 480, 487
boosting 26
bootstrapping 31, 37–38, 45, 294, 491
breast cancer drug 87
BRENDA 126
brewing 511
broker model 200
BSAlign 301
building blocks 240, 391, 409, 412

C

Caco-2 cell line 350
Caco-2 cell permeability 349, 350
CACTVS 387, 389–391
CADEX method 35, 36
calibration 11, 320, 486
Cambridge Structural Database 319, 390
CAMEO program 97
cancer 87, 196, 198, 199, 386, 445, 518, 556
cancer cell line encyclopedia (CCLE) 199
canonical variables 479
carbinolamine 378–381
carbocyclic coformycin 122, 123
carbohydrates 106, 505
carcinogenicity 336, 429, 431, 434, 435, 442, 443
Carcinogenic Potency Database (CPDBAS) 434
cardiotoxicity 350
Cartesian coordinates 135, 137
Case Ultra 430
CAS numbers 253
CASPER program 144
CAST 286, 287
catalysts 263, 278, 549, 554, 558
catalytic cycle 366
catalytic site 275, 283, 365
catechol 381
CATH 298
CavBase 299, 301
CCR8, 408
CDK 449
CDK2 complex 268, 270, 297
cDNA microarray analysis 536
cell
  adhesion 533, 535, 551
  communication 534
  extracellular matrix interactions 534
  migration 534
  proliferation 536
cellular and molecular processes 106, 536
cellular disease models 199
Center for Food Safety and Applied Nutrition (CFSAN) 443, 506
centering 19, 472
central metabolism 106, 108
central nervous system (CNS) 172, 342
ceramics 554
Cerius2, 347
cetirizine 73
Chamming 27
charge density 55
charge distribution 24, 248, 507
CHARGE program 144
ChEBI database 111, 112
CheMagic.org 388
Chematica 100–101
ChemAxon 400, 449
ChemBank 198
ChEMBL database 13, 196, 198, 395
ChemDraw/SymxyDraw 253
Chemical Abstract Services (CAS) 386
Chemical Abstracts 96, 255, 386
Chemical Activity Predictor (CAP) 391
chemical descriptors, *see* descriptors
chemical environment effect 135, 136, 141, 149, 293, 301, 338, 367

chemical identifier resolver (CIR) 387–388
chemical information 3, 134, 137, 168, 385, 506, 511
chemical properties 5, 10, 11, 53, 135, 300, 587
chemical reactors 572
chemical shifts 136, 141–149
chemical space 34, 95, 171, 292, 322, 334, 365, 462–464, 506–510, 549
Chemical Structure Lookup Service (CSLS) 388, 389
chemical subgraphs and reaction mark-up language (CSRML) 453, 457
chemical systems biology 125, 588
CHEMICS 153
ChemInform 99
ChemNavigator iResearch™, 391
ChemOffice 143, 144
chemogenomics 171, 199
chemome 239
chemometrics 2, 17, 471–496
chemoselectivity 92–93, 360
ChemoText 399, 400
ChemoTyper 58, 187, 188, 454
chemotypes 187, 239, 240–243, 445, 453–459
ChemScore 320
ChemSpider project 14, 256, 400
Chinese herbal medicine 172, 221, 237–243
CHIRON computer program 100
chlorpromazine 179
cholestasis 397
cholesterylester transfer protein (CETP) 223
chorismate 112, 114
chromatography 93
chromosome 316, 486
cinnamyl acid 240–243
circular fingerprints 459, 461
cityblock distance 42, 252
classification 18–20, 25, 28, 293, 333, 346, 400, 475
  activity 210
  blood brain barrier permeability 343
  models 11
  random forest 293
  supervised 479
classification and regression trees (CART) 480
click chemistry 90
clinical trials 188, 219, 549
clique detection 299, 301, 316
Cliquer 301
CLOGP 65, 347
cluster(ing) 17, 27, 36, 293, 397, 447
  analysis 150, 477
  genome 123–125, 129
  hierarchical 124, 270
  structure-based 430
cluster-based methods 35, 36
COCOA 154
CODESSA 555
coenzyme Q10, 486
co-factors 121, 266, 360, 363, 364, 427
cohesive energy 558
collision-induced dissociation (CID) 151
column distillation 579–581
combinatorial chemistry 53, 134, 141, 166, 175
combinatorial libraries 172, 175, 183, 513
COMBINE program 153
CoMFA 22–25, 181
comparative molecular field analysis (CoMFA) 22, 365, 420, 421
comparative molecular moment analysis (CoMMA) 24
comparative molecular similarity indices analysis (CoMSIA) 24, 181, 420, 421
COMPASS 24
compatible solutes 528, 529, 533
compound databases 13, 225, 239, 246
compound identifier (CID) 246, 249
compound libraries 153, 171, 182, 202, 219, 272, 313, 423, 506 *see* libraries
compound-mediated toxicities 199
computational chemistry 3, 7, 150, 208, 334, 440, 588

computational tools 86, 102, 208, 211, 220, 228, 229
computer-aided drug design (CADD) 259, 385–391
computer-aided pharmacophore modeling and screening 278
computer-aided structure elucidation (CASE) 153–156
computer-aided synthesis design (CASD), *see* computer-assisted synthesis design
Computer Assisted Evaluation of industrial chemical Substances According to Regulation (CAESAR) 435
computer-assisted molecular design 406
computer-assisted structure elucidation (CASE) 3, 133–157, 587
computer-assisted substance identification 139
computer-assisted synthesis design (CASD) system 5, 84, 94, 96, 100, 101, 587
computer-assisted synthesis planning 409
computer automated structure evaluation (CASE), MCASE 430
computer graphics 99
ConCavity 288
conductivity 551–558
conductor-like screening model (COSMO) 55
conformational ensemble methods 316–317
conformational flexibility 137, 269, 316, 533
conformational search, systematic 419
conformational space 24, 181, 365, 367
conformations 16, 55, 63, 211, 243, 249, 269–272, 291, 315–317, 419, 423, 428, 529
conformer generation 316
confusion matrix 45
CONGEN 153
congeneric series of compounds 20, 58, 64
connection tables 89, 110, 137
connectivity map concept 200
Connolly surface 287, 299
consensus model 72, 433, 434
consensus scoring 288, 320
conservation 208, 285, 287, 364, 503
constraints 95, 153, 267–268, 323, 375, 425
contiguous blocks 489
continuum solvation model for real solvents (COSMO-RS) 55
contour maps 421
CORINA 47, 110, 112, 121, 122, 137, 138, 172, 339, 389, 464
corosolic acid 222
correction factors 57, 64, 65
correlation 135, 337, 552
  analysis 335, 485
correlation coefficient 20, 335, 421
concordance (CCC) 40
  Kendall 485
  Pearson 321, 485
  Spearman 485
correlative structure-activity relationship methods 429
cosmetics 7, 350, 441, 452, 525–541, 586
COSMOS 445, 449, 464
COSMOS DB 449
COSY spectra 154
cross-validation 30, 31, 36–37, 44, 186, 335, 398, 400, 421, 433, 485, 491, 555
crystals 69, 71, 320
crystallization 72, 93
CSI-FingerID 153
CTD2 project 198, 199
curation 9, 13–14, 201
curcumin 242, 513
custom development project 253
cyclooxygenase-2 (COX-2) inhibitor 373
cyclopeptides 533–541
cytochrome , 360, 363–367

d

DARC-EPIOS 153
data analysis 17, 18, 150, 472
data augmentation 249
data avalanche 196
databases 34, 96, 99, 100, 142, 144, 151, 386, 391
  BRENDA 126
  ChEMBL 395
  chemical shift 142
  ConSurf-HSSP 287
  drug 171
external 178, 247, 388, 392
  Food additives 505
Fragrance and Flavor 505
  HDB 238
  KEGG REACTION 111
  mining 200, 272
  monitoring 578
  Natural products 215
  PAFA 452
  NCI 386
  PEDANT 123
  Pocketome 297
  PubChem 175, 245, 247
  reaction 91, 96–97, 110
  Regulatory 505
  searching 142, 143
  SIDER 401
  storage 249
  toxicity 448
  toxicological 431
  VigiBase 398
data compression 17
data curation 13, 14
data cycle 189
data-driven 195
data mining 17, 134, 166, 175, 443, 464, 513–521
data modelling 20, 449–453
  methods, non-linear models *see* non-linear models
  regression-based QSAR approaches *see* regression, analysis
  3D QSAR *see* 3D QSAR
data reduction 372, 474
data silos 201
data splitting 32
dataset 21, 26, 32–34, 71, 119, 135, 137, 181, 183, 334, 338, 489, 512, 549, 552
DataWarrior 520
DAVID 199
Daylight DayCart 449
N-dealkylation biotransformations 378
deamination of AMP 122, 123
death-associated protein kinase 3 (DAPK3) 409, 411
decision tree (DT) 18, 25–27, 138, 147, 227, 335, 347, 444, 552
decoys 226, 273
Deductive Estimation of Risk from Existing Knowledge (DEREK) 430, 432
deep learning 200, 396, 495
degrees of decomposition 55
dehydrogenase 363
Delaunay triangulation 287
delocalization stabilization of charge 119
O-demethylation 376, 379
denaturation 529
DENDRAL Project 5, 134, 153
dendrogram 477
denitration 578
de novo design 101, 169, 179, 182, 276, 407–409, 413, 423
density functional theory (DFT) 55, 142, 367
dereplication 209, 219
descriptors 14, 19, 20, 58, 59, 224, 293, 317, 335, 443
  1D 15
  2D 71
  3D 16, 137
  3D Zernike 300
  binding site 284
  chemical descriptors 13, 119, 121, 135, 144, 347, 365, 400, 555
  Dragon 434
  fragment-based 135
  geometric 16–17, 144
  GRIND 16

descriptors (*contd.*)
herbal prescription 238
mathematical 551
molecular, *see also* molecular descriptors
pharmacophore 365
physicochemical 119, 121, 144, 347, 365
quantum-chemical 555
selection 552
shape 346
spectral property 150
topological 16, 144, 338, 555
uninformative 551
whole-molecule 65
design of experiments (DOI) 553
detoxification 361
Developmental Therapeutics Program (DTP) 386, 387
diabetes 222, 225, 237 *see also* type 2 diabetes
dibenzofuran 445
dibenzo-p-dioxin 445
Dice 252
2,6-dichlorophenol 187
dietary supplements 486, 516
Difference of Gaussian (DoG) 286
diffusion 167, 343, 347, 533, 554
dihydrofolate reductase enzyme 260, 261
para-dihydroquinone 379
dimension reduction 476, 477
dimethylsulfide 128
DIOS database 222
1,2-diphenylethane 418
dipole and quadrupole moments 24, 54, 346, 550, 558
discriminant analysis 294, 480
discrimination functions 19, 227, 293
disruptive technology 548
dissolving process 69, 336
distance-based methods 42
distance, Manhattan 487
distribution 5, 14, 24, 30, 34, 42, 170, 185, 186, 274, 333, 349, 359, 507
diverse libraries 178
diversity 67, 506
Django 449
DNA methyltransferase (DNMT) 509, 518–521
DOCK 221, 230, 276, 315, 327
docking 16, 168, 179, 221–225, 240, 274, 287, 296, 315–317, 367, 406, 423
DoGSite 286
DoGSiteScorer 293, 294
Dolabriferol 90
domain of applicability 336, 552
domestic substance list (DSL) 441, 452
donor 16, 68, 178, 249, 263, 264, 268, 279, 299–301, 318, 347, 430, 507, 519
dopamine agonists (DAA) 173, 174
D-optimal design 23, 36
DRAGON software 400, 432, 516, 555
drug approval 167, 190
DrugBank 277, 400, 508, 514
drug design 165–168, 179, 276, 385–391, 406–13 *see* computer-aided drug design
drug discovery 5, 165–190, 333, 371, 405, 586
applications 501
data-driven 201
indication expansion 200
target mining 196
toxicity prediction 199
druggability 198, 284, 288, 292–296
druggability dataset (DD) 294
drug2gene 196, 198
Drug Information System (DIS) 386
drug like density (DLID) 294
drug likeness 171, 178, 274, 387, 521, 522
drug metabolism 360, 361
drug metabolism and pharmacokinetics (DMPK) 371–372
drug-metabolizing enzymes 360, 363
drug permeability 350
DrugPred 293, 294
drug-receptor interactions 261, 262
drug targets 13, 170, 171, 196, 198
DrugScore 319
DrugSite 287, 288

DUPLEX algorithm 36
dynophores 276–278

e

EC number 108, 111, 112, 117
ecotoxicology 184, 429
EC-PDB databases 113
ectoine 528–533
edge of chaos 414
EDIA 321
effective core potentials (ECPs) 142
Eigenvectors algorithm 490
Elaboration of Reactions for Organic Synthesis (EROS) 97–98, 139
electron density 24, 321
electron impact (EI) mass spectra 150, 151
electronegativity 119, 144, 173, 346, 368
electronic effects 23, 119, 342, 346, 456
electronic health records 201
electronic laboratory notebooks 102
electrophiles 186
electrophoresis, single cell gel 474
electrospray soft ionization (ESI) 151
electrostatic field 22, 421
electrostatics 23, 24, 319
  potential 528, 533
electrotopological state (E-states) 430
embryonic stem cells (ESCs) 562
empirical scoring functions 318–320, 423
enantiomers 248, 261
enantioselectivity 93
encoding 98, 143, 147, 248–250, 302, 566
endocrine-disrupting chemicals (EDCs) 429, 544
endogenous metabolism 106, 171
endoplasmic reticulum 367
englerin A 412
enrichment 199, 269, 272, 324, 536
enrichment factor (EF) 274, 324
Ensembl 201
ensemble docking (ED) 317
ensemble learning 25, 26, 576
enthalpy (ΔHf°) 60, 181, 318, 321
entity-attribute-value model 450
Entrez 246, 250, 255
entropy (S°) 60, 179, 181, 265, 318, 320, 326
enumeration 134, 243
environmental fate 429, 447
enzymes 11, 106, 108, 111, 118–124, 167, 284, 367
epalrestat 241, 242
(–)-epigallocatechin gallate (EGCG) 518
epigenetics 516–521
epoxide hydrolase (EH) 363, 365
equilibrium solubility 69
Eribulin/Halaven 87
error back-propagation algorithm 29
E-state indices 71, 338
esterases 363
estradiol 261
estrogenic agent 261
eTOX project 189, 200, 449, 452
Euclidian distance 27, 36, 224, 252, 464, 473, 486
European Chemical Agency (ECHA) 447, 448
European Food Safety Authority (EFSA) 444
EU Scientific Committee of Consumer Safety (SCCS) 444
Everything Added to Food in the United States (EAFUS) databases 506, 510, 520, 522
evolution 13, 21, 208, 503
evolutionary algorithms 29, 101
exclusion volumes 267, 268, 270
excretion 184, 333, 349, 359, 364
expectation maximization (EM) 489
expert systems (ES) 134, 135, 138, 154, 368, 429, 431
EXPIRS 138
explicit water 321, 428
extensible markup language (XML) 250
extracellular matrix (ECM) 535, 538, 539

f
factor analysis (FA) 19
  factorial design 559
false negatives (FN), positives (FP) 45, 273, 324, 343, 370, 483
Farnesyl-diphosphate 115, 116
FAst MEtabolizer (FAME) software 368
Fasudil 409, 411
FATCAT 298
feature bitvector 251
feature extraction 27
feature selection 347, 552, 557, 565, 566
federated database system (FDBS) 449
fibroblasts 535, 536, 540, 541
file formats 248, 253, 388, 390
filtering 178, 211, 220, 274, 322, 334, 485
FINDSITE 287
fingerprints 15, 147, 176, 251, 299, 369, 454, 459, 507
  bitvectors 251
chemical 153
dynamic 459–461
extended connectivity 459, 461, 507, 508
  MACCS 400
  pharmacophore-based 300
  radial 508
first-in-class drugs 195, 196
first order approximation 56, 61
Fisher criterion (F value) 485
fitness function 271, 316, 411, 486
fitness landscapes 407, 409, 411
FITTED 317
fitting 552
flagging 58
FLAP 300
flattened protoporphyrin IX 428
flavin adenine dinucleotide (FAD) 121, 424, 427, 517
flavin monooxygenases (FMO) 363, 365
flavonoids 240
flavor cliffs 511–512, 522
flavor compounds 125, 126, 128
flavor-forming pathways 125–128
Flavornet 506, 509
FlexX 315, 316, 317, 424
FlexX-Pharm 425
Flory Huggins parameters 557
FlowerPower 298
food additives 444, 502, 505
food analysis 492
food-contact substances 443
Foodinformatics 502
food-related chemogenomic space 513, 521
food science 501–522
force field 16, 137, 319, 418, 428
  HINT interaction 320
  OPLS-2005, 530
force field based scoring functions 319
forensic toxicology 372, 374–375
forest model 368
formal reaction generators 97
forward search 97
forward-selection regression 21
forward synthetic route 99
Fpocket 287, 293
fractional factorial design 23
fragment 15, 25, 58, 135, 138, 263, 335
  constant approach 64, 65
  fragmentation 316, 343
    approaches 58
    MS patterns 375
fragment-based *de novo* design 276
Fragment Reduced to an Environment that is Limited (FREL) 135
FRED 315, 316
free binding energy 320, 424
free energy 11, 14, 23, 58, 91, 110, 181, 265, 317, 321
free energy perturbation (FEP) 262, 321
Free-Wilson method 11
frequent hitters 58
Fructus Lycii (Gouqizi) 238, 240
Fructus Schisandrae (Wuweizi) 238, 240
FSSP/Dali 298
Fullerenes 554, 555

functional groups 23, 54, 57, 65, 91, 135, 143, 266, 337, 362–364, 369, 560
FuzCav 300
fuzzy data 154
fuzzy logic 138

g

Gaucher disease 254
Gauge-Invariant Atomic Orbital (GIAO) method 142
Gaussian 142
  function 301
  processes 59, 432
Gaussian radial basis function 481
GEN 154
GenBank 246
gene expression 199, 535, 564
GeneGo cellular processes 536
Genentech dataset 199
Generalized Born Surface Area (GBSA) 181, 320
generally recognized as safe (GRAS) 452, 502
generative topographic mapping (GTM) algorithm 509
gene set enrichment analysis 199
gene signatures 200
genetic algorithm (GA) 17, 18, 21, 123, 154, 175, 179, 316, 335, 347, 486, 555, 557, 572
genetic algorithm-based process variables and dynamics selection (GAVDS) 572
genetic algorithm for multiple molecule alignment (GAMMA) 123, 179
genetic function approximation (GFA) 47, 347, 353
genetic linkage 198
genetic programming 540, 557
GENIUS 154
GENOA 153
genome clustering 123–125, 129
genomic(s) 106, 166, 199, 304, 407
genotoxicity 443, 474–475
geometrical features 144, 560
geometric methods 42
GERM 24
ghecom 288
Gibbs free energy (ΔG) 317
Gibbs–Helmholtz 318
GIF Creator 389–390
glass transition temperature 553, 554, 557, 558
Glide 315, 316, 426
glucose 222, 240
glucuronidation 363, 376, 379
glutamate 112, 113, 124, 180, 516
glutamate racemase (MurI) 180–181
glutathione conjugation 363, 364
glycans 143, 144
glycation mechanisms 240
glycine 128, 533, 534
glycoprotein 350, 540
glycoside 240, 242, 243, 509
Glycyrrhiza glabra L. (Fabaceae) 226
GlyNest program 143
GoFigure 298
GOLD 315, 316, 327
GoldScore 319
GOLPE 23
goodness of prediction 39–41, 421
G-protein coupled receptor (GPCR) 272, 409, 423
gradient boosted trees (GBTs) 346, 347, 349
gram-positive 124
GraphicsMagick 389
graph theory 14, 16, 137, 173, 251, 299, 301, 421, 460
GRAS compounds 506, 508, 510, 513, 517, 520
green chemistry 71, 413
GRID 17, 23, 24, 269
grid-based algorithm 285
GRIND descriptors 16
group additivity 61–63
group contribution methods 63, 337–338 *see also* group additivity
guanidine 240, 266
GUSAR 391

h

halophilic organisms 528, 529
Hammett equation 456
Hamming distance 172
Hartree–Fock 142
hash codes 249, 250, 252, 369
hash-coding algorithm 15
hash function 299
HazardExpert 431
HDon_O 347, 349
heat capacity (Cp°) 60, 61, 558
heat of adsorption 560
heats of atomization (ΔHa) 62
heats of formation (ΔHf°) 60–63
heats of reactions 98
Helicobacter pylori, 180
hemiacetal 97
hemoproteins 363
Henry's Law, coefficient 566
hepatic metabolism 376
hepatotoxicity 376, 378, 379, 397, 398, 431
herb-chemome-MOA network (HCMN) 239, 243
herbs 239–242
heterocyclic systems 63
heuristics 97, 99, 153, 301, 316, 388, 406, 488
hierarchical cluster analysis (HCA) 119, 124, 227, 270, 477
Hierarchical Ordered description of the Substructure Environment (HOSE) 136, 143, 144, 147
hierarchical trees 73
highest occupied molecular orbital (HOMO) 16, 550
high production volume (HPV) 446, 452
HINT interaction force field 320
histidine degradation 125
histogram, 300, 301, 324, 325, 456, 459
histone acetyltransferase p300, 224
histone deacetylases (HDACs) 517–518
historical Hungarian coins 480–483
HiT QSAR 14
hit rate 172, 229, 274, 292, 293, 325, 326, 412, 459
HIV inhibitors 101
HIV-1 protease 297
HMBD 375
holoprotein 428
homology 143, 168, 261, 284, 291, 313, 315
HOSE code 136, 143, 144, 147
host-microbial co-metabolism 128
hot spot diagram 287, 368, 371
HSQC and HMBC spectra 154
Hudlicky's analysis of syntheses 96
HumanCyc 375
human genome 106, 170
human health 7, 429, 431, 442, 517, 586, 588
human intestinal absorption (HIA) 334, 346–350
human oral bioavailability (HOBA) 349, 353
human plasma protein binding (PPB) 343, 349, 350, 353
human serum albumin (HSA) 349, 350, 353
human skin permeability 349, 350
hybridization 60
hydantoins 175–177
HYDE 320
hydrogen-bond acceptors 16, 270, 302, 338, 339, 347, 428, 430, 464, 519
hydrogen-bonding 176, 177, 261, 264, 318, 321, 343, 349
hydrogen-donor feature 270
hydrogen encoding 248
hydrolases 117–119, 121, 363
hydrophobic contacts 226, 264–265, 320, 424
hydrophobicity 24, 176, 184, 264, 293, 294, 296, 318–320, 336, 346, 350, 353
11β-hydroxysteroid dehydrogenase 1 (11β-HSD1) inhibitors 222, 223
HyperChem 142
hyperplane 19, 29
hyperstructure 154
hypothetical activity models 166

i

IC_{50}, 221, 222, 225, 240, 241, 247, 374, 429, 535
ICH M7, 441, 444
ICM 315
ICSynth 99–100
IGOR 97
ill-conditioned problem 476
impact resistance 557
impurities 72, 441, 444, 445
InChI 91, 112, 249, 253, 387, 388, 450
InChI (Chemical Identifier) 89, 91, 112, 249, 253, 387, 388, 450
InChIKey 112, 387, 388
increment-based methods 143–144
incubation 373, 376, 378, 536
index
  database monitoring (DMI) 579
  E-state 71, 338
  Tanimoto 112
  topological 16, 71, 338, 430, 558
indication expansion 195, 196, 200–201
indoleamine 2,3-dioxygenase 226
indomethacin 373
indoyl 240
induced fit 317, 426, 428
induced fit docking (IFD) 317, 426
inductive effect 75
inductive learning 84
inflammation 200, 274, 275, 535, 536, 538, 540, 541, 544
inflammatory bowel disease (IBD) 200
InfoChem 100
infrared (IR) spectrum 135, 137–140
inhibitor 101, 121–123, 129, 167, 222, 226, 242, 284, 353, 420, 426, 510, 517, 519, 520
inner salts 67
Innovative Medicine Initiative (IMI) eTOX project 451
innovative syntheses 96
inosine monophosphate (IMP) 121–123
in silico design and data analysis (ISIDA) 58, 400
in silico inspired synthesis 97
in silico toxicology 417, 429, 432–435
insulin 240
integrins 534, 535, 539, 540, 542
interactions 318
  cation-π, 265, 266
  cis-trans isomer 136
  drug-drug 365
  energies 55
ionic 266, 423
  ligand-receptor 264
  π-π, 318
  protein-ligand 317
  protein-protein 189
interactive computer graphics 98–99
intercellular adhesion molecule-1 (ICAM-1) 530
internal validation 31–33, 36, 43–45
International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH M7) 441, 444
interpolations 41, 143, 582
interval PLS (iPLS) 486
intestinal absorption 333, 334, 346–347, 353
IntOGen 198
intralaboratory reproducibility 55
intrinsic solubility 69
inverse Boltzmann technique 423
inverse soft sensor-based feed forward (ISFF) control method 582
in vivo animal studies 441
in vivo/In Silico Metabolites Database (IIMDB) 375
in vitro tests 184, 459, 535
ion channel 167, 423
ionic liquids 71, 554
ionization 68, 73–76, 151, 456
Iris dataset 35, 36
ISAC approach 320
ISIDA/Duplicates 14
IsoCleft 299
isomerases 117
isomer generator software 153
isomeric hydrocarbons 61
isostere 260

iso-surface 269
isotope 249, 250, 252
IspD 428
iterative stochastic elimination 181

j
jackknifing 485, 489
JavaScript 253
JChem 449
Joint FAO/WHO Expert Committee on Food Additives (JECFA) 444, 505, 506, 516
JSME 112, 389
Just Exploring Druggability at protein Interfaces (JEDI) 296

k
kaempferol 277–278
KEGG 112, 375
Kennard-Stone algorithm 35–36
keratinocytes 530, 535, 536, 541
kernel function 29, 481
k-fold cross-validation 36, 37, 335
kinase-inhibitors 220, 409
kinetic solubility 69
k-nearest neighbor (kNN) 27, 138, 145, 147, 335, 432, 482, 577
knowledge-based expert systems 368–369, 431
knowledge-based scoring functions 319, 423
knowledge extraction 84, 151
Kohonen maps 173, 174, 478
Kohonen neural network 17, 19–20, 119, 138, 173
kosmotropic 533
Kulczynski 252
Kyoto Encyclopedia of Genes and Genomes 111

l
lab-on-a-chip 408
lactate-2-monooxygenase 126
lactic acid bacteria (LAB) 126, 128
Laidler scheme 62
Langerhans cells 529, 530
latent TGF-β binding protein (LTB,P1) 539
latent variables 19, 22, 476, 487
lattice energy 70, 336
LC50-value 184
LD50-value 184
lead compounds 178, 183, 313, 333, 371, 435, 527, 542
lead finding 167, 195
  ligand-based drug design (LBD) 171–175
  structure-based drug design (SBD) 179–182
lead hopping 172, 174, 183
lead identification 183, 417, 423
LeadIT 316
lead optimization 167, 182, 195, 417
  ADMET properties 183
  toxicity 184–187
leadscope 430, 431
lead structure 167, 172–175, 182, 190, 272, 279, 283, 418, 588
learning 11 *see also* classification methods; non-linear models; regression analysis; 3D QSAR
  statistical 29
  supervised 18, 28, 411, 492
  unsupervised 17, 173, 492
least absolute shrinkage and selection operator (LASSO) 486–488, 572
leave-many-out (LMO) 491
leave-one-out (LOA) 37, 181, 421, 433, 489, 491
legacy encoding 248
Lennard–Jones potential 24, 319
leoligin 223
Leontopodium alpinum, 223
leucine 126
L-glutamate 108, 109, 112, 180
LHASA 98–100, 391
L-histidine 125
libraries
combinatorial 172, 175, 183, 513
  corporate compound 171
  design 195, 337, 338
  diverse 178

focused 172, 178, 240
screening 209, 240, 241, 513
LIBRARY mode 66, 68
lichen 208, 221
licorice 226
life cycle assessment 96
ligandability 292, 294, 296
ligand-based approach 270, 273, 409, 419–422, 436
ligand-based drug design (LBDD) 169–179
ligand conformers 270, 272
LigandScout 263, 268–270, 275, 276, 278
ligand shape 267–268
ligand/target binding mechanism 259
ligases 117, 121
lignan 223, 226
LIGSITE 285, 287
linear discriminant analysis (LDA) 19, 138, 227, 479
linear free energy relationship (LFER) 11, 20, 58, 59, 64, 65, 73, 184
linear regression 20, 59, 65, 71, 293, 433, 487
Lipinski's rule 94, 210, 294, 347
lipophilic contacts 264, 265
lipophilicity 53, 63, 64, 178, 184, 342, 361, 365, 367, 369, 508
lipoprotein 350
liquid chromatography–mass spectrometry (LC-MS) 151, 227, 441, 448
L-lactate oxidase (LOX) 126, 128
loading vectors 476, 485
logical rules 138, 397
logistic regression (LR) 21
loquat 222
lowest unoccupied molecular orbital (LUMO) 16, 550, 558
LSD 154, 156
LUDI 269, 270, 423
lyases 117

m

MACCS key 507–510, 517
machine learning 26, 29, 30, 59, 71, 76, 77, 99, 144, 149, 157, 227, 495
macroconstants 73
magic constants 64
magnetic field 142
magnolol 225
Mahalanobis distance 42
maitotoxin 87, 102
majority voting 26
Mannich bases 139
mapping methods *see* classification, methods
Mass Frontier 151, 152, 156
mass spectrometry (MS) 150, 151
mastic gum 225
matairesinol 226
matching 297
atom-by-atom 110, 252
materials science 553–566
matrix,
confusion matrix 45
decomposition 476
loading 19
metalloproteinases (MMPs) 541, 542
Matthews correlation coefficient (MCC) 45
Maximal Affinity Predicted for Passively absorbed Oral Drug (MAPPOD) 293
maximal electroshock seizure test 515
maximum common subgraph 299, 301
mean absolute error (MAE) 31, 41, 145, 147
mean centering 19, 472
mean molecular polarizability 59–60, 338, 339
mechanism of action (MOAs) 185, 186, 238, 239, 242, 418
mechanistic analysis 97
medicinal chemistry 166, 219, 259, 260, 303, 365, 371, 373, 395, 406–409, 413
MEDLINE 397
melting point (MP) 54, 55, 70–72, 337
membrane bioreactor (MBR) 573
membrane partition coefficients 442, 456

membrane transport 11, 336
MeSH 246, 400, 401
Mestrelab 147
MetabolExpert 368
metabolic pathways prediction 125–128
metabolics tree 369, 372
metabolism 11, 106, 333, 359, 376, 413
metabolites 110, 117, 125, 145, 153, 208, 219, 260, 365, 370, 371, 374, 375, 378, 444
metabolome 360, 375
metabolomics 106, 108, 151–153, 209, 227, 372, 375, 412, 441, 448, 453, 535
MetaCore™ 535, 536
metadata 397–398
MetaDrug 368
metagenomics 209
metaheuristic algorithm 495
metal complexation 248, 266–267
metal ions 265, 266, 295
metalloproteases 266, 541
metal organic frameworks (MOFs) 548, 560f
MetaPocket 288
MetaSite 367, 373
Meteor Nexus 368, 369, 376, 377, 379, 381
methanogenesis 124
methotrexate 261, 388
methylparaben 516
methyltransferase 518–521
metric, distance 27, 483
Metropolis criterion 315
microarrays 535, 536, 564–566
microbial cell factories 128
microconstants 73, 74
microfluidics 408, 409
microsomes 376, 432
microstates analysis 350
mixture
  multicomponent 219
  non-ideal 57
MMFF 319
model
  acceptability criteria 41–43
  degradation 574, 576, 577
  just-in-time (JIT) 576
  molecular 366
  moving window (MW) 576
performance 334, 335, 340, 344, 345, 348, 351
  pharmacophore 268
  quantum mechanical 366
  validation 30–31, 36, 209, 400, 433, 574
mode of action (MoA) 185, 186, 266, 270, 418
Modgraph's NMRPredict 143
moisturizer 528–530, 542
MoKa program 76
molecular databases 209, 220, 228, 395, 502–506, 512, 513, 521
molecular data handling 134
molecular descriptors 10, 11, 14, 135–137, 172, 334, 549
molecular docking 274–276, 315, 515
molecular dynamics (MD) 149, 240, 272, 276, 277, 291, 296, 320, 367, 407, 428, 529–531
molecular features 101, 134, 412, 419
molecular informatics 2, 407
molecular interaction field (MIF) 23, 24, 76, 269, 275, 367
molecular libraries initiative project 246, 395
molecular mechanics (MM) 137, 181, 316, 319, 320, 367, 428
molecular modeling 166, 168, 211, 278, 366–368, 385, 418, 439, 548
3D molecular models 135
Molecular Operating Environment (MOE) 263, 278, 508, 509
molecular property (P), 54, 56, 61, 388, 449
molecular refraction 57
molecular scaffolds 412, 507, 508, 520
molecular shapes 268
molecular surface 16, 55, 176, 177, 186, 224, 285, 320
molecular weight 57, 112, 178, 249, 250, 284, 292, 338, 370, 379, 405, 430, 507

3D-molecule representation of structures based on electron diffraction (3D-MoRSE) codes 16
Molfiles 253, 389ff
MolSql 449
molybdenum oxygenases 363
moment, dipolar 560
monoamine oxidases 171, 185–187, 238, 363, 418
monooxygenase enzymes 112, 114, 363, 366, 374
Monte Carlo method 181, 316, 491, 560, 561
Moore–Penrose 484
morphine 96
MOSES 353, 449
multi-class classification 483
multidrug resistance (MDR) 353
multilayer perceptron (MLP) 335
multi-learning approaches 77
multi-linear regression (MLR) 11, 18, 20–22, 61–65, 139, 147, 335, 338, 350, 353, 420, 484, 555
multiple sequence alignments 285, 291, 298
multiplicity 141, 153
multiprotic 73
multivariate classification 150
multivariate curve resolution (MCR) 492
multivariate analyses 150, 471
MurI inhibitors 180f
muscle relaxants 179
mushroom 208
mutagenicity 183, 429, 431–433, 441, 445
mutation 196, 198, 432
MutSigCV 198
mutual overlap 288
Mycobacterium smegmatis, 126
m/z values 137

n

NADP *see* nicotinamide adenine dinucleotide (NAD)
naive Bayes (NB) classifier 27–28
nanomaterials 548, 550, 551, 553–557
nanotechnology 548
nanotubes 554
narcosis 185
National Cancer Institute (NCI) 6, 385–391
National Center for Biotechnology Information (NCBI) 175, 201, 246, 250, 253, 255, 256
National Institutes of Health (NIH) 6, 13, 175, 246, 385, 395
natural products 151, 172, 207–212, 215, 220, 223–228, 411
NCI Database Browser 386, 387, 392
negative contoured maps 23, 573
neural networks 19, 26, 59, 144, 553
  architecture 559
  artificial 28, 135, 138, 502, 511, 571
  back-propagation (BPG) 154, 338
  counter propagation (CPG) 29, 138, 139, 145, 186, 224, 434
  deep 77
  feed-forward 145, 147, 151, 492, 555
general regression (GRNN) 29
  multilayer feed-forward 558
neuraminidase (NA) inhibitors 226
neuron 28
new approach methods (NAMs) 448, 453, 462
new chemical entities (NCEs) 166f, 406
new drug application (NDA) 452
new molecular entities (NMEs) 167, 198, 407
Newton optimizer 179
nicotinamide adenine dinucleotide (NAD) 121, 517
Nipagin 516
Nipasol 516
NMR chemical shifts 142–145, 147
NMR coupling constants 142
NMRPredict 143
NMRShiftDB database 145
NMR spectra 134, 136, 141

NMR spectroscopy 134, 140, 141, 150, 313
automatic structure elucidation 141
machine learning methods 144, 145, 147, 149
no effect level (NOEL) 444, 465, 543
NOESY 154
Nomenclature Commission of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) 117
non-electrolyte activity 57
non-hydrogen atoms (NHA) 66, 67, 136, 154, 512
non-ideal mixtures 57
non-linear approaches 71
non-linear iterative partial least squares (NIPALS) 487
non-linear mapping (NLM) 19
non-linear models 20, 25–30, 556, 560, 562
norethindrone 318
normalization 44, 249, 451, 473, 536
Notopterygium incisum, 225
nuclear magnetic moment 142
nuclear receptor modulators 272
nuclear spin energy 142
nucleic acids 106, 144, 407, 528
database 390
nucleosides 390
nutriepigenomics 517

o

obesity 275
obligate anaerobe 124
OCSS, 98, 99
octanol/water distribution coefficient (log D) 67–68
octanol/water partition coefficient (log P) 54, 63–67, 184, 333, 335, 337, 550
odorants 503, 505, 512, 513, 521
off-target effects 13, 199, 517
oleoresin 225
olfactory system 509, 512, 513
olivetoric acid 222
OMEGA 316
oncogenes 196, 518
OncoLogic 431
oncology 196
Online Chemical Modeling Environment (OCHEM) 48, 71
OntoBrowser project 451
ontology 298, 397, 443, 451, 536
Ontology Lookup Service 451
open access 102
open circuit photovoltage 557
open data 48, 102
Open3DQSAR 24, 47
open-source 24, 305, 389
opentargets.org 196, 200
Optical Structure Recognition Application (OSRA) 388–389
OptiSim 36
oral bioavailability 63, 293, 349, 353
OrChem 449
ordinary least squares (OLS) 47, 484, 488
organic synthesis 94, 97, 99, 102, 129, 175, 182, 408, 409
Organization for Economic Co-operation and Development (OECD) 41, 54, 431, 445
organ-on-a-chip devices 408
osteoporosis 565, 566
outlier 17, 397, 485, 553, 573, 574
overdosing 363
overfitting 25, 27, 29, 38, 482, 488, 552, 554, 574
overtraining 455, 553
oxidoreductases 117, 119, 121

p

PaDEL 459
PAFA (Priority-based Assessment of Food Additives) 452–464
PAIRS 138
p-aminobenzoic acid 260f
Panax quinquefolius L., 227
parachor 57
parallel synthesis 53, 134
parsing 252, 391
partial charges 16, 98, 119, 145, 187, 368, 457, 460, 555

partial least squares (PLS) 18, 21–23, 73, 147, 181, 294, 335, 420, 479, 502, 511, 571
  discriminant analysis (PLS-DA) 480
  regression (PLSR) 181, 420, 485ff
particle swarm optimizer (PSO) 21
partition coefficients 11, 55, 63–67, 320, 338, 347, 442, 456, 520
partitioning, recursive 480
PASS 286, 288, 387
passenger alterations 198
PAST 298
PATENTS dataset 72
pathogen 170, 274, 516, 540, 562–564
pathological mechanisms 200
pathways 111, 375
  biochemical 106
  metabolic 228
  synthetic 86
patient stratification 196, 199
pattern recognition 17, 135, 138, 462, 475–480
peak intensity 150
Pearson 252, 321, 485
PEDANT database 123, 124
penicillin 166
periodontal disease 124f
perlatolic acid 222
permeability 342–347, 349, 350
peroxidases 363
perturbation 142, 198, 199, 262
pesticides 4, 431, 444, 453, 459, 465
petrochemical process 572, 582
PFAM 298
P-glycoprotein (P-gp) 343, 514, 515
pH 67–69, 266, 349, 506
pharmaceutical and food preservatives 514–516
pharmacodynamic effects 359, 381
pharmacognosy 219–220, 230
pharmacokinetics 167, 183, 190, 359, 365, 371–372
pharmacophore 11, 16, 24, 169, 178, 221, 224, 259–279, 320, 419
Pharmacovigilance 398–401
PharmDock 276
phase I and II biotransformations 362–364
phenols 186–187, 376
PhenomicDB 196
phenotype 124, 198, 418, 462, 586
phenylethyl amide 373
phenylpyrazole inhibitor (INH) 424f
phloretin 221
photo-ageing 538
photovoltaic cell 556f
Phydbac 298
Phyllanthusengleri, 412
phylogeny 298
physicochemical properties 73–76, 135, 186, 349
physiologically-based pharmacokinetic (PBPK) models 543
physiological properties 349
physodic acid 222
phytochemicals 222, 226–228
phytotherapy 230
Pipeline Pilot 400, 432
π-π interaction 265, 266
PIQOR™ 536
Pistacia lentiscus, 225
Pistoia Ontologies Mapping project 201
pK_a 54, 68, 73, 75, 76, 349, 350, 457
Plant Protection Products Regulation 429, 435
PLANTS 315, 316
plasma protein binding (PPB) 343, 349, 350, 353
platform technologies 547
pleiotropic 220
PM3, 66
POCASA 288
POCKET 285
pocket detection 285, 288
PocketMatch 300
PocketPicker 285, 288, 293
polarity 294, 336, 346, 509
polarizability 59–60, 75, 98, 144, 338, 339, 346, 550
polarization 55, 149
polar surface area 347, 349, 430
polyacetylenes 225, 226

polyhalogenated dibenzo-p-dioxins 445
polymers 554, 557, 558
polymerization 561
polymorphic forms 55, 69, 71
polymorphisms, genetic 363
polypharmacology 171, 303, 379, 407, 411
poremean curvature 560
posees 275, 315–317, 321, 425, 428
positive contoured maps 23
potential energy 16, 419, 428, 531
potential of mean force (PMF) 319, 423
Power User Gateway (PUG) 255
PPARγ partial agonists 225
precision medicine 406, 409–411
preclinical and clinical trials 188
preclinical data science 196
preclinical research 199, 201
preclinical testing 167
prediction 10, 18, 25, 39, 387, 421
  accuracy 335, 343, 574, 578
  active site 284
  activity 391
  bioactive molecules 273
  druggability 292
  error 489
  physicochemical properties 53
  polypharmacology 411
  pose 317
  property 333
  reaction 84
  reliability 143
  spectra 135
  structure 139
  target 406
  toxicity 199
predictive error of sum of squares (PRESS) 39, 488
preferential exclusion model 528, 531
pre-processing methods 19, 475
PrGen 420, 422
primary cancer cells 199
principal component analysis (PCA) 17, 19, 35, 42, 150, 227, 335, 420, 462–464, 476, 487, 502, 511, 558
principal component regression (PCR) 18, 22, 487, 488, 571
principal components 19, 22, 35, 420, 462, 476, 477, 491, 559
principal moments of inertia 24
PRINT 298
prior, Gaussian 555
privileged structures, 219
probabilistic neural network (PNN) 29
ProBis 299
PROCAT 298
process analytical technology (PAT) 573
process control 7, 571, 581
pro-drug 361, 381
projection 21, 35, 36, 294, 462, 463, 472
projection of latent structures (PLS) 18, 21–24, 147, 335, 420, 486–488, 571
promiscuity 305, 364
property prediction 10, 11 *see also* QSAR/QSPR
proportional-integral-derivative (PID) controller 581f
proprietary methods 143
propylparaben 516
PROSHIFT 147, 149, 157
PROSITE 298
prostaglandin E2 synthase1 (mPGES-1) inhibitors 221
ProSurfer 299
Protein Data Bank (PDB) 285, 389, 390, 423
protein-ligand binding 428
proteins 106, 108, 143, 144, 147, 264, 266, 269, 276
  flexibility 291, 303, 317
  fold 298
proteins-based pharmacophoric filters 323
protein-solvent-protein (PSP) 285, 287
proteomics 106, 108, 166, 170, 535
protonation states 326
protoporphyrin IX 366, 419–428
protoporphyrinogen 419–428
Pseudomonas aeruginosa (PA) 562

pseudoreceptor model 422
Pseudo-Rotational Online Service and Interactive Tool (PROSIT) 298, 390
pseudo-synthesis 409
PubChem 13, 112, 175, 245, 391, 395, 512
PubMed 126, 246, 250, 375, 396, 400
pulmonary hypertension 221
purities of compounds 55
pyrazolopyrimidinediones 180
Python 449

q

3D QSAutogrid 24
q^2-GRS method 23
Q-SiteFinder 287
qualitative structure–property relationship (QSPR) 333, 548, 585
quality by design (QbD) 573
quality control 227–228, 493, 543, 585
quality criteria 32, 39–41, 294
quantitative high throughput screening (qHTS) 441
quantitative structure-activity relationships (QSAR) 10–46, 183, 221, 365, 395, 442, 484, 543, 551, 585
quantitative structure-odor relationships 512–513
quantitative structure–property relationship (QSPR) 10–13, 59, 71, 333, 342–349, 548, 549, 551, 554, 571, 588, 589
quantitative structure-spectrum relationships 139
quantitative weight of evidence assessment 447
quantum chemistry
  calculations 138
  descriptors 66
  methods 55
  prediction of NMR properties 142
quantum efficiency 557
quantum mechanical (QM) methods 97, 137, 142, 367
molecular mechanical methods (QM/MM) 367
quantum mechanics 54, 142
queries 138
  data 250
  processing 254
quetiapine 374–375

r

radial basis function (RBF) 29
radial distribution function (RDF) 16, 135–145, 224, 339
Radix Astragali (Huangqi) 238, 240
Radix Ginseng (Renshen) 240
Radix Ophiopogonis (Maidong) 238, 240
Radix Puerariae (Gegen) 238, 240
Radix Rehmanniae (Dihuang) 238, 240, 242
Radix Rehmanniae (H02) 241
Radix Trichosanthis (Gualou) 238, 240
Raman spectrum 138
Random forest (RF) 17, 18, 26–27, 68, 144, 147, 149, 293, 335, 368, 400, 432, 433, 480, 487
randomization tests 271, 490
random sampling 34
random splitting 34
random subsets 27, 489
range-based methods 42
rapid automated materials and processing (RAMP) 549
Rapid Overlay of Chemical Structures (ROCS) 226
RApid Pocket MAtching using Distances (RAPMAD) 301
rational selection methods 34
RDKit 449
reactions 83–91
  biochemical 106
reaction center 91, 111, 125, 365, 367, 370
reaction databases 84, 86, 91, 96–97, 110, 111, 125, 128
Reaction InChI (RInChI) 91
reaction mechanisms 90, 92, 119
reaction pathways 95, 97, 113

reaction planning 84, 86
reaction prediction 87–97
reaction retrieval 111
reaction sites 110, 111, 123
reaction tree 412
reactivity 90, 92, 93, 102, 261, 365, 445
  chemical 10, 11, 20, 58, 97, 365, 372, 445, 447
Reactome 111, 197
Read-Across (RA) 442, 445–448, 543
real time release testing (RTRT) 573
Reaxys 96
RECAP 175
Receiver Operating Characteristic (ROC) curves 45, 226, 272, 274, 323, 483
receptor 167, 361
  binding 11, 166, 178
  protein 506, 534
  surface 24
receptor-based deduction 269
receptor-based pharmacophores 269–270
recoding structural features 248
recursive partitioning 25, 430, 480
redocking 323, 324
refractive index 554, 557, 558
regenerative medicine 548
regiochemistry 92, 93, 375
Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) 441, 446, 586
regression
  analysis 20–22
  backward-elimination 21
  ensemble online support vector (EOSVR) 577
  forward-selection 21
  ill-posed 552
  linear 20, 71, 293
  multi-linear, *see* multi-linear regression analysis
  multiple linear 420, 555
  ridge 486
  support vector 29, 571
  tree 480
  univariate 484
regression/correlation models 11
regulatory and metabolic capabilities 123
regulatory databases 247, 505
regulatory science 439–464
relational database management system (RDBMS) 449
repeated double cross-validation 491
repositories 196
representational state transfer (REST) 449
rescoring 320
Research Institute for Fragrance Materials (RIFM) 505
residual error 484
residual sum of squares 39
resonance effect 58
response randomization 38
retrieval system 111
retroaldols 97
retrosynthetic analysis 94ff
reverse docking approach 224
reverse pathway engineering (RPE) 125–128
Rhea 113
Rhizoma Anemarrhenae (Zhimu) 238, 240
Rhizoma Dioscoreae (Shanyao) 238, 240
rho kinase inhibitors 220, 221
ribonuclease 261
ridge regression 486
ring effects 63
risk assessment 350, 417, 441–448
ROBIA program 97
robustness 336, 421
ROESY 154
root mean squared error of cross-validation (RMSECV) 488
root mean square deviation (RMSD) 324
root-mean-square error (RMSE) 41, 486, 555, 577
rotatable bonds 316, 430
rotation forest 27

R squared (coefficient of determination), 554
R statistical package 27
rule-based systems 133, 368, 429
Russel 252
Ruta graveolens, 223
rutamarin 223
RXN00173, 112

S

Salmonella reverse mutation assay 432
Salmonella typhimurium, 432
salt bridges 318
saponins 227
SARpy 459
scaffold 365, 430
  hopping 262, 272
  molecular 507
  Murcko 507
scaling 19, 472–474
scatterplot 150
Schrödinger's equation 92, 142
SciFinder 96
SCOP 298
score-based metabolic reconstruction 124
score histogram 324, 325
scoring,
  binding affinity prediction 317
  consensus 320
  function 181, 276, 294, 299, 315, 317–330, 407, 423
scPDB 303, 304
scrambling 554
SCREEN 293
screening 181, 245, 251, 298, 323, 391
  commercial libraries 513
  experimental random 422
  in silico 292
  NMR-based fragment 292
  phenotypic 405
  shape-and feature-based 274
  similarity 226
  virtual 172, 209, 211, 226, 240, 423
scree plot 477
SDF-format 248, 389, 390, 396
search(ing)
  full structure 91, 112, 250
  reaction 91
  similarity 91, 172, 252
  substructure 251, 387
  superstructure 251
secondary metabolites 125
second order approximation 56, 61
selection 211, 486
selectivity 177, 183, 263, 273
self-organizing map (SOM) 17, 19, 28, 119–121, 173, 228, 293, 335, 338, 343, 478
semantics 397
sensitivity 45, 324, 370, 432
sequence homology 143
sequence search algorithms 297
serotonin 224, 254
SESAMI 154
sesquiterpene lactones (STLs) 224, 228
Setubal principles 336
shake algorithm 531
shape-matching 406
shared encoding 250
side effects 188, 199, 298, 326, 343, 361, 401, 540
SIENA 297
SIFTER 298
signal-to-noise ratio 488
similarity 10, 19, 250, 397, 446
  analysis 41
  maps 512
  searching 252
  Tanimoto 369
similarity ensemble approach (SEA) 199
simplified molecular-input line-entry system fingerprint (SMIfp) 509
SIMPLS 22
simulated annealing 17, 21, 315, 316, 430, 495
simulation
  molecular dynamics (MD) 530
  spectra 133, 138
single nucleotide polymorphisms 363
SiteAlign 300
SiteBase 299
SiteEngine 299

SiteFinder 287
SiteMap 287, 293
site of mechanism (SoM) , 176, 177, 367, 369, 377
sketcher 253, 254
skin aging 528, 529, 542
skin sensitization 431, 443
SLN strings 252, 253
smallest set of smallest rings (SSSR) 63
SMARTCyp algorithm 367
SMARTS 253, 254, 367, 445, 459
SMILES 112, 252, 387, 396, 450, 459
  canonical 249
  translator 389
  unique 389
soft sensors 571–582
solar cells 548, 556, 557
sole suspect 398
solubility 54, 69, 72, 334, 336, 361, 551, 554
solvent
  accessibility 285, 301
  supercritical 554
solvent accessible solvent area (SASA) 76
SONNIA 173, 176, 186
SOPHIA 97
SPARTA+, 147, 149
spectra 133–156
spectral data 5, 133, 134, 150, 461, 473, 483
spectrum prediction and comparison 134
spectrum–structure correlation 135
sphere exclusion 36
spherical volume 425
spin–spin couplings 154
SPINUS 145
splitting, data 32
SQL 450
π-stacking 265
standard deviation (SD) 38, 61, 65, 66, 144, 149, 172, 287, 335, 451, 472
standard deviation error of prediction (SDEP) 41
standard error 38, 554, 558, 563, 564
standardization 201, 247–250, 255, 413, 472–475, 478, 503
Staphylococcus aureus (SA) 562
StarDrop P450, 367
static fingerprints 459
statistical analysis 178, 269, 319, 429, 442, 447, 533
statistical modelling 23, 30, 34, 76, 200, 561, 571
statistical significance analysis 271
steady-state 342, 372
stereochemistry 110, 136, 154, 211, 248, 249
stereoselectivity 93
steric clash 267
steric contours 23, 24
steric effects 58
steric hindrance 75
steroids 445
sterol 224
Stevens–Johnson Syndrome (SJS) 398–401
stochastic methods 315
stratification 34
stratum corneum 529
StromBone^{TM} , 565
structural identity 140
structural keys 15
structure-activity relationships (SAR) 4, 11, 260, 418, 419, 442
  quantitative, *see* quantitative structure-activity relationships
structure-based approaches 422–429
structure-based drug design (SBD) 168, 179–182
structure-based pharmacophore model 269
structure-based virtual screening (SBVS) 313–326
structure elucidation 133–135, 140, 209
structure-flavor relationships 511
structure generator 110, 172
structure ID (SID) identifier 246, 387
structure–metabolism relationships (SMRs) 368

structure–property relationships (SPR), *see* quantitative structure-property relationships
structure query 253
structure representation 248
structure searching 91, 112, 250
structure-spectra correlation 133, 137
structure validation 134
subgraph 456
substance database 246
substrate 167, 360
substructure recognition 150
substructure-subspectrum database 136
sulfotransferases 277
sulphonation 363
sulphone-di-oxygenation 374
sulphotransferase(SULT) 365
sulphoxide-mono-oxygenation 374
sum of ranking differences (SRD) 474
SuMo 299
superfeature 277
superimposition 123, 179, 419
SuperScent 506, 509
SuperStar, 269
superstructure search 251
SuperSweet database 505, 509
supervised learning 18, 335, 475
support vector machine (SVM) 18, 29–30, 59, 119, 144, 147, 294, 335, 343, 400, 432, 480, 488, 560
support vector regression (SVR) 29, 571
surface 181
  hydrophobic 327
  molecular 224
  GBSA 181, 320
  protein 283
SURFNET 286
sweeteners 514
SYLVIA 101, 182
SynChem 100
SynGen 99
SYNOPSIS 101
synthesis
  biology-oriented (BIOS) 412
  continuous flow 408
  in silico 375
short 96, 407
  total 87
synthesis design 84, 86, 94–102
synthesis planning programs 99
synthesis routes 407
synthetic accessibility 101, 182
Synthetically Accessible Virtual Inventory (SAVI) 391
System for Organic reaction Prediction by Heuristic Approach (SOPHIA) 97
systems biology 201

t

tailored/target-specific scoring functions 320
Tanimoto coefficient 36, 112, 172, 252, 369, 370, 510, 517
targetability 292
target-based *drug* discovery 405
targeted/focused libraries 178
target fishing 198, 223
target identification 167ff, 196
target-mediated toxicities 199
target mining 196–198
target validation 170
taste 503
tautomerism 55, 249, 252, 326, 428, 445
TCGA database 196
Tenascins 540
TESS 299
test set 17, 334, 490
2,3,4,5-tetrachlorophenol 187
Tetrahymena pyriformis, 433
tetrazole 409
text mining 198, 396–398, 400, 401
Theonella swinhoei, 224
thermal stability 151
thermodynamic integration 321
thermodynamic properties 60–63
3-methybutanoic acid 126
  threading algorithm 287
three-point contact model 261
3R philosophy 429
Threshold of Regulation (TOR) 443

threshold of toxicological concern (TTC) 443, 444, 446, 543
Tikhonov regularization 487
TIMES 368
tissue inhibitor of metalloproteinases (TIMP-1) 541
tizanidine 179
TOCSY spectra 154
Toll-like receptors, subtype 2 (TLR2) 274, 275
tolperisone 179
top-down approach 442
Topiramate 200
topoisomerase 540
topological autocorrelation 173, 186
topological descriptors 16, 144, 338, 555
topological distances 73
topological indices 16, 71
topological polar surface area (TPSA) 347, 349
topological properties 553
topological spheres 73, 75
topomer 25
torsional energy 319
total body clearance 349
total synthesis 87
totrain 320
Tox21 algorithmic groups 200
toxalerts 58
ToxCAST™ 395, 431, 441
toxicity 11, 58, 183, 184, 361, 372, 391, 442
  aquatic 433
  assessment 441
  chronic 431
  genetic 431
  prediction 199
*to*xicity *p*rediction by *k*omputer *a*ssisted *t*echnology TOPKAT 430
Toxicity Reference Database (ToxRefDB) 432
toxic modes of action 185
toxicological alerts 187
toxicology 14, 177, 184, 200, 254, 350, 365, 370, 372, 374–375, 398, 417, 429, 431–435, 441, 586
ToxPrint 58, 187, 453–458
ToxTree 431, 445
trachelogenin 226
Traditional Chinese Medicines 171, 238 *see also* Chinese herbal medicine
training set 11, 17, 145, 200, 264, 334, 555
trans-diethylstilbestrol 261
transepidermal water loss (TEWL) 528
transferases 117, 363
transition states 121, 123
transmembrane protein 534
transport system 167, 208
tripeptoid library 175
Triton-X100, 426
TrixP 302
TrixX 316, 317
true actives 324
true negatives 45, 273, 324, 343
true positives 45, 273, 324, 343, 483
t-test 485, 536
tumor suppressor 198, 518
tumor therapy 353
turnover number 554, 558, 559
Tversky 252
Tylenol 254
type 2 diabetes (T2D) 222, 225, 237

u

ultra high-throughput screening (u-HTS) 166
uniform design method 35
uninformative descriptors 551
UNIQUAC Functional-group Activity Coefficients (UNIFAC) 57
univariate regression 484
unsupervised learning 17–18, 119, 173, 335, 475
uridine 5'-diphospho-glucuronosyltransferase (UGT) 365
uropathogenic *Escherichia coli* (UPEC) 562
US Food and Drug Administration (FDA) 238, 405, 443, 502

v

valency states 248
validation 11, 30, 31, 271, 323, 334, 335, 400
  applicability domain 41, 42
  data compilation 32–37
external 30–32, 37, 41, 43, 44, 442, 492
  goodness of prediction 39
valproic acid 517
van der Waals (vdW)
  energy 16, 287
  radii 260, 285
variables 472–488
variable importance (VIP) 486
variable selection methods 23, 44, 485
variable subset selection (VSS) 488
variance analysis (ANOVA) 474ff
VAST 298
VB model 248
vector 135
  score 476
venetian blinds 489
vibrational spectroscopy 137
VigiBase 398, 400
virtual hits 182, 274
virtual organic reactions 409
virtual screening 10, 101, 166, 175, 177, 210, 211, 220, 272, 278, 313
  see also structure-based virtual screening (SBVS)
virulence factors 170
viscosity 554, 557
VolSurf 16, 24, 267
VSN algorithm 536

w

Waikato Environment for Knowledge Analysis (WEKA) 27
wall chart 106
water-ectoine mixture models 529ff
water-glycerol complex 531ff
WaterMap 321
water-binding activity 528
water solubility (log S) 55, 57, 69–71, 336–342
weak polar interactions 318
weighted averaging 26
whole-molecule descriptors 65
wisdom of crowds concept 26
Workbench for the Organization of Data for Chemical Applications (WODCA) 99
World Drug Index 432

x

xenobiotic metabolism 359–375
xenografts 199
Xiaoke 238–240, 242
XLogP 57, 65, 346, 464
X-ray
  crystallography 168, 321, 323, 423, 425, 428
  structure 168, 180, 269, 411, 418
X-Score 320
xyz-format 390

y

Yalkowsky's general solvation equation (GSE) 70
yield
  of actives 274
  of reaction 93
Y-permutation 38
Y-randomization 31, 32, 38–39, 45
Y-scrambling 30, 38–39, 490
Yule 252

z

zeolites 559
zeolitic imidazolate framework (ZIF) 560
zero-order approximation 56
ZINC databases 514
Z-test 38, 39
zwitterions 67, 75, 528