

Computer-Aided Molecular Design: Fundamentals, Methods, and Applications

Athanasios I Papadopoulos, Centre for Research and Technology Hellas, Thessaloniki, Greece

Ioannis Tsivintzelis, Aristotle University of Thessaloniki, Thessaloniki, Greece

Patrick Linke, Texas A&M University at Qatar, Doha, Qatar

Panos Seferlis, Aristotle University of Thessaloniki, Thessaloniki, Greece

© 2018 Elsevier Inc. All rights reserved.

Introduction	4
Molecular Representation and Property Prediction	5
GC Approach	5
Other Molecular Representations	6
Property Prediction Models	7
Pure component properties	7
Activity coefficient models	7
Cubic equations of state	7
EoS/G ^E models	9
Advanced statistical thermodynamic models—The statistical associating fluid theory approach	10
Computational chemistry models—The conductor-like screening model approach	12
Main Classes of CAMD Problems	13
Classification Overview	13
Single-Objective Computer-Aided Molecular Design	14
Computer-Aided Molecular and Process Design	15
MultiObjective Computer-Aided Molecular Design	15
Computer-Aided Mixture Design	16
Computer-Aided Molecular Design Under Uncertainty	16
Computer-Aided Molecular, Process, and Control Design	16
Computer-Aided Molecular Design for Reactive Systems	18
Computer-Aided Molecular Design Using Quantum Mechanical Models	18
Summary of Classes	18
General Solution Approaches	19
Challenges in the Solution of Molecular Design Problems	19
Classification of Solution Approaches	19
Generate and Test Approaches	20
Deterministic Optimization Algorithms	21
Overview of main algorithms	21
Deterministic optimization approaches in molecular design	23
Stochastic Optimization Algorithms	25
Overview of main algorithmic concepts	25
Genetic algorithms (GA)	25
Simulated annealing	26
Tabu search	27
Ant colony optimization	27
Property Clustering Approaches	27
Merits and Shortcomings of Solution Approaches	28
Methodologies for Solution of CAMPD Problems	30
Generate and Test Approaches	30
Deterministic Optimization Approaches	30
Stochastic Optimization Approaches	33
Continuous Molecular Representations	34
Property Clustering and/or Reverse Formulations	36
Methodologies for Various Problem Classes	36
Multiobjective Optimization	36
Mixture and Blend Design	38
Molecular Design Under Uncertainty	43
Molecular, Process, and Control Design	46
Exploitation of Advanced Computing Infrastructures	47
Methodologies for Reactive Systems or QM-Based Approaches	47
Molecular Design	51

CAMD in reactive systems	51
QM models in reactive systems	52
QM models in nonreactive systems	53
Molecular and Process Design	54
Applications	55
Solvents for Industrial Separations, Reactive Separations, and Promotion of Reactions	56
Catalysts, Adsorbents, and Ionic Liquids	56
Materials for CO ₂ Capture	63
Heat Exchange Fluids	64
Polymers	64
Bio-Based Fuels and Chemicals	64
Formulated Products and Other Chemicals	66
Future Outlook and Further Reading	66
References	69
Further Reading	76

Nomenclature

A	Helmholtz free energy
A_k	Matrix of number of occurrences of molecular groups in molecule k
ACO	Ant colony optimization
ADI	Absorption-desorption index
AHP	Analytic hierarchy process
a_{mn}	Group-group interaction parameter used in calculation of γ_i^c
AUP^M	Augmented property index
AP	Augmented penalty
ASOG	Analytical Solution of Groups
BB	Branch and bound
BM	Big-M formulation
b_{mn}	Group-group interaction parameter for higher order approximation used in calculation of γ_i^c
C	Vector of concentration of components/molecules
CADD	Computer-aided drug design
CAM^pD	Computer-aided mixture design
CAMD	Computer-aided molecular design
CAMP _{CD}	Computer-aided molecular, process and control design
CAMP _D	Computer-aided molecular and process design
$CAM^R D$	Computer-aided molecular design for reactions
$CAM^u D$	Computer-aided molecular design under uncertainty
CAOS	Computer-aided organic synthesis
C_j	Cluster for property j
c_k	Concentration of k^{th} component/molecule
c_{mn}	Group-group interaction parameter for higher order approximation used in calculation of γ_i^c
CoMT	Continuous molecular targeting
COSMO	Conductor-like Screening Model
D	Vector of design variables
DEA	Data envelopment analysis
DFO	Derivative-free optimization
DFT	Density functional theory
d_i	i^{th} design variable
D_j	Contributions of second-order groups
E	Vector of disturbances
EACO	Efficient ACO
EBS	Environmentally benign solvent
E_i	Contributions of third-order-groups
EoS	Equation of state
ER	Equality relaxation
F	Vector of objective functions
F_j	j^{th} objective function
GA	Genetic algorithms
GBD	Generalized Benders decomposition
GC	Group contributions