# John B O Mitchell

List of Publications by Year
in descending order



| 113 | 6,119 | 76326 | 74163 |
|---|---|---|---|
| papers | citations | 40 h-index | 75 g-index |

| 117 | 117 | 117 | 6750 |
|---|---|---|---|
| all docs | docs citations | times ranked | citing authors |

| # | Article | IF | Citations |
|---|---|---|---|
| 1 | A machine learning approach to predicting proteinâ€“ligand binding affinity with applications to molecular docking. Bioinformatics, 2010, 26, 1169-1175. | 4.1 | 619 |
| 2 | A review of methods for the calculation of solution free energies and the modelling of systems in solution. Physical Chemistry Chemical Physics, 2015, 17, 6174-6191. | 2.8 | 389 |
| 3 | Machine learning methods in chemoinformatics. Wiley Interdisciplinary Reviews: Computational Molecular Science, 2014, 4, 468-481. | 14.6 | 331 |
| 4 | Amino/Aromatic Interactions in Proteins: Is the Evidence Stacked Against Hydrogen Bonding?. Journal of Molecular Biology, 1994, 239, 315-331. | 4.2 | 319 |
| 5 | Random Forest Models To Predict Aqueous Solubility. Journal of Chemical Information and Modeling, 2007, 47, 150-158. | 5.4 | 277 |
| 6 | Protein folds and functions. Structure, 1998, 6, 875-884. | 3.3 | 207 |
| 7 | BLEEP?potential of mean force describing protein-ligand interactions: I. Generating potential. Journal of Computational Chemistry, 1999, 20, 1165-1176. | 3.3 | 194 |
| 8 | Why Are Some Properties More Difficult To Predict than Others? A Study of QSPR Models of Solubility, Melting Point, and Log P. Journal of Chemical Information and Modeling, 2008, 48, 220-232. | 5.4 | 165 |
| 9 | Melting Point Prediction Employing k-Nearest Neighbor Algorithms and Genetic Parameter Optimization. Journal of Chemical Information and Modeling, 2006, 46, 2412-2422. | 5.4 | 154 |
| 10 | In Silico Target Predictions: Defining a Benchmarking Data Set and Comparison of Performance of the Multiclass NaÃ¯ve Bayes and Parzen-Rosenblatt Window. Journal of Chemical Information and Modeling, 2013, 53, 1957-1966. | 5.4 | 131 |
| 11 | Protein Recognition of Adenylate: An Example of a Fuzzy Recognition Template. Journal of Molecular Biology, 1996, 263, 486-500. | 4.2 | 125 |
| 12 | Understanding the Functional Roles of Amino Acid Residues in Enzyme Catalysis. Journal of Molecular Biology, 2009, 390, 560-577. | 4.2 | 117 |
| 13 | BLEEP?potential of mean force describing protein-ligand interactions: II. Calculation of binding energies and comparison with experimental data. Journal of Computational Chemistry, 1999, 20, 1177-1185. | 3.3 | 112 |
| 14 | The nature of the N ? H?O?C hydrogen bond: An intermolecular perturbation theory study of the formamide/formaldehyde complex. Journal of Computational Chemistry, 1990, 11, 1217-1233. | 3.3 | 104 |
| 15 | Predicting Intrinsic Aqueous Solubility by a Thermodynamic Cycle. Molecular Pharmaceutics, 2008, 5, 266-279. | 4.6 | 104 |
| 16 | Towards an understanding of the arginine-aspartate interaction. Journal of Molecular Biology, 1992, 226, 251-262. | 4.2 | 103 |
| 17 | Protein Ligand Database (PLD): additional understanding of the nature and specificity of protein-ligand complexes. Bioinformatics, 2003, 19, 1856-1857. | 4.1 | 96 |
| 18 | The Determination of the Crystal Structure of Anhydrous Theophylline by X-ray Powder Diffraction with a Systematic Search Algorithm, Lattice Energy Calculations, and13C and15N Solid-State NMR:Â A Question of Polymorphism in a Given Unit Cell. Journal of Physical Chemistry B, 2001, 105, 5818-5826. | 2.6 | 92 |

| # | Article | IF | Citations |
|---|---|---|---|
| 19 | Ligand-Target Prediction Using Winnow and Naive Bayesian Algorithms and the Implications of Overall Performance Statistics. Journal of Chemical Information and Modeling, 2008, 48, 2313-2325. | 5.4 | 92 |
| 20 | Is Experimental Data Quality the Limiting Factor in Predicting the Aqueous Solubility of Druglike Molecules?. Molecular Pharmaceutics, 2014, 11, 2962-2972. | 4.6 | 91 |
| 21 | First-Principles Calculation of the Intrinsic Aqueous Solubility of Crystalline Druglike Molecules. Journal of Chemical Theory and Computation, 2012, 8, 3322-3337. | 5.3 | 84 |
| 22 | Gaussian multipoles in practice: Electrostatic energies for intermolecular potentials. Journal of Computational Chemistry, 1994, 15, 1187-1198. | 3.3 | 77 |
| 23 | Hierarchical virtual screening for the discovery of new molecular scaffolds in antibacterial hit identification. Journal of the Royal Society Interface, 2012, 9, 3196-3207. | 3.4 | 68 |
| 24 | Uniting Cheminformatics and Chemical Theory To Predict the Intrinsic Aqueous Solubility of Crystalline Druglike Molecules. Journal of Chemical Information and Modeling, 2014, 54, 844-856. | 5.4 | 68 |
| 25 | On the relative strengths of amideâ€¦amide and amideâ€¦water hydrogen bonds. Chemical Physics Letters, 1991, 180, 517-523. | 2.6 | 64 |
| 26 | MACiE (Mechanism, Annotation and Classification in Enzymes): novel tools for searching catalytic mechanisms. Nucleic Acids Research, 2007, 35, D515-D520. | 14.5 | 64 |
| 27 | On the electrostatic directionality of Nî—¸Hâ€¦Oî—»C hydrogen bonding. Chemical Physics Letters, 1989, 154, 267-272. | 2.6 | 63 |
| 28 | Computational toxicology: an overview of the sources of data and of modelling methods. Expert Opinion on Drug Metabolism and Toxicology, 2009, 5, 1-14. | 3.3 | 63 |
| 29 | Drug Design for CNS Diseases: Polypharmacological Profiling of Compounds Using Cheminformatic, 3D-QSAR and Virtual Screening Methodologies. Frontiers in Neuroscience, 2016, 10, 265. | 2.8 | 62 |
| 30 | Can we predict lattice energy from molecular structure?. Acta Crystallographica Section B: Structural Science, 2003, 59, 676-685. | 1.8 | 58 |
| 31 | Electrostatic factors in DNA intercalation. Biopolymers, 1999, 52, 84-93. | 2.4 | 53 |
| 32 | D-amino acid residues in peptides and proteins. Proteins: Structure, Function and Bioinformatics, 2003, 50, 563-571. | 2.6 | 53 |
| 33 | Amino/aromatic interactions. Nature, 1993, 366, 413-413. | 27.8 | 50 |
| 34 | Predicting proteinâ€“ligand binding affinities: a low scoring game?. Organic and Biomolecular Chemistry, 2004, 2, 3267-3273. | 2.8 | 49 |
| 35 | Predicting Phospholipidosis Using Machine Learning. Molecular Pharmaceutics, 2010, 7, 1708-1714. | 4.6 | 49 |
| 36 | Predicting the mechanism of phospholipidosis. Journal of Cheminformatics, 2012, 4, 2. | 6.1 | 49 |

| # | Article | IF | Citations |
|---|---|---|---|
| 37 | MACiE: a database of enzyme reaction mechanisms. Bioinformatics, 2005, 21, 4315-4316. | 4.1 | 47 |
| 38 | Comments on â€œLeave-Cluster-Out Cross-Validation Is Appropriate for Scoring Functions Derived from Diverse Protein Data Setsâ€: Significance for the Validation of Scoring Functions. Journal of Chemical Information and Modeling, 2011, 51, 1739-1741. | 5.4 | 47 |
| 39 | Can human experts predict solubility better than computers?. Journal of Cheminformatics, 2017, 9, 63. | 6.1 | 46 |
| 40 | The Chemistry of Protein Catalysis. Journal of Molecular Biology, 2007, 372, 1261-1277. | 4.2 | 43 |
| 41 | The Relationship between the Sequence Identities of Alpha Helical Proteins in the PDB and the Molecular Similarities of Their Ligands. Journal of Chemical Information and Computer Sciences, 2001, 41, 1617-1622. | 2.8 | 40 |
| 42 | Support vector inductive logic programming outperforms the naive Bayes classifier and inductive logic programming for the classification of bioactive chemical compounds. Journal of Computer-Aided Molecular Design, 2007, 21, 269-280. | 2.9 | 40 |
| 43 | Using Reaction Mechanism to Measure Enzyme Similarity. Journal of Molecular Biology, 2007, 368, 1484-1499. | 4.2 | 39 |
| 44 | A Systematic Nonempirical Method of Deriving Model Intermolecular Potentials for Organic Molecules:Â Application To Amides. Journal of Physical Chemistry A, 2000, 104, 10958-10971. | 2.5 | 38 |
| 45 | Non-randomness in side-chain packing: the distribution of interplanar angles. , 1997, 29, 370-380. | | 37 |
| 46 | Development and Comparison of hERG Blocker Classifiers: Assessment on Different Datasets Yields Markedly Different Results. Molecular Informatics, 2011, 30, 443-458. | 2.5 | 37 |
| 47 | Quantitative Comparison of Catalytic Mechanisms and Overall Reactions in Convergently Evolved Enzymes: Implications for Classification of Enzyme Function. PLoS Computational Biology, 2010, 6, e1000700. | 3.2 | 33 |
| 48 | A novel hybrid ultrafast shape descriptor method for use in virtual screening. Chemistry Central Journal, 2008, 2, 3. | 2.6 | 32 |
| 49 | Multiple Solution Conformations of the Integrin-Binding Cyclic Pentapeptide Cyclo(-Ser-d-Leu-Asp-Val-Pro-). Analysis of the (phi,psi) Space Available to Cyclic Pentapeptides. FEBS Journal, 1996, 242, 352-362. | 0.2 | 31 |
| 50 | Chemoinformatics-Based Classification of Prohibited Substances Employed for Doping in Sport. Journal of Chemical Information and Modeling, 2006, 46, 2369-2380. | 5.4 | 30 |
| 51 | Characterizing the complexity of enzymes on the basis of their mechanisms and structures with a bioâ€computational analysis. FEBS Journal, 2011, 278, 3835-3845. | 4.7 | 30 |
| 52 | The Natural History of Biocatalytic Mechanisms. PLoS Computational Biology, 2014, 10, e1003642. | 3.2 | 30 |
| 53 | Design, synthesis and structure of a zinc finger with an artificial Î²-turn. Journal of Molecular Biology, 1998, 279, 973-986. | 4.2 | 29 |
| 54 | A structureâ€“odour relationship study using EVA descriptors and hierarchical clustering. Organic and Biomolecular Chemistry, 2004, 2, 3250-3255. | 2.8 | 28 |

| # | Article | IF | Citations |
|---|---|---|---|
| 55 | Are the Sublimation Thermodynamics of Organic Molecules Predictable?. Journal of Chemical Information and Modeling, 2016, 56, 2162-2179. | 5.4 | 28 |
| 56 | A Random Forest Model for Predicting Allosteric and Functional Sites on Proteins. Molecular Informatics, 2016, 35, 125-135. | 2.5 | 27 |
| 57 | Simultaneous feature selection and parameter optimisation using an artificial ant colony: case study of melting point prediction. Chemistry Central Journal, 2008, 2, 21. | 2.6 | 26 |
| 58 | A comparison of three theoretical approaches to the study of side-chain interactions in proteins. Journal of the Chemical Society, Faraday Transactions, 1993, 89, 2619. | 1.7 | 25 |
| 59 | Classifying Molecules Using a Sparse Probabilistic Kernel Binary Classifier. Journal of Chemical Information and Modeling, 2011, 51, 1539-1544. | 5.4 | 24 |
| 60 | Predicting Melting Points of Organic Molecules: Applications to Aqueous Solubility Prediction Using the General Solubility Equation. Molecular Informatics, 2015, 34, 715-724. | 2.5 | 24 |
| 61 | One origin for metallo-Î²-lactamase activity, or two? An investigation assessing a diverse set of reconstructed ancestral sequences based on a sample of phylogenetic trees. Journal of Molecular Evolution, 2014, 79, 117-129. | 1.8 | 23 |
| 62 | Anisotropic Repulsion Potentials for Cyanuric Chloride (C3N3Cl3) and Their Application to Modeling the Crystal Structures of Azaaromatic Chlorides. Journal of Physical Chemistry A, 2001, 105, 9961-9971. | 2.5 | 22 |
| 63 | Evaluation of a knowledge-based potential of mean force for scoring docked protein-ligand complexes. Journal of Computational Chemistry, 2001, 22, 673-688. | 3.3 | 22 |
| 64 | Enzyme Informatics. Current Topics in Medicinal Chemistry, 2012, 12, 1911-1923. | 2.1 | 20 |
| 65 | Rational Drug Design of Antineoplastic Agents Using 3D-QSAR, Cheminformatic, and Virtual Screening Approaches. Current Medicinal Chemistry, 2019, 26, 3874-3889. | 2.4 | 20 |
| 66 | 4273Ï€: Bioinformatics education on low cost ARM hardware. BMC Bioinformatics, 2013, 14, 243. | 2.6 | 19 |
| 67 | How To Winnow Actives from Inactives:â€‰ Introducing Molecular Orthogonal Sparse Bigrams (MOSBs) and Multiclass Winnow. Journal of Chemical Information and Modeling, 2008, 48, 306-318. | 5.4 | 17 |
| 68 | From sequence to enzyme mechanism using multi-label machine learning. BMC Bioinformatics, 2014, 15, 150. | 2.6 | 17 |
| 69 | Is EC class predictable from reaction mechanism?. BMC Bioinformatics, 2012, 13, 60. | 2.6 | 16 |
| 70 | Predicting targets of compounds against neurological diseases using cheminformatic methodology. Journal of Computer-Aided Molecular Design, 2015, 29, 183-198. | 2.9 | 16 |
| 71 | Communication and re-use of chemical information in bioscience. BMC Bioinformatics, 2005, 6, 180. | 2.6 | 15 |
| 72 | PFClust: a novel parameter free clustering algorithm. BMC Bioinformatics, 2013, 14, 213. | 2.6 | 15 |

| # | Article | IF | Citations |
|---|---|---|---|
| 73 | Toward Physics-Based Solubility Computation for Pharmaceuticals to Rival Informatics. Journal of Chemical Theory and Computation, 2021, 17, 3700-3709. | 5.3 | 15 |
| 74 | Theoretical Study of the Reaction Mechanism of Streptomyces coelicolor Type II Dehydroquinase. Journal of Chemical Theory and Computation, 2009, 5, 1284-1294. | 5.3 | 14 |
| 75 | Full â€œLaplacianisedâ€•posterior naive Bayesian algorithm. Journal of Cheminformatics, 2013, 5, 37. | 6.1 | 14 |
| 76 | Random graphs with arbitrary clustering and their applications. Physical Review E, 2021, 103, 012309. | 2.1 | 14 |
| 77 | SATIS:â€‰ Atom Typing from Chemical Connectivity. Journal of Chemical Information and Computer Sciences, 1999, 39, 751-757. | 2.8 | 13 |
| 78 | Enzyme function and its evolution. Current Opinion in Structural Biology, 2017, 47, 151-156. | 5.7 | 13 |
| 79 | Chemistry in bioinformatics. BMC Bioinformatics, 2005, 6, 141. | 2.6 | 12 |
| 80 | The Geometry of Interactions between Catalytic Residues and their Substrates. Journal of Molecular Biology, 2007, 369, 1140-1152. | 4.2 | 12 |
| 81 | Scoring functions and enrichment: a case study on Hsp90. BMC Bioinformatics, 2007, 8, 27. | 2.6 | 11 |
| 82 | Informatics, machine learning and computational medicinal chemistry. Future Medicinal Chemistry, 2011, 3, 451-467. | 2.3 | 11 |
| 83 | Predicting the protein targets for athletic performance-enhancing substances. Journal of Cheminformatics, 2013, 5, 31. | 6.1 | 11 |
| 84 | Probing the average distribution of water in organic hydrate crystal structures with radial distribution functions (RDFs). CrystEngComm, 2017, 19, 641-652. | 2.6 | 11 |
| 85 | Knowledge Based Potentials: the Reverse Boltzmann Methodology, Virtual Screening and Molecular Weight Dependence. QSAR and Combinatorial Science, 2005, 24, 527-536. | 1.4 | 10 |
| 86 | Greedy and Linear Ensembles of Machine Learning Methods Outperform Single Approaches for QSPR Regression Problems. Molecular Informatics, 2015, 34, 634-647. | 2.5 | 10 |
| 87 | Toxicological relationships between proteins obtained from protein target predictions of large toxicity databases. Toxicology and Applied Pharmacology, 2008, 231, 225-234. | 2.8 | 9 |
| 88 | Artificial intelligence in pharmaceutical research and development. Future Medicinal Chemistry, 2018, 10, 1529-1531. | 2.3 | 9 |
| 89 | Applications of crystal structure prediction â€“ organic molecular structures: general discussion. Faraday Discussions, 2018, 211, 493-539. | 3.2 | 8 |
| 90 | Diisophorone and related compounds. Part 19 Synthesis and reactions of 6,8-dibromodiisophorones. Monatshefte FÃ¼r Chemie, 1988, 119, 195-213. | 1.8 | 7 |

| # | Article | IF | Citations |
|---|---|---|---|
| 91 | Crystal structure evaluation: calculating relative stabilities and other criteria: general discussion. Faraday Discussions, 2018, 211, 325-381. | 3.2 | 7 |
| 92 | Three machine learning models for the 2019 Solubility Challenge. ADMET and DMPK, 2020, 8, 215-250. | 2.1 | 7 |
| 93 | Two-pathogen model with competition on clustered networks. Physical Review E, 2021, 103, 062308. | 2.1 | 7 |
| 94 | Percolation in random graphs with higher-order clustering. Physical Review E, 2021, 103, 012313. | 2.1 | 7 |
| 95 | The Parzen Window method: In terms of two vectors and one matrix. Pattern Recognition Letters, 2015, 63, 30-35. | 4.2 | 6 |
| 96 | Applications of crystal structure prediction â€“ inorganic and network structures: general discussion. Faraday Discussions, 2018, 211, 613-642. | 3.2 | 6 |
| 97 | PFClust: an optimised implementation of a parameter-free clustering algorithm. Source Code for Biology and Medicine, 2014, 9, 5. | 1.7 | 5 |
| 98 | Verifying the fully â€œLaplacianisedâ€•posterior NaÃ¯ve Bayesian approach and more. Journal of Cheminformatics, 2015, 7, 27. | 6.1 | 5 |
| 99 | Exact formula for bond percolation on cliques. Physical Review E, 2021, 104, 024304. | 2.1 | 5 |
| 100 | Multipole-based calculation of the polarization energy. Theoretica Chimica Acta, 1996, 94, 287-295. | 0.8 | 4 |
| 101 | Why do Sequence Signatures Predict Enzyme Mechanism? Homology versus Chemistry. Evolutionary Bioinformatics, 2015, 11, EBO.S31482. | 1.2 | 4 |
| 102 | Allosteric Inhibition of Acinetobacter baumannii ATP Phosphoribosyltransferase by Protein:Dipeptide and Protein:Protein Interactions. ACS Infectious Diseases, 2022, 8, 197-209. | 3.8 | 4 |
| 103 | Triazinone tautomers: solid phase energetics. CrystEngComm, 2003, 5, 498. | 2.6 | 3 |
| 104 | Cooperative coinfection dynamics on clustered networks. Physical Review E, 2021, 103, 042307. | 2.1 | 3 |
| 105 | Modelling the interactions of protein side-chains. Molecular Engineering, 1995, 5, 89-105. | 0.2 | 2 |
| 106 | Winnow based identification of potent hERG inhibitors in silico: comparative assessment on different datasets. Journal of Cheminformatics, 2012, 4, . | 6.1 | 2 |
| 107 | Symbiotic and antagonistic disease dynamics on networks using bond percolation. Physical Review E, 2021, 104, 024303. | 2.1 | 2 |
| 108 | Degree correlations in graphs with clique clustering. Physical Review E, 2022, 105, 044314. | 2.1 | 2 |

| # | Article | IF | Citations |
|---|---|---|---|
| 109 | Enzyme mechanism prediction: a template matching problem on InterPro signature subspaces. BMC Research Notes, 2015, 8, 744. | 1.4 | 1 |
| 110 | 3. In Silico methods to predict solubility. , 2019, , 71-112. | | 1 |
| 111 | Modelling the Interactions of Protein Side-Chains. Jerusalem Symposia on Quantum Chemistry and Biochemistry, 1995, , 119-135. | 0.2 | 1 |
| 112 | A note on utilising binary features as ligand descriptors. Journal of Cheminformatics, 2015, 7, 58. | 6.1 | 0 |
| 113 | A Bayesian network structure learning approach to identify genes associated with stress in spleens of chickens. Scientific Reports, 2022, 12, 7482. | 3.3 | 0 |