J Comput Aided Mol Des (2012) 26:153–157
DOI 10.1007/s10822-011-9492-3

PERSPECTIVE

The *Journal of Computer-Aided Molecular Design*: a bibliometric note

Peter Willett

Received: 20 October 2011 / Accepted: 8 November 2011 / Published online: 16 November 2011

Abstract Summarizes the articles in, and the citations to, volumes 2–24 of the *Journal of Computer-Aided Molecular Design*. The citations to the journal come from almost 2000 different sources that span a very wide range of academic subjects, with the most heavily cited articles being descriptions of software systems and of computational methods.

Keywords Bibliometrics · Citation analysis · Publication analysis

Introduction

The editorial of the very first issue of the *Journal of Computer-Aided Molecular Design* (hereafter JCAMD) stated that "a new discipline has emerged which has invaded almost all aspects of the study of molecular structure and recognition. Its applications range through modern medicinal chemistry, protein engineering, catalytic design, polymer studies and many other facets of organic, inorganic, physical and biological chemistry" [1]. The journal was established to provide a mechanism for research findings in this new discipline, and it has done this with conspicuous success for a quarter of a century. The first volume appeared in 1987 with this and the subsequent three volumes all containing four issues. The journal moved to having six issues a volume in 1991, to eight in 2000, and to its current monthly publication in the following year following its merger with the closely related review journal *Perspectives in Drug Discovery and Design*. The steady growth in the number of issues (and, consequently, in the number of articles, with 22 in the first volume and 91 in the most recently completed volume for 2010) demonstrates that the journal is successfully meeting the needs of an increasingly widespread academic community.

This brief communication uses the methods of bibliometrics [2, 3] to provide an overview of the articles that have appeared in JCAMD and, more importantly, of JCAMD's scientific impact as denoted by the citations to it from the broader scientific literature. Such bibliometric studies of journals have been reported in very many disciplines [4] including, recently, computer-aided drug discovery [5–7]. The data for the present study was obtained from the *Web of Science*, a database from Thomson Reuters that contains the publications in, and the citations to, ca. 10,000 of the world's leading academic journals and ca. 110,000 conference proceedings. It also contains tools that allow very rapid processing of large volumes of publication and citation data from the database, enabling analyses to be carried out in a few hours that would previously have taken many weeks of effort. The *Web of Science* has covered JCAMD since the start of volume 2 in 1988, and the analyses below are based on the articles that have appeared from then till the end of 2010 (i.e., up to and including volume 24). After excluding editorials, reviews, letters etc. a total of 1,175 articles appeared in these 23 volumes, and this corpus formed the basis for the discussion below.

Publications in the journal

The articles were the work of 4,081 authors, with these including many of the leading workers in the field of computer-aided drug discovery. Eleven authors each

P. Willett (✉)
Information School, University of Sheffield, 211 Portobello Street, Sheffield S1 4DP, UK
e-mail: p.willett@sheffield.ac.uk

contributed ten or more articles, these comprising the following: Clark, Dean (the most productive with no less than 28 contributions), Gasteiger, Höltje, Jain, Kuntz, Marshall, Murray, Oprea, Orozco, and Willett. Conversely, no less than 3,829 individuals made two or less contributions to the journal: this highly skewed distribution is known as Lotka's Law and is apparent in very many bibliometric analyses of publication data.

A skewed distribution is also apparent when one considers the 55 countries represented by the authors, where three countries dominate. These are the USA, the UK and Germany with 417, 217 and 132 contributions, respectively; taken together, these countries represent 64.8% of the total country data for the period. However, significant changes are taking place in the international organization of science. Thus, if one sub-divides the country data into the periods 1987–2000 and 2001–2010, then the first four countries are USA, UK, Germany and Spain (in that order) for both periods, with France, Italy, Sweden and Switzerland also appearing in the ten most productive countries for both periods. However, the dominance of the USA is far more pronounced in this century, and China and India have moved from joint-27th and joint-22nd positions to fifth and sixth positions, respectively. In addition to such expected countries, the journal's international scope is demonstrated by contributions from countries as diverse as Lithuania, Uruguay and Uzbekistan.

The authors came from 972 different institutions. Hardly surprisingly, most of these are academic institutions but it is interesting to note that no less than 17 of the 77 institutions that have contributed at least five articles are industrial in character. These 17 include both software companies (e.g., Molsoft, OpenEye and Tripos) and pharmaceutical companies (e.g., Abbott, Merck and Novartis), with several of the latter now defunct following the many mergers that have taken place in the pharmaceutical industry over the last few years (e.g., SmithKline Beecham, Wellcome and Wyeth Ayerst).

Citations to the journal

The articles attracted a total of 31,262 citations (as of late-September 2011) with an *h*-index of 71, i.e., 71 of the articles had attracted at least 71 citations [8]. JCAMD's impact is widespread with the citations coming from a total of 19,468 articles in 1944 different sources (overwhelmingly academic journals). The 20 journals that provided the greatest numbers of citations are listed in Table 1. Very crudely, one can divide the papers in JCAMD into two broad (albeit often overlapping) categories. The first category includes articles that describe a method of some sort (e.g., a matching algorithm for molecular alignments, or a novel criterion for evaluating the predictive power of a QSAR method) or a computer (normally software) system that has been developed to carry out some particular modelling task (e.g., identifying the hot-spots in a protein binding site, or computing melting points). The second category comprises articles that describe the practical use of a method or a system in a molecular design project (e.g., a QSAR analysis of a set of Factor Xa inhibitors, or the development of a pharmacophore model for a set of dopamine D_4 receptor antagonists).

It will be seen that Table 1 contains both chemoinformatics journals, which are likely to provide many of the citations to method/system articles, and more general chemical journals (albeit with the expected preponderance of medicinal and biological chemistry), which are likely to provide many of the citations to applications articles; both types of journal are likely to make extensive reference to

Table 1 The 20 journals that cite *Journal of Computer-Aided Molecular Design* most frequently

Journal	Citations	Impact factor
Journal of Medicinal Chemistry	1,154	5.18
Journal of Chemical Information And Modeling[1]	1,071	3.72
Journal of Computer Aided Molecular Design[2]	745	3.52
Bioorganic Medicinal Chemistry	481	3.11
Journal of Computational Chemistry	363	5.04
Journal of the American Chemical Society	356	8.98
Molecular Informatics[1,3]	294	1.87
Bioorganic Medicinal Chemistry Letters	281	2.59
Journal of Physical Chemistry A	277	2.84
Journal of Molecular Graphics and Modelling[1]	273	2.27
Journal of Physical Chemistry B	254	4.43
Journal Of Molecular Structure Theochem	226	1.21
Biochemistry	225	3.15
European Journal of Medicinal Chemistry	216	3.36
Proteins Structure Function and Bioinformatics	213	3.18
Journal of Biological Chemistry	207	3.14
Journal of Molecular Modeling	189	2.24
Journal of Molecular Biology	183	4.10
Journal of Chemical Physics	181	2.93
Current Medicinal Chemistry	147	4.96

[1] The totals for *Journal of Chemical Information and Modeling*, *Journal of Molecular Graphics and Modelling* and *Molecular Informatics* include citations from their predecessor journals, i.e., *Journal of Chemical Information and Computer Sciences*, *Journal of Molecular Graphics*, and both *QSAR* and *QSAR and Combinatorial Science*, respectively

[2] The total for *Journal of Computer-Aided Molecular Design* includes citations from *Perspectives in Drug Discovery and Design*

[3] The 5-year impact quoted for *Molecular Informatics* is that for its processor *QSAR and Combinatorial Science*, as a figure for the renamed journal is not yet available

the many system-description articles published in JCAMD if the "standard" reference (see below) is cited. The 5-year journal impact figures (i.e., the number of times that a typical article in a particular journal has been cited in a particular time-frame) are also listed in Table 1. This data comes from the Thomson Reuters *Journal Citation Reports* database, and demonstrates clearly the scientific quality of the journals that are drawing on work published in JCAMD.

The 1944 different sources cover no less than 170 different *Web of Science* subject categories, with the 20 most frequently occurring categories listed in Table 2. These categories make clear that JCAMD articles are of relevance not just to medicinal and biological chemists but also more broadly across the chemical sciences. Indeed, the journal's influence extends beyond chemistry, as the 150 categories not listed in the table are as diverse as Business Finance, Fisheries, Medical Ethics, and Sports Science, *inter alia*. An economic analogy has been noted by Cronin and Pearson for the crossing of subject boundaries by ideas, research and methods [9]. These authors suggest that if a citing journal is from a subject area that is different from that of the cited article then the citation can be regarded as an example of an export process. Viewed in this light, JCAMD can be seen to be exporting successfully to a range of other subject areas, hence increasing the visibility of the journal and of the discipline that it describes.

Table 2 The 20 Web of Science subject categories that are most frequently assigned to articles citing *Journal of Computer-Aided Molecular Design* articles

Web of Science subject category	Citations
Biochemistry Molecular Biology	4,851
Chemistry Multidisciplinary	3,979
Chemistry Medicinal	3,644
Computer Science Interdisciplinary Applications	2,785
Pharmacology Pharmacy	2,181
Chemistry Physical	2,158
Biophysics	2,016
Chemistry Organic	1,665
Computer Science Information Systems	1,114
Physics Atomic Molecular Chemical	1,109
Biochemical Research Methods	801
Mathematical Computational Biology	610
Crystallography	553
Biotechnology Applied Microbiology	427
Chemistry Inorganic Nuclear	418
Chemistry Analytical	388
Toxicology	366
Cell Biology	298
Mathematics Interdisciplinary Applications	291
Chemistry Applied	256

On average, each JCAMD article attracts 26.6 citations, but this is again a very skewed distribution in which just a few articles attract a very large amount of interest. The 20 most highly cited articles are listed in Table 3, and readers of this journal, whatever their specific area of research, are likely to be familiar with at least some of these. In all, the 20 articles attracted 9,301 citations, i.e., 29.8% of the total number of citations to the journal. As has been noted in previous studies [5, 7], descriptions of chemoinformatics methods or systems can rapidly attract very large numbers of citations if an article becomes the "standard" reference, i.e., the one that is normally cited when a particular algorithm or piece of software is used in a project. This is again the case here, with descriptions of DISCO, DOCK, GASP, LUDI, MOLDEN and MOPAC appearing amongst the 20 most-cited JCAMD articles that are listed in Table 3; descriptions of CAVEAT, CoMSIA, IsoStar and SPROUT *inter alia* appear as one moves further down the list of highly-cited articles. The extensive citation of method/system articles is hardly surprising since a method/system may have wide applicability both within and without the general modeling area, thus enabling it to attract citations from a wide range of disciplines. Interest in an applications article, conversely, is likely to be principally of interest to the molecular design community (or some part thereof), and such articles hence tend to have relatively lower citations rates.

Method/system articles may continue to be cited over a considerable period, even when new versions become available in the case of operational systems; for example, Stewart's MOPAC article was first published over two decades ago, in 1990, but has still attracted an average of over 50 citations *per annum* for the last 5 years. This extended citation lifetime further increases the citation yield of such articles whereas the information in an applications article may date quite quickly as new biological or chemical knowledge becomes available. Such long-term effects are lessened if we consider the 20 most-cited articles from amongst those that have been published over just the last 5 years, i.e., the period 2006–2010. Methods and systems are again well represented (e.g., descriptions of the systems GALAHAD and PHASE, and techniques for the evaluation and comparison of virtual screening experiments) but four of the articles focus on specific applications (e.g., homology modelling of the human histamine H-3 receptor, and QSAR modeling of tylophorine derivatives as anti-cancer agents). In conclusion, it should be noted that while method/system articles attract the most citations, it is applications that have always provided the principal focus of the majority of the articles published in the journal. Thus, if we consider the issues of

Table 3 The 20 most-cited articles from *Journal of Computer-Aided Molecular Design*

Article	Citations
Stewart, J.J.P. MOPAC—a semiempirical molecular-orbital program. JCAMD, 4 (1990) 1–45	1,958
Schaftenaar, G. and Noordik, J.H. Molden: a pre- and post-processing program for molecular and electronic structures. JCAMD, 14 (2000) 123–134	1,483
Eldridge, M.D. et al. Empirical scoring functions. 1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. JCAMD, 11 (1997) 425–445	648
Böhm, H.-J. The development of a simple empirical scoring function to estimate the binding constant for a protein ligand complex of known 3-dimensional structure. JCAMD, 8 (1994) 243–256	599
Morris, G.M. et al. Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. JCAMD, 10 (1996) 293–304	484
Ewing, T.J.A. et al. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. JCAMD, 15 (2001) 411–428	461
Böhm, H.-J. The computer-program LUDI—a new method for the denovo design of enzyme-inhibitors. JCAMD, 6 (1992) 61–78	377
McMartin, C. and Bohacek, R.S. QXP: Powerful, rapid computer algorithms for structure-based drug design. JCAMD, 11 (1997) 333–344.	347
Wang, R.X. et al. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. JCAMD, 16 (2002) 11–26.	306
Böhm, H.-J. LUDI—Rule-based automatic design of new substituents for enzyme-inhibitor leads. JCAMD, 6 (1992) 593–606.	290
Martin, Y.C. et al. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. JCAMD, 7 (1993) 83–102.	288
Tetko, I.V. et al. Virtual computational chemistry laboratory—design and description. JCAMD, 19 (2005) 453–463	274
Kellogg, G.E. et al. HINT—A new method of empirical hydrophobic field calculation for CoMFA. JCAMD, 5 (1991) 545–552	258
Gerber, P.R. and Müller, K. MAB, a generally applicable molecular-force field for structure modeling in medicinal chemistry. JCAMD, 9 (1995) 251–268	246
van Aalten, D.M. F. et al. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. JCAMD, 10 (1996) 255–262	241
Jones, G. et al. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. JCAMD, 9 (1995) 532–549	217
Bush, B.L. and Nachbar, R.B. Sample-distance partial least-squares—PLS optimized for many variables, with application to CoMFA. JCAMD, 7 (1993) 587–619	211
Storer, J.W. CLASS-IV charge models—a new semiempirical approach in quantum-chemistry. JCAMD, 9 (1995) 87–110	209
Brady, G.P. and Stouten, P.F.W. Fast prediction and visualization of protein binding pockets with PASS. JCAMD, 14 (2000) 383–401	207
Oprea, T.I. Property distribution of drug-related chemical databases. JCAMD, 14 (2000) 251–264	197

the journal published in 1990, 2000 and 2010 (i.e., volumes 4, 14 and 24) 63, 64 and 71% respectively of them were applications articles, as reflected by mention in the title of a specific class of structures and/or of a specific biological activity (these percentages are after excluding perspectives and special-issue articles).

Conclusions

In conclusion, the publication and citation data presented here demonstrate very clearly that JCAMD has met fully the objectives that the founding editors set themselves when the journal was established in 1987. The journal attracts high-quality articles that are of relevance not only to the international molecular modeling community but also more broadly across the chemical sciences and beyond. This is especially the case with articles describing computational methods and software systems, these complementing the applications studies that form the principal component of the journal.

References

1. Marshall GR, Vinter JG, Höltje H-D (1987) The Tower of Babel. J Comput-Aid Mol Des 1:1–2
2. Bar-Ilan J (2008) Informetrics at the start of the 21st century–a review. J Informet 2:1–52

3. Nicolaisen J (2007) Citation analysis. Ann Rev Inform Sci Technol 41:609–641
4. Anyi KWU, Zainab AN, Anuar NB (2009) Bibliometric studies on single journals: a review. Malay J Libr Inform Sci 14:17–55
5. Willett P (2007) A bibliometric analysis of the Journal of Molecular Graphics and Modelling. J Mol GraphModel 26:602–606
6. Willett P (2009) A bibliometric analysis of Quantitative Structure-Activity Relationships and QSAR & Combinatorial Science. QSAR Combin Sci 28:1231–1236
7. Al Jishi R, Willett P (2010) The Journal of Chemical Documentation and the Journal of Chemical Information and Computer Science: publication and citation statistics. J Chem Inform Model 50:1915–1923
8. Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Nat Acad Sci USA 102:16569–16572
9. Cronin B, Pearson S (1990) The export of ideas from information science. J Iform Sci 16:381–391