Index

a
absorption, distribution, metabolism, 
excretion, and toxicity (ADME/Tox) 39, 283
Accelrys FCFP4 fingerprint 110
activity Cliff-forming scaffolds 71, 72
– activity Cliff concept 71
– multitarget Cliff-forming scaffolds 71–73
AllChem 94
N-alpha-(2-naphthyl-sulphonyl-glycyl)-or-
pamidinophenylalanyl-piperidine (NAPAP)
ligand 238, 239
amino acid
– interaction types 223
– pharmacophoric properties 221
– residues, responsible for binding of 261
– sequence 172, 260
– specific binding 223
AstraZeneca 17, 87
– virtual library 95
acyclic fragments 233
AZ-VL system 97

b
bacterial neuraminidase, X-ray crystal
structure 241
BCI fingerprint 110
benzimidazole scaffold 3
BI CLAIM system 97
binding affinity 3, 48, 221, 222, 264, 265, 268,
269, 273
binding conformation, superimposition
of 287
BindingDB 64
bioisostere encoding 144
bioisosteric replacements 247, 259
BIOSTER database 144
bis(trifluoromethyl)phenyl rings 266
BROOD system 179
– bioisosteric similarity methods 180
calcium-dependent NOS activity 290
calmodulin (CaM) 284
Cambridge Structural Database (CSD) 271
CAS. see Chemical Abstracts Service (CAS)
CAS (Chemical Abstracts Service) registry
43, 46
– nonredundant library 46
CATS. see chemically advanced template search
(CATS)
CATS 2D autocorrelation function 111
CATS fingerprints 108, 111, 114
cavity-independent fingerprints (APIF) 222
cerebral palsy, prevention 289
CG. see chemical graph (CG)
ChEMBL database 44, 49, 52, 53, 55,
62, 97
ChEMBL data set 111
Chemical Abstracts Service (CAS) 16, 43
chemical clichés 44
chemical graph (CG) 131, 133
– paracetamol molecular structure 132
chemically advanced template search
(CATS) 119–122
– CATS2 atom types 121
– descriptor
– generation 122
– median BEDROC performance 125
– fuzzy molecular representations 120
– modifie ATS descriptor 120
chemical patent 15, 17, 28, 34, 36
chemical space 83, 84
– comprehensive enumeration 85
– BOOMSLANG algorithm 88
– fragments 85, 86
– Kohonen map 86
– reagents 87, 88
– ring systems 86, 87
– SMIRKS manipulation 87
INDEX

VEHICLE 86
– iterative stochastic search 85
– perspective 99
chemoinformatics 107
claimed intellectual property 17
COBRA database 111, 114, 123
community-selective scaffolds 69
computed molecular properties 107
computer encoded chemical reaction 94
consensus shapes
– and center molecules, calculations 187, 188
– to probe CCR5 extracellular pocket 190–192
COX-2 inhibitor 4
CP-99994
– benzyl amino group 265
– binding mode of 265
Cresset's field technology 270
cumulative scaffold frequency plots (CSFP) 49
cyan sticks 216
cycle graph 149
cyclin-dependent kinase (CDK) family 242
cyclooxygenase-2 (COX-2) 235
data visualization 55
Daylight fingerprints (DFPs) 108, 110, 138, 139, 142
diarylthiazolotriazole 4
dictionary-based approach 108
2,5-dioxopiperazine 267
DOGS program 94
DrugBank 43
drug discovery
– application of Spark 206, 207
– automatic devices 99
– and design 3
– Drug Guru 36
– similarity searching in 155–157
Drug Guru program 36, 89, 90, 99
drug metabolism and pharmacokinetic (DMPK) 247
3D scaffold replacement method 175
– BROOD 179, 180
– ChEMBLstructures 177
– commercial structure 177
– CSD structures 176
– databasess 176, 177
– filtering 177, 178
– fragment generation 177, 178
– fragment replacement search 178
– molecule databases 175
– ReCore 179
– scoring 178

– SHOP, as novel replacement scaffolds 178, 179
– virtual structures 177
– workflow 175, 176
2D structure 98
e
ECFP4 fingerprints 111, 114
electronic laboratory notebooks (ELNs) 95
enantiospecific inhibition 252
endothelial nitric oxide synthase (eNOS)-derived NO 283
ER activity, designed compounds with 236
ester linkage, in tryptophan-derived NK1 antagonists 269
estrogen receptor (ER) 241
– affinity 235
– site point constraints 240
ethylendiamine fragment 286, 288
Extended Connectivity Record (ECTR), programs 16
eXtended electron distribution (XED) force field. see XED force field
f
feature point pharmacophores (FEPOPS) 157–159
– active recall, and scaffold hopping descriptor assessment 166
– alignment of protein kinase C inhibitor 165
– alternative alignment 168–170
– atomic properties tocentroids, assignment of 161
– average similarity of probes 168
– chemical space uniqueness 171, 172
– computation 159–162
– defining scaffold uniqueness 163, 164
– descriptor 160
– generation of feature point pharmacophores 161
– k-means and k-medoids calculations 160, 162
– clustering 158
– multiscale reduction of ATP atoms 159
– multiple flexible conformers 160
– Omega software, use of 162
– perspective 172, 173
– recovery of validated HTS actives, clusters, scaffolds, and ring systems 167
– scaling and correlations 162
– in silico target prediction 170, 171
– in similarity searching and scaffold hopping 164–168
Index 299

feature trees 149
-- applications 152, 153
-- CoLibri software, use of 153
-- comparison 150
-- FTrees–FS system
-- exploration of combinatorial chemistry spaces 152, 153
-- using variant of Feature Trees representation 153
-- FTreesXL 152
-- generation 149, 150
-- implementations 152
-- MTree 152
-- RECAP fragmentation scheme 152, 153
-- retrospective validation 151, 152
FEPOPS. see feature point pharmacophores (FEPOPS)
FieldStere 270
fingerprints 108. See also molecular interaction fingerprints
-- based virtual screening 109, 112
-- for ligands and proteins (FLAP) method 217
-- scenarios 224
fragment-based drug design (FBDD) 238
fragment connectivity rules 238
fragment hopping
-- to design novel inhibitors, for neuronal nitric oxide synthase 283–288
-- minimal pharmacophoric elements and 281–283
-- nNOS inhibitors 287
-- to optimize neuronal nitric oxide synthase inhibitors 288, 289
-- schematic flow diagram 282
freedom-to-operate (FTO) inquiry 17
Free–Wilson analysis 17
functional vs. structural molecular scaffolds 7

g
Gaussian function 224
GENSAL language 16
Glu193 265
glycine transporter type 1 (GlyT1) inhibitors 247
G protein-coupled receptor (GPCR) 64, 280
-- ligand 220
graphical user interface plugin 231
graph matching 143, 144
graph reduction
-- attempts to simplify objects 131
-- generation 133
-- extended reduced graph (ErG) approach 136, 137
-- node labeling 134, 135
-- reduction scheme 133, 134
-- Sheffield implementations 135, 136
GRID, computational methods
-- based similarity searches 178
-- understanding selectivity of inhibitors for 281
GTPcS functional assay 236

h
halogen-substituted phenyl groups, usages.
see fragment hopping
H-bond acceptor 217
heteroatoms 107
hierarchical scaffolds 63
high concentration screening (HCS) 279
high-throughput screening (HTS) 9, 155, 166, 168, 279
-- applications 279
-- Merck’s effort 253
-- T-type Ca$^{2+}$ inhibitors 249
-- VU/MLCPN, chemical optimization 250
His 265
-- Ala-replacement 265
-- significance 265
HTS. see high-throughput screening (HTS)

i
IADE software 89
intellectual property (IP) space 247
interaction annotated structural features (IASF) method 216
inverse QSAR technique 89, 90

k
Kohonen networks 54

l
L-732138 266, 269
L-737488, analog 269
ligand-annotated target fingerprints 218, 222
ligand-based hypothesis 155
ligand–water–protein interactions 241
lipophilicity 3, 88
L-$\text{N}^\circ$-nitroarginine methyl ester (L-NAME) 290
L-732138, template molecule 270
Lys194 265

m
Markush, Eugene 15
Markush Input Language (MIL) 18
Markush structure 15, 16
Index

- correspondence between MIL file and 21, 22
- economic importance 16
- encoding 18
  -- exact R groups 19
  -- fused R groups 19, 20
  -- inexact R groups 19
  -- r_group records 19
- Menguin program 20, 21
- MARPAT systems 17
- maximum common substructure (MCS) 9–11, 44, 84, 91
- algorithms, challenges 9
- MDDR database 109
- activity classes from 109
- scaffold-hopping searches 110
- MDL fingerprint 108, 110, 111
- medicinal chemistry scaffolds, enumerated possibilities 47
- integer descriptors 46
- Lipkus’ topology analysis 47
- nonredundant library, in CAS registry 46
- ring systems 46, 47
- topologies of voids 46, 47
- VEHICLe database 47
- MEGx database 114
- Menguin program 18
- Merged Markush Service (MMS) 16, 17
- MIL file 18
- molecular cloning studies 248
- molecular geometry 107
- molecular interaction fingerprints
  -- ligand-annotated target 217–219
  -- protein–ligand complexes
  -- three-dimensional (3D) structures of 215
- protein-ligand pharmacophores 217
- target-annotated ligand 215, 216
  -- interacting atom/fragment 216, 217
  -- true target-ligand 220
  -- association of 220–222
  -- interaction pattern 222–225
- molecular libraries production center network (MLPCN) 248
- molecular libraries screening center network (MLSCN) 248
- molecular topology 107
- molecular weight 84–86, 88, 107, 261, 279
- molecule-based approach 108
- Molecule Evoluator 89
- MotifScore 223
- multiple objective optimization 90
- multiple solvent crystal structure (MSCS) method 281
- Murcko scaffold 110
- MUV data set 111

n
natural products, source of novel scaffolds 45, 46
- NATx database 114
- neurokinin 1 (NK1) receptor 260
  -- antagonists 262
  -- residues, involved in binding of nonpeptide antagonists 265
- neuronal nitric oxide synthase (nNOS) 283
- neuropeptide SP, NMR structure 265
- neuropeptide substance 261
- nitric oxide synthase (NOS), chemical structures 285
- nitroarginine-containing dipeptide/peptidomimetic inhibitors 284
- nNOS
  -- binding conformation 289
  -- crystal structure 288
  -- dipeptide/peptidomimetic inhibitors 284
  -- inhibitors, minimal pharmacophoric elements 286
- N5SB polymerase 3
- NVP-AUY922, HSP90 inhibitor 6

o
objective and invariant scaffold representations 7
  -- molecular frameworks 7, 8
  -- scaffold tree 8
- oxazolidinedione 268
- 2-oxopiperazine 267

p
ParaFit program 188–190
ParaSurf program 188–190
- PGVL system 97
- pharmacological similarity 195
  -- chemotype trap 195, 196
  -- G protein-coupled receptors 195
  -- ligand 195, 198
  -- pattern 196–199
  -- shape compatibility 195
  -- van der Waals forces 195
- pharmacophore-based interacting fingerprints 217
- pi-cation interaction 219
- Pipeline Pilot software 110
- piperazine ring system 268
polarity 3
polar surface area (PSA) 107, 286
potency-based scaffold discontinuity score (PScS) 71
privileged scaffolds 11
privileged substructures 63
Pro271 265
programmable compounds 99
promiscuous chemotypes 68
promiscuous scaffolds, in drugs 70, 71
protein and ligand fingerprints (PLFPs) 221
protein–ligand association fingerprint 220
protein ligand interaction explorer (PROLIX) 223
protein–ligand interaction fingerprints 218
protein–ligand interaction pharmacophores 217
protein–ligand structural alignments 222
pseudo-SMILES encoding 138
PubChem 32, 33, 43, 62, 86
Putative ligand structures 232
PyMOL graphical user interface 232
pyrrolidine ring 286

q
quantitative structure–activity relationship (QSAR) models 152
– feature tree representations of molecules 152
– Free–Wilson analysis 33
– inverse 89, 90, 98

r
rabbit kits
– neurobehavioral assessment 290, 291
– nitric oxide synthase/nitric oxide concentration 290
random virtual compounds 83
RDKit fingerprint 124
reaction SMARTS 93, 94
readily synthesizable compounds 94
– construction 94, 95
– iterative approaches 96
– outside big pharma 96
– searching 95
real-life reactions, for in silico molecule construction 93, 94
ReCore program 179
reduced graph (RG) 131, 133, 141
– comparison and usage 137, 138
– augmenting fingerprints, with edit distance 140, 141
– bioisostere encoding 144
– conventional fingerprinting 138, 139
– dopaminergic activity retrieved 138
– extended reduced graph fingerprints 141–143
– graph matching approaches 143, 144
– RG-specific fingerprints 139, 140
– structural diversity of 5HT reuptake inhibitors 142
– key challenges, to assess molecular similarity 145
reference structure 107
Reflex method 240
R/F reduction scheme 135
RG, see reduced graph (RG)
ring fragments 233
rotating spherical polar fourier expansions 185, 186

s
SAR data 97
scaffold composition
– drugs and small-molecule libraries 45, 46
– to interpret bioactivity data 48
– SCINS code 48
– structure–activity relationships 48
– medicinal chemistry libraries 46, 47
– medicinal chemistry space 41–45
– molecular framework analysis 45, 46
– drug and drug-like library 46
– natural product library 46
– study based on MCS 44
scaffold definition 64, 65, 84
scaffold diversity
– defining 41
– in medicinal chemistry 39–41
– medicinal chemistry space
– metrics for quantifying 48–53
– visualizing scaffold diversity 53–56
– of pharmaceutical targets 76–78
– structural relationships between scaffolds 76, 77
scaffold hopping 3, 31, 76, 78, 107, 240–242, 247–248
– access novel calcium T-type channel inhibitors 250–253
– access novel glycine transporter type 1 (GlyT1) inhibitors 253–254
– application of 235
– bioisostere 259
– capability 110
– from a CCR3 antagonist to existing library compound 173
– into chiral 3-aminopiperidine cores 251
Index

- into [3.1.0] cores 252
- de novo design 231
- Drug Guru program 36
- fast-follower project 36
- fragment-based drug design 279–281
  - minimal pharmacophoric elements/fragment hopping 281–283
- neuronal nitric oxide synthase 283–288
- optimize neuronal nitric oxide synthase inhibitors 288–289
- Free–Wilson analysis 33–35
- neurokinin 1 receptor (NK_1R) 260, 261
  - target active site and binding mode 264, 265
- neurokinin 1 receptor antagonists 261–264
  - bioisosteric replacements 266–273
- neurokinin 1 (NK_1) therapeutic areas 259, 260
- neuronal nitric oxide synthase inhibitors, application of
  - cerebral palsy prevention 289–291
  - periscope system assisting 32
  - potential 76
- CATS1 and CATS2 descriptors 125
- descriptors, actives-retrieving performance categories 124, 125
  - determining 125
  - heatmap depicting p-values 124
  - retrospective evaluation of enrichment and 122–126
- ROC-related BEDROC score 123
- virtual screening benchmark 123
- predictive studies, using 2D fingerprints 112–114
- prospective applications 126–128
- substructure searching 32, 33
- T-type calcium channel inhibitors 248–250
- using 2D fingerprints, retrospective studies 109–112
  - in virtual screening 78
  - scaffold network graphs, of drugs 56
  - scaffold representations
    - history 4–7
    - of kinase inhibitor lapatinib 40
    - libraries A and B 42
  - scaffold space, visualizations 96, 97
  - scaffolds, with defined activity progression 73
  - activity profile sequences 73–75
  - conserved scaffolds 75
  - scaffold-target family profiles 70
  - scaling bit positions 217
  - SCINS (scaffold identification and naming systems) 48, 52, 163, 164
  - descriptor 163, 172
  - ScreenScore 234
  - search algorithm 22–25
  - i3am search algorithm 23
  - matching R groups 25
    - exact R groups 25–27
    - fused R groups 28
    - hydrogen atoms 29, 30
    - inexact R Groups 27, 28
    - managing multiple fragment/R group matches 30, 31
  - molecules matched against Markush 24, 25
    - indole core 24
    - matche C_{1-6} alkoxy 25
    - methoxycarbonyl and chloride groups 25
    - R groups and portions of molecule unmatched 24
    - total score 25
    - Xenon atoms 24
  - selectivity, of scaffolds 63
    - privileged substructures 63, 64
    - target community-selective scaffolds 64–66
    - target-selective scaffolds 67
  - self-organizing maps (SOMs) 54, 126, 281
  - Shannon entropy 51
  - shell residue fingerprint (SRFP) 223
  - SHOP program 178, 179
  - similarity searching, in drug discovery 155–157
    - 4D chemical similarity method (see feature point pharmacophores (FEPOPS))
    - 2D similarity methods 156, 157
    - graph-like representations, of compounds 156
    - high throughput screening (HTS) 155
    - ligand-based hypothesis 155
    - Markush structures 156
    - scaffold hopping 156
    - structure–activity relationships 156
  - Skelgen 231
    - cyclin dependent kinase 2 (CDK2) 235
    - cyclooxygenase-2 (COX-2) 235
    - estrogen receptor 235, 236
    - histamine H3 inverse agonists 236
    - principles 231
    - PyMOL graphical user interface 232
    - retrospective validation study 235
    - scaffold hopping 240–242
      - using fixed fragments 237, 238
      - using site points 238–240
    - structure generation/optimization
      - fragments/fragment sets 232–233
      - ligand-based design 234–235
-- scoring 234
-- validation studies 235
Skelgen fragment 237
SMIRKS 87, 89
SOMs. see self-organizing maps (SOMs)
Spark
-- in drug discovery scenarios, applications 206, 207
-- P38 kinase inhibitor fragment growing using 207–209
-- working 202
-- selecting fragment replacement site 206
-- software 202, 205
SparkV10 270, 271, 272
-- heterocyclic bioisosteres of the ester linkage 273
-- template molecule L-732138 271
spherical harmonic surface
-- shape similarity 186, 187
spherical harmonic surfaces 183–185
stearoyl-CoA desaturase inhibitor 3
steric interactions 272, 285
structural interaction fingerprint (SIFt) 218
structure–activity relationships (SARs) 61, 84
structure-based de novo design 90–92
substance P receptor (SPR). see neurokinin 1 receptor
substrate activity screening (SAS) 279
Surflex-Dock score 225
SVM model 221
SYNOPSIS tool 93

t
Tachykinin NK1, family 264
tachykinin receptor 1 (TACR1). see neurokinin 1 receptor
Tanimoto coefficient 107, 108
-- values 78
tan ribbons 216
target-annotated ligand fingerprints 216
target community-selective scaffolds 64–66
target–ligand fingerprints 220
target promiscuity of scaffolds 67
-- promiscuous BM scaffolds and CSKs 67–70
-- promiscuous scaffolds in drugs 70, 71
-- scaffold–target family profiles 70
target selectivity (TS) value 67
thiophene S-oxide 4
TIFP fingerprint 224
topological structure, autocorrelation 119
toxicity 3
tree maps 54
Tripos Unity fingerprint 110

tryptophan-derived NK1 receptor antagonists 266, 270
-- binding affinity 273
-- ester linkages, case study 266, 273
-- heterocyclic bioisosteres 267
T-type Ca2+ inhibitors
-- high-throughput screen (HTS) 249
-- Merck piperidine series 251
-- second-generation 249
T-type calcium channel family 248
Tversky coefficient 108

u
Unity fingerprint 108

v
valid molecular fields, generation of 199.
See also pharmacological similarity
-- ACC approximation 199
-- Coulombic interactions 199
-- MM force 199
-- molecular mechanics 199
-- validating position and pattern of field points, tools for 200
-- Cambridge Crystallographic Database 200
-- XED model 199
Vanderbilt GlyT1 inhibitors 254
vBROOD system 180
virtual chemistry 83
virtual compounds, iterative generation 88
-- analog generators 89
-- chemical modifications, manual selection of 88, 89
-- creation, different ways to 92
-- inverse QSAR 89, 90
-- multiple objective optimization 90
-- structure-based de novo design 90–92
-- transformations 88
virtual libraries 88, 92
-- perspectives 98
virtual synthesis 92
-- construction 94, 95
-- iterative approaches 96
-- outside big pharma 96
-- readily synthesizable compounds 94
-- real-life reactions 93
-- searching 95
-- synthetic tractability 92, 93
visualizations, of scaffold space 96, 97

w
WOMBAT database 144
World Drug Alerts 242
x

XED force field 195, 200–202
– applications 200
– applied to scaffold hopping in Spark 202
– to benzene 203
– constructs
-- polarizable, and allowing for nuclear charge 201
-- types 200, 201
-- creating new molecules 208–210
-- desymmetrization of the ether oxygen atoms by 203
-- electrons distribution 200

-- far-reaching consequences of using molecular fields
-- as measures of similarity 212, 213
-- improvement in electrostatic behavior 202, 204
-- model reflecting intuitive molecular properties 199
-- new potential inhibitors 209–212
-- XED MM method 200
X-ray crystallography 279

z

ZINC database 111, 152