John P Overington

List of Publications by Citations

103 papers	16,851 citations	44 h-index	110 g-index
110 ext. papers	19,793 ext. citations	12.6 avg, IF	6.4 L-index

#	Paper	IF	Citations
103	How many drug targets are there?. *Nature Reviews Drug Discovery*, **2006**, 5, 993-6	64.1	2624
102	ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, **2012**, 40, D1100-7	20.1	2257
101	The ChEMBL database in 2017. *Nucleic Acids Research*, **2017**, 45, D945-D954	20.1	1059
100	The ChEMBL bioactivity database: an update. *Nucleic Acids Research*, **2014**, 42, D1083-90	20.1	1052
99	A comprehensive map of molecular drug targets. *Nature Reviews Drug Discovery*, **2017**, 16, 19-34	64.1	1032
98	The genome of the blood fluke Schistosoma mansoni. *Nature*, **2009**, 460, 352-8	50.4	822
97	An atlas of genetic influences on human blood metabolites. *Nature Genetics*, **2014**, 46, 543-550	36.3	695
96	The promise and peril of chemical probes. *Nature Chemical Biology*, **2015**, 11, 536-41	11.7	523
95	X-ray analysis of HIV-1 proteinase at 2.7 A resolution confirms structural homology among retroviral enzymes. *Nature*, **1989**, 342, 299-302	50.4	428
94	HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Science*, **1998**, 7, 2469-71	6.3	421
93	Can we rationally design promiscuous drugs?. *Current Opinion in Structural Biology*, **2006**, 16, 127-36	8.1	421
92	JOY: protein sequence-structure representation and analysis. *Bioinformatics*, **1998**, 14, 617-23	7.2	356
91	Probing the links between in vitro potency, ADMET and physicochemical parameters. *Nature Reviews Drug Discovery*, **2011**, 10, 197-208	64.1	333
90	A structural basis for sequence comparisons. An evaluation of scoring methodologies. *Journal of Molecular Biology*, **1993**, 233, 716-38	6.5	266
89	Derivation of rules for comparative protein modeling from a database of protein structure alignments. *Protein Science*, **1994**, 3, 1582-96	6.3	261
88	Genomic-scale prioritization of drug targets: the TDR Targets database. *Nature Reviews Drug Discovery*, **2008**, 7, 900-7	64.1	244
87	PSICQUIC and PSISCORE: accessing and scoring molecular interactions. *Nature Methods*, **2011**, 8, 528-9	21.6	227

86	ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Research*, **2015**, 43, W612-20	20.1	215
85	The druggable genome and support for target identification and validation in drug development. *Science Translational Medicine*, **2017**, 9,	17.5	212
84	18th Sir Hans Krebs lecture. Knowledge-based protein modelling and design. *FEBS Journal*, **1988**, 172, 513-20		211
83	Comprehensive characterization of the Published Kinase Inhibitor Set. *Nature Biotechnology*, **2016**, 34, 95-103	44.5	191
82	Alignment and searching for common protein folds using a data bank of structural templates. *Journal of Molecular Biology*, **1993**, 231, 735-52	6.5	157
81	Unexplored therapeutic opportunities in the human genome. *Nature Reviews Drug Discovery*, **2018**, 17, 317-332	64.1	156
80	From comparisons of protein sequences and structures to protein modelling and design. *Trends in Biochemical Sciences*, **1990**, 15, 235-40	10.3	134
79	Modeling alpha-helical transmembrane domains: the calculation and use of substitution tables for lipid-facing residues. *Protein Science*, **1993**, 2, 55-70	6.3	127
78	Fragment ranking in modelling of protein structure. Conformationally constrained environmental amino acid substitution tables. *Journal of Molecular Biology*, **1993**, 229, 194-220	6.5	110
77	SureChEMBL: a large-scale, chemically annotated patent document database. *Nucleic Acids Research*, **2016**, 44, D1220-8	20.1	102
76	UniChem: a unified chemical structure cross-referencing and identifier tracking system. *Journal of Cheminformatics*, **2013**, 5, 3	8.6	90
75	Ligand efficiency indices for an effective mapping of chemico-biological space: the concept of an atlas-like representation. *Drug Discovery Today*, **2010**, 15, 804-11	8.8	84
74	Activity, assay and target data curation and quality in the ChEMBL database. *Journal of Computer-Aided Molecular Design*, **2015**, 29, 885-96	4.2	79
73	New open drug activity data at EBI. *Chemistry Central Journal*, **2009**, 3,		78
72	Comparative modelling of major house dust mite allergen Der p I: structure validation using an extended environmental amino acid propensity table. *Protein Engineering, Design and Selection*, **1994**, 7, 869-94	1.9	75
71	Minimum information about a bioactive entity (MIABE). *Nature Reviews Drug Discovery*, **2011**, 10, 661-9	64.1	69
70	Computational and Practical Aspects of Drug Repositioning. *Assay and Drug Development Technologies*, **2015**, 13, 299-306	2.1	64
69	Chemical, target, and bioactive properties of allosteric modulation. *PLoS Computational Biology*, **2014**, 10, e1003559	5	63

68	Improved large-scale prediction of growth inhibition patterns using the NCI60 cancer cell line panel. *Bioinformatics*, **2016**, 32, 85-95	7.2	60
67	PCSK9 monoclonal antibodies for the primary and secondary prevention of cardiovascular disease. *The Cochrane Library*, **2017**, 4, CD011748	5.2	57
66	Improving the odds of drug development success through human genomics: modelling study. *Scientific Reports*, **2019**, 9, 18911	4.9	54
65	The prediction and orientation of alpha-helices from sequence alignments: the combined use of environment-dependent substitution tables, Fourier transform methods and helix capping rules. *Protein Engineering, Design and Selection*, **1994**, 7, 645-53	1.9	53
64	Drug Target Commons: A Community Effort to Build a Consensus Knowledge Base for Drug-Target Interactions. *Cell Chemical Biology*, **2018**, 25, 224-229.e2	8.2	51
63	Benchmarking of protein descriptor sets in proteochemometric modeling (part 2): modeling performance of 13 amino acid descriptor sets. *Journal of Cheminformatics*, **2013**, 5, 42	8.6	48
62	Molecular recognition in protein families: a database of aligned three-dimensional structures of related proteins. *Biochemical Society Transactions*, **1993**, 21 (Pt 3), 597-604	5.1	47
61	Open Source Drug Discovery: Highly Potent Antimalarial Compounds Derived from the Tres Cantos Arylpyrroles. *ACS Central Science*, **2016**, 2, 687-701	16.8	44
60	Nicastrin, a presenilin-interacting protein, contains an aminopeptidase/transferrin receptor superfamily domain. *Trends in Biochemical Sciences*, **2001**, 26, 213-4	10.3	44
59	Antibody informatics for drug discovery. *Biochimica Et Biophysica Acta - Proteins and Proteomics*, **2014**, 1844, 2002-2015	4	43
58	Discrimination of common protein folds: application of protein structure to sequence/structure comparisons. *Methods in Enzymology*, **1996**, 266, 575-98	1.7	43
57	Target prediction for an open access set of compounds active against Mycobacterium tuberculosis. *PLoS Computational Biology*, **2013**, 9, e1003253	5	42
56	Comparison of three-dimensional structures of homologous proteins. *Current Opinion in Structural Biology*, **1992**, 2, 394-401	8.1	41
55	PDBLIG: classification of small molecular protein binding in the Protein Data Bank. *Journal of Medicinal Chemistry*, **2004**, 47, 3807-16	8.3	40
54	Global analysis of small molecule binding to related protein targets. *PLoS Computational Biology*, **2012**, 8, e1002333	5	35
53	Artificial intelligence, drug repurposing and peer review. *Nature Biotechnology*, **2020**, 38, 1127-1131	44.5	35
52	Unprecedently Large-Scale Kinase Inhibitor Set Enabling the Accurate Prediction of Compound-Kinase Activities: A Way toward Selective Promiscuity by Design?. *Journal of Chemical Information and Modeling*, **2016**, 56, 1654-75	6.1	35
51	Chemical databases: curation or integration by user-defined equivalence?. *Drug Discovery Today: Technologies*, **2015**, 14, 17-24	7.1	32

50	Mycobacterial dihydrofolate reductase inhibitors identified using chemogenomic methods and in vitro validation. *PLoS ONE*, **2015**, 10, e0121492	3.7	32
49	The relationship between target-class and the physicochemical properties of antibacterial drugs. *Bioorganic and Medicinal Chemistry*, **2015**, 23, 5218-24	3.4	31
48	Synthesis of macrocyclic, potential protease inhibitors using a generic scaffold. *Journal of Organic Chemistry*, **2002**, 67, 4882-92	4.2	31
47	Release of 50 new, drug-like compounds and their computational target predictions for open source anti-tubercular drug discovery. *PLoS ONE*, **2015**, 10, e0142293	3.7	30
46	diXa: a data infrastructure for chemical safety assessment. *Bioinformatics*, **2015**, 31, 1505-7	7.2	27
45	Annotating Human P-Glycoprotein Bioassay Data. *Molecular Informatics*, **2012**, 31, 599-609	3.8	27
44	Collation and data-mining of literature bioactivity data for drug discovery. *Biochemical Society Transactions*, **2011**, 39, 1365-70	5.1	27
43	UniChem: extension of InChI-based compound mapping to salt, connectivity and stereochemistry layers. *Journal of Cheminformatics*, **2014**, 6, 43	8.6	25
42	Structural and Functional View of Polypharmacology. *Scientific Reports*, **2017**, 7, 10102	4.9	24
41	A drug target slim: using gene ontology and gene ontology annotations to navigate protein-ligand target space in ChEMBL. *Journal of Biomedical Semantics*, **2016**, 7, 59	2.2	21
40	Prioritizing the proteome: identifying pharmaceutically relevant targets. *Drug Discovery Today*, **2002**, 7, 516-21	8.8	21
39	Design of selective thrombin inhibitors based on the (R)-Phe-Pro-Arg sequence. *Journal of Medicinal Chemistry*, **2002**, 45, 2432-53	8.3	21
38	Chemogenomics approaches for receptor deorphanization and extensions of the chemogenomics concept to phenotypic space. *Current Topics in Medicinal Chemistry*, **2011**, 11, 1964-77	3	18
37	A ligand's-eye view of protein similarity. *Nature Methods*, **2013**, 10, 116-7	21.6	16
36	Identification of Allosteric Modulators of Metabotropic Glutamate 7 Receptor Using Proteochemometric Modeling. *Journal of Chemical Information and Modeling*, **2017**, 57, 2976-2985	6.1	16
35	Insights into protein function through large-scale computational analysis of sequence and structure. *Trends in Biotechnology*, **2001**, 19, 61-66	15.1	16
34	myChEMBL: a virtual machine implementation of open data and cheminformatics tools. *Bioinformatics*, **2014**, 30, 298-300	7.2	15
33	Rapid analysis of pharmacology for infectious diseases. *Current Topics in Medicinal Chemistry*, **2011**, 11, 1292-300	3	15

32	Target Identification of Phenotypic Hits Using a Concerted Chemogenomic, Biophysical, and Structural Approach. *Frontiers in Pharmacology*, **2017**, 8, 681	5.6	14
31	The EBI enzyme portal. *Nucleic Acids Research*, **2013**, 41, D773-80	20.1	13
30	Cheminformatics. *Communications of the ACM*, **2012**, 55, 65-75	2.5	13
29	Insights into protein function through large-scale computational analysis of sequence and structure. *Trends in Biotechnology*, **2001**, 19, S61-6	15.1	13
28	A large-scale crop protection bioassay data set. *Scientific Data*, **2015**, 2, 150032	8.2	12
27	Scientific Lenses to Support Multiple Views over Linked Chemistry Data. *Lecture Notes in Computer Science*, **2014**, 98-113	0.9	12
26	MyChEMBL: A Virtual Platform for Distributing Cheminformatics Tools and Open Data. *Challenges*, **2014**, 5, 334-337	3.4	11
25	Mapping small molecule binding data to structural domains. *BMC Bioinformatics*, **2012**, 13 Suppl 17, S11	3.6	11
24	Protein sequence analysis in silico: application of structure-based bioinformatics to genomic initiatives. *Current Opinion in Pharmacology*, **2002**, 2, 574-80	5.1	11
23	ADME SARfari: comparative genomics of drug metabolizing systems. *Bioinformatics*, **2015**, 31, 1695-7	7.2	10
22	Rational design of non-resistant targeted cancer therapies. *Scientific Reports*, **2017**, 7, 46632	4.9	9
21	Classification and analysis of a large collection of in vivo bioassay descriptions. *PLoS Computational Biology*, **2017**, 13, e1005641	5	9
20	Setting Our Sights on Infectious Diseases. *ACS Infectious Diseases*, **2020**, 6, 3-13	5.5	9
19	PPDMs-a resource for mapping small molecule bioactivities from ChEMBL to Pfam-A protein domains. *Bioinformatics*, **2015**, 31, 776-8	7.2	8
18	A document classifier for medicinal chemistry publications trained on the ChEMBL corpus. *Journal of Cheminformatics*, **2014**, 6, 40	8.6	7
17	The functional therapeutic chemical classification system. *Bioinformatics*, **2014**, 30, 876-83	7.2	5
16	Brain: biomedical knowledge manipulation. *Bioinformatics*, **2013**, 29, 1238-9	7.2	5
15	Protein three-dimensional structural databases: domains, structurally aligned homologues and superfamilies. *Acta Crystallographica Section D: Biological Crystallography*, **1998**, 54, 1168-77		5

14	Transporter assays and assay ontologies: useful tools for drug discovery. *Drug Discovery Today: Technologies*, **2014**, 12, e47-54	7.1	4
13	The Molecular Basis of Predicting Druggability1315-1334		3
12	Chapter 28. Recent development in cheminformatics and chemogenomics. *Annual Reports in Medicinal Chemistry*, **2003**, 38, 285-294	1.6	3
11	Modelling of the lignin peroxidase LIII of Phlebia radiata: use of a sequence template generated from a 3-D structure. *Protein Engineering, Design and Selection*, **1993**, 6, 177-82	1.9	3
10	Towards predictive resistance models for agrochemicals by combining chemical and protein similarity via proteochemometric modelling. *Journal of Chemical Biology*, **2014**, 7, 119-23		2
9	ChEMBL Beaker: A Lightweight Web Framework Providing Robust and Extensible Cheminformatics Services. *Challenges*, **2014**, 5, 444-449	3.4	2
8	Pleiotropic Effects of Statins. *Annual Reports in Medicinal Chemistry*, **2004**, 39, 239-258	1.6	2
7	Chapter 19. Expanding and exploring cellular pathways for novel drug targets. *Annual Reports in Medicinal Chemistry*, **2002**, 37, 187-196	1.6	2
6	Knowledge-Based Protein Modelling: Human Plasma Kallikrein and Human Neutrophil Defensin **1990**, 567-574		2
5	Structural and Functional View of Polypharmacology		2
4	Repurposing vandetanib plus everolimus for the treatment of ACVR1-mutant diffuse intrinsic pontine glioma. *Cancer Discovery*, **2021**,	24.4	2
3	**1994**,		1
2	Symposium 1: Structure and engineering of proteins: New developments. *FreseniussJournal of Analytical Chemistry*, **1990**, 337, 1-3		
1	Knowledge-Based Protein Modeling and the Design of Novel Molecules **1990**, 209-227		