Bibliometric Analysis of Chinese Research on Cyclization, MALDI-TOF, and Antibiotics

Jiang Li

Department of Information Management, Nanjing University, China

Peter Willett*

Department of Information Studies, University of Sheffield, United Kingdom

Received August 26, 2009

This paper reports a bibliometric analysis of the impact of research in China on cyclization, MALDI-TOF, and antibiotics, comparing this research with that in the USA, Germany, and Japan. It is shown that the productivity of the Chinese research (in terms of numbers of publications) is growing rapidly; however, this growth has not, to date, been accompanied by an analogous growth in impact (in terms of citations to the published work). A citation analysis of national and international collaboration patterns shows that collaborative research does not invariably result in a larger number of citations; in part, at least, this is shown to be due to the dominant role played by the publications of the Chinese Academy of Sciences.

INTRODUCTION

Bibliometrics was first defined by Pritchard as "The application of mathematical and statistical methods to books and other media".[1] It involves the numeric analysis of data such as the authors of papers, places of publication, citations to a specific paper or papers, links between Web sites, collaborations between different authors or nations, the comparison of bibliometric and peer-review estimates of research, and the impact of different academic journals inter alia.[2−5] Many applications in chemistry have been reported,[6−15] with recent analyses focusing on individual academic departments,[16,17] individual journals,[18−21] 3D structure databases,[22,23] the literature of chemoinformatics in general,[24] and productivity in drug discovery[25] inter alia. Here, rather than a department, journal, or database, we report a bibliometric analysis of a country, specifically China, as reflected in its contributions to three important areas of chemistry.

Recent developments in China have enabled it to make an increasing contribution to academic knowledge.[26−28] However, the degree of growth seems to vary widely between fields, with Cyranoski, indeed, suggesting chemistry as one of several fields where Chinese research lacks impact,[29] and with Glänzel et al. ranking China 11th in a survey of the 32 most productive countries in chemistry research, based on publications from 1990 to 1998.[30] More recently, in 2005, the National Science Library of the Chinese Academy of Sciences [at http://www.cas.cn/ (in Chinese)] reported that China was the 11th most influential country in terms of the top 1% highly cited chemical papers in the *Science Citation Index*; in 2008, the National Science Foundation reported that China was the second most productive nation in engineering and chemistry in 2005;[31] and in 2009 the Institute of Scientific and Technical Information of China [at http://scitech.people.com.cn/GB/8504234.html (in Chinese)] noted that Chinese chemical publications comprised 12.5% of the chemical publications in the *Science Citation Index* from 1998 to 2008. At the disciplinary, macro level, then, it would seem that China is making an increasing contribution to chemical knowledge, but there has been less interest in the extent of the contribution at the micro level, where specific chemical topics are considered; to date, we have identified only a study by Kumari of the output and impact of 77 countries' research in synthetic organic chemistry, which noted that Chinese research in this area had increased linearly during the period 1998−2004.[32] This paper reports a more detailed study, focusing on important topics in medicinal chemistry (antibiotics), analytical chemistry (MALDI-TOF, i.e., matrix assisted laser desorption/ionization time-of-flight mass spectrometry), and synthetic organic chemistry (cyclization reactions).

METHODS

Data Sources. This study is based on publication and citation data extracted from Thomson Reuters' Web of Knowledge (WOK) database, which contains bibliographic data for, and citations to, publications in over 10 000 of the world's most important academic journals dating back to 1900; the database has also recently started to include data for important conference proceedings.

Searches for publications on cyclization during the period 1900−2008 were performed using the search statement

Topic = (CYCLI?ATION*) AND Address = (CHINA)

In this query, the asterisk (*) represents any group of characters or no character, the question mark (?) represents any single character, and "Topic" refers to at least one of the fields Title, Abstract, Author Keywords, and Keywords Plus. The search hence retrieves outputs on cyclization, cyclisations etc. produced from 1900 to 2008 and published by at least one institution with a Chinese address. The two other sets of publications were generated by using MALDI-TOF or ANTIBIOTIC* in the query statement above. All

* Corresponding author. Telephone: 0044-114-2222633. E-mail: p.willett@sheffield.ac.uk.

10.1021/ci9003199
Published on Web 12/23/2009

of the WOK searches were carried out in April 2009, as were analogous sets of searches for publications with at least one address in Germany, Japan, and the USA. These nations were chosen since they all publish extensively in these three areas. This was demonstrated by carrying out searches without the address criterion and then ranking in decreasing order of appearance all of the nations represented in each of the search outputs: the USA was ranked first in all three areas; Germany was ranked second in MALFI-TOF and antibiotics, and third in cyclization; and Japan was ranked second, fourth, and fifth in cyclization, MALDI-TOF, and antibiotics. The corresponding ranks for China were seventh, third and 16th, respectively. It is worth noting the comprehensiveness of the country data here: only 3.3% of the WOK records for cyclization did not have country data, with the corresponding figures for MALDI-TOF and antibiotics being 0.0% and 4.4%, respectively. A further point of note is that the great majority of the retrieved articles are written in English: for cyclization, MALDI-TOF, and antibiotics, the percentages are 88.9%, 78.5%, and 93.5%, respectively, with all of the remainder being in Chinese (with the exception of two Japanese articles). Non-English publications are known to attract fewer citations, other things being equal, than English publications:[33] it is hence likely that the Chinese citation counts discussed below are a slight underestimate of those that would have been obtained if all of the Chinese articles had been written in English.

The Chinese publications retrieved were then filtered before the analyses described later in the paper. First, some universities have multiple names in WOK, for example (using the WOK abbreviations): PEKING UNIV and BEIJING UNIV; SUN YAT SEN UNIV and ZHONGSHAN UNIV; UNIV HONG KONG and HKU; and TSING HUA UNIV, TSINGHUA UNIV, and QINGHUA UNIV. A more complex example is ACAD SINICA, the WOK abbreviation for Academia Sinica. Academia Sinica was originally founded in mainland China in 1928 but split into the Chinese Academy of Sciences (CHINESE ACAD SCI) in mainland China and Academia Sinica (ACAD SINICA) in Taiwan after the Chinese Civil War. However, a few scientists in mainland China have continued to use the old name: accordingly, publications from ACAD SINICA with an address in mainland China have had their institution changed to CHINESE ACAD SCI. Second, some publications were retrieved with addresses in Taiwan; these were eliminated, as were those that did not have a complete address field. Third, publications with addresses that include HONG KONG were removed if they had been published prior to 1997, when sovereignty was transferred from the United Kingdom back to the People's Republic of China. The resulting outputs contained 2065 cyclization publications, 850 MALDI-TOF publications, and 2034 antibiotics publications.

In addition to looking at the Chinese literature for our three chosen topics, we have also considered the extent to which Chinese chemical research is collaborative in nature and the extent to which such collaborations affect the impact of this research. The information required for this part of the study is available from WOK since the great majority of the publications there are associated with address details for the places where the work was carried out. It was hence a simple task to classify each item as belonging to one of the three categories identified by van Leeuwen:[34] no collaboration (if an institution published on its own), national collaboration (if an institution published in collaboration with another institution in the same country), or international collaboration (if an institution published in collaboration with an institution in another country). We shall refer to these three categories subsequently as NO, NAT, and INT, respectively.

Bibliometric Indicators. A very wide range of bibliometric indicators has been described in the literature,[2−5] with some of these suggested specifically for use in comparative studies.[35,36] Van Raan[37] has noted that it is unwise to use just a single indicator in a bibliometric analysis, and we have thus made use of multiple indicators in the work reported here: number of publications, *P*; number of citations, *C*; Hirsch index, *h*; mean number of citations per publication, *CPP*; percentage of publications not cited, *PNC*; and impact factor, *IF*. These indicators are easy to calculate (see below) and provide an overview of research performance from a range of different aspects.

The basic data retrieved from the WOK is the set of publications meeting the search criteria: this gives *P* directly, then *C* by using these *P* publications for a WOK citation search, and *CPP* is then computed from *P* and *C*. These three "commonly used" indicators[38] and *PNC* are among the set of bibliometric indicators that has been developed at the Centre for Science and Technology Studies at the University of Leiden (at http://www.cwts.nl) and that is now widely used in bibliometric studies. To these, we have added two others that have aroused considerable discussion in the literature: the Hirsch index and the impact factor.

The Hirsch index is the largest value of *h* for which an author (or an institution, or a country, or whatever) has published *h* outputs that have each attracted at least *h* citations.[38] The impact factor, *IF*, measures the extent to which a particular journal is cited in the literature, with values being listed in the *Journal Citation Report* (JCR) part of WOK; specifically, it is the average number of times papers from a journal published in the past two years (e.g., 2006 and 2007) have been cited in the JCR year (e.g., 2008). We have used this idea to compute the average number of times papers published by a country are cited in a particular year. Citation rates vary from subject to subject, and to normalize for this effect, we have chosen as a time-window the cited half-life of a subject category, i.e., the median age of articles cited by journals in that specific JCR category. We have chosen the category containing the largest number of outputs for each of our three subjects to compute the half-life, and hence the *IF* values for each of these subjects. For cyclization, the most important JCR category is "Chemistry, Organic", which yields a citation window of 5 years; the corresponding figures for MALDI-TOF and for antibiotics are 6 and 5 years, respectively.

RESULTS

Publications and Citations. The first Chinese papers on cyclization appeared in WOK in the 1980s, but it was not until the early 1990s that appreciable numbers began to appear, with the subsequent growth being extremely rapid. This is illustrated in Figure 1, which plots *P* for the four chosen countries (the single point at the bottom-left denotes early USA papers, the first of which dates from 1933). The growth over the past few years has been particularly rapid,

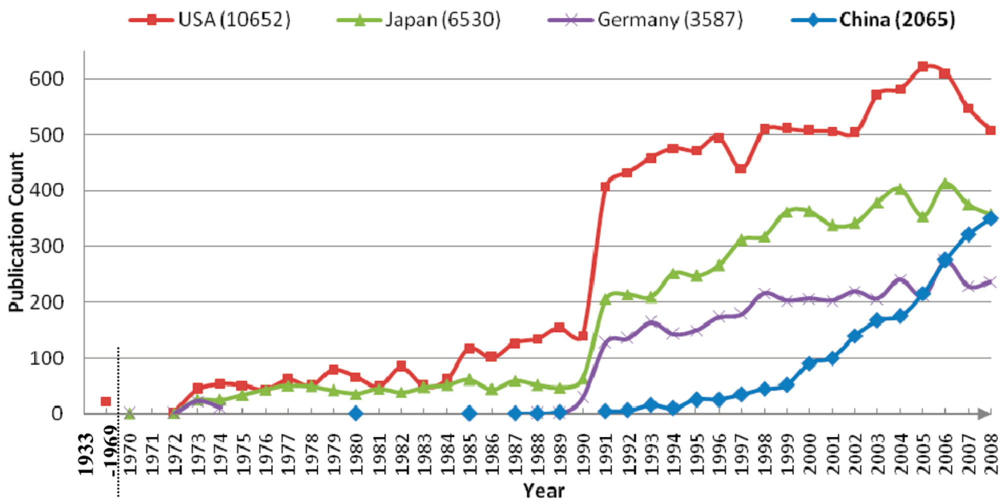

Figure 1. Publication growth of China and the comparison countries in cyclization.

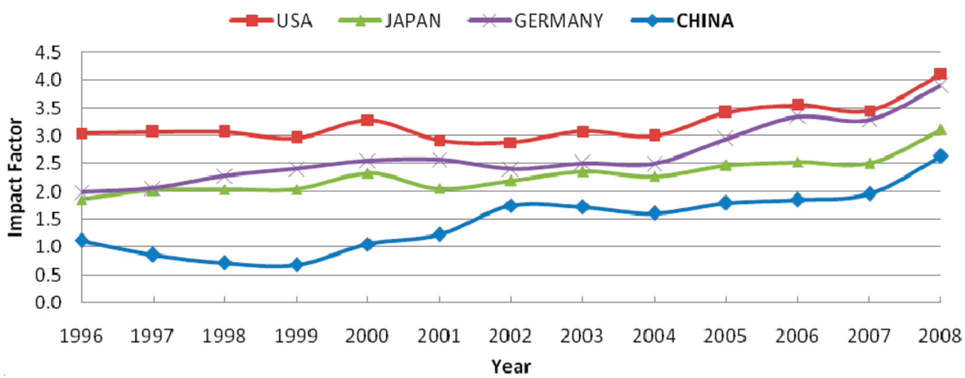

Figure 2. *IF* changes for China and the comparison countries in cyclization with a 5-year citation window.

Table 1. Bibliometric Indicators for Research into Cyclization, MALDI-TOF, and Antibiotics in China, USA, Germany, and Japan[a]

	cyclization				MALDI-TOF				antibiotics			
indicator	China	USA	Germany	Japan	China	USA	Germany	Japan	China	USA	Germany	Japan
P	2065	10 652	3587	6530	850	2186	1321	765	2034	45 274	9417	10 117
C	13 839	252 227	60 027	97 273	3650	36 540	19 100	8139	13 363	994 181	154 396	150 468
CPP	6.7	23.7	16.7	14.9	4.3	16.7	14.5	10.6	6.6	22	16.4	14.9
PNC	27.2	11.2	7.2	7.3	34.8	14.6	12.4	13.3	30.3	14.4	16.3	11.5
h	40	148	85	84	26	74	52	38	42	262	126	117

[a] The indicators are the number of publications, *P*; number of citations, *C*; mean number of citations per publication, *CPP*; percentage of publications not cited, *PNC*; and Hirsch index, *h*.

with China overtaking Germany around 2006 and likely to overtake Japan in 2009. A feature that is very obvious in Figure 1 (and also in the subsequent Figure 5) is the dramatic rise in *P* values in 1991: this arose from a large increase that year in the number of journals covered in WOK. It will also be seen that there is a drop for two of the countries in the 2008 values as compared to 2007: this may be due to incomplete coverage of all issues of 2008 journals when the searches were carried out, and similar comments apply to Figures 3 and 5.

The *IFs* for the four nations' publications are shown in Figure 2, where it will be seen that all of them have grown over the period (1996−2008). China's *IF* values have always been the lowest of the four, but China exhibits the largest percentage increase (135.4%) over the period 1996−2008. The figure only goes back to 1996 so as to exclude (given the 5-year window) data prior to 1991 when the WOK coverage changed (and similar comments apply to Figure 6 below). Although China's research productivity (as represented by the publication count) is growing rapidly, the impact of this research (as indicated by the citation-based indicators, all of which are listed in Table 1) still lags far behind the other three countries. China thus has the largest *PNC*, and smallest *C*, *CPP*, and *h* values in the top left-hand part of Table 1; for example, Germany generates 1.7 times as many outputs as does China, but these attract 4.3 times as many citations.

The *P* plot for the MALDI-TOF research is shown in Figure 3 (where both the total numbers of publications and the time duration are much less for this relatively new spectroscopic technique than they were for the cyclization data in Figure 1). China has the third largest total number of publications in this subject area; its annual productivity has already passed Japan and Germany and will overtake the USA in the next few years if current trends continue. This rapid catching-up can also be seen in the *IF* plot of

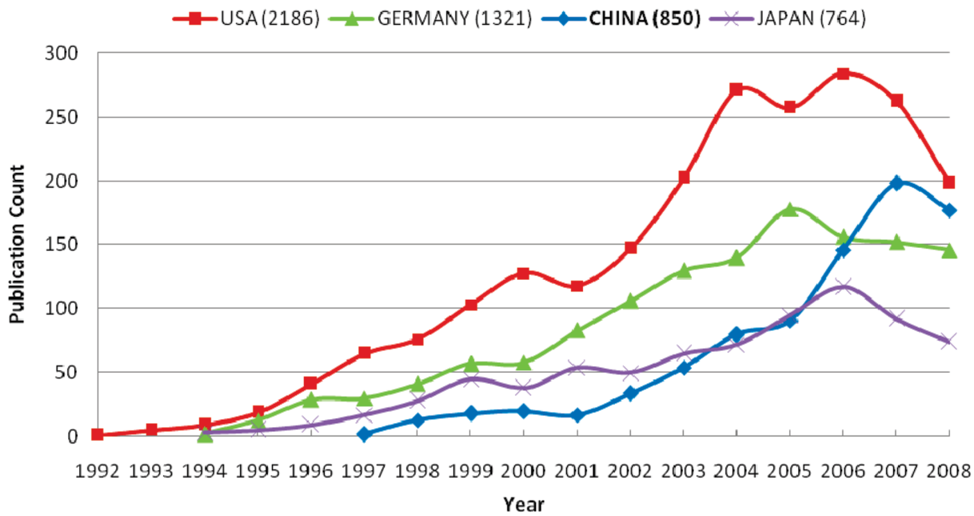

Figure 3. Publication growth of China and the comparison countries in MALDI-TOF.

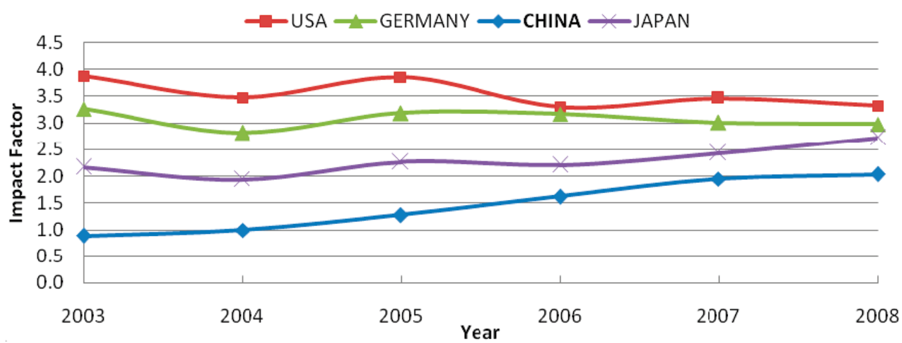

Figure 4. *IF* changes for China and the comparison countries in MALDI-TOF with a 6-year citation window.

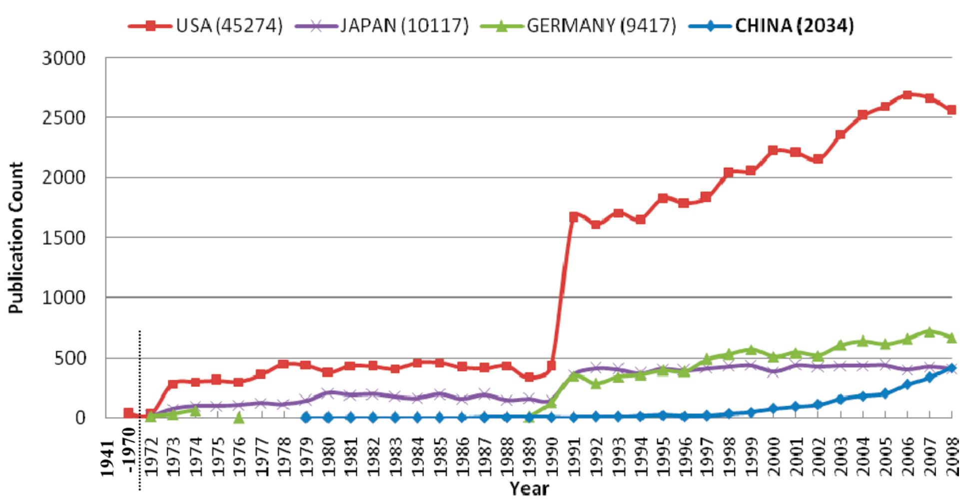

Figure 5. Publication growth of China and the comparison countries in antibiotics.

Figure 4: the percentage increase for China here is 129.7%, which is again the largest increase for the four nations. However, the central part of Table 1 shows that the impact of the Chinese research is still noticeably inferior to that of the other three nations; for example, over one-third of the Chinese outputs have never received a single citation (a percentage that is more than twice that for any of the other three nations).

Figure 5 demonstrates clearly the USA's world leadership in antibiotics research (especially since the changes in WOK coverage), producing over twice as many outputs as the other three nations combined over the entire period. China's overall *P* value is low (as noted previously, it is ranked only 16th among the 65 nations identified in WOK as having published on antiobiotics); however, the recent rapid growth seen in Figures 1 and 3 is also apparent here, with China having overtaken Japan in 2008 and now rapidly catching Germany. China's *IF* value is increasing only very slowly with respect to the other three nations (though it still has the largest percentage increase at 70.3%), and Table 1 again emphasizes the very limited impact of the Chinese research to date.

Table 1 presents data for the entire period under review, i.e., up to and including 2008. Sections a and b of Table 2 show this data subdivided to cover the periods up to 2005 and 2006−08, respectively. It will be seen that during the three years, 2006−08, China produced approximately one-

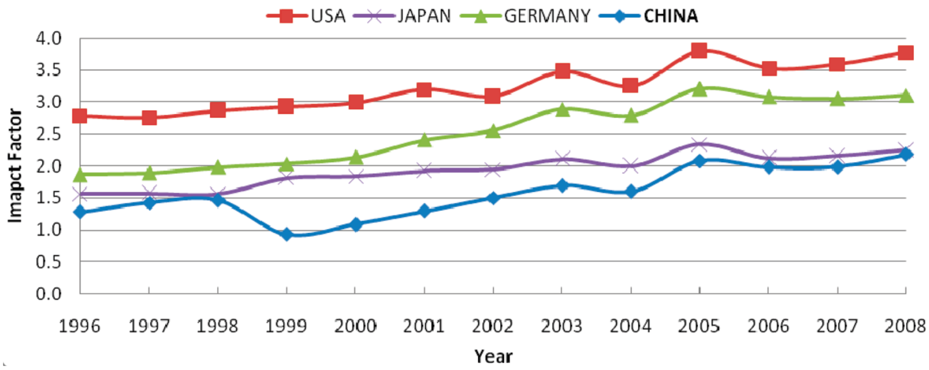

Figure 6. *IF* changes for China and the comparison countries with a 5-year citation window.

Table 2. Bibliometric Indicators for Research into Cyclization, MALDI-TOF, and Antibiotics in China, USA, Germany, and Japan for (a) pre-2006 and (b) 2006−08[a]

	cyclization				MALDI-TOF				antibiotics			
indicator	China	USA	Germany	Japan	China	USA	Germany	Japan	China	USA	Germany	Japan
					(a) Pre-2006							
P	1117	8915	2832	5342	329	1415	853	474	1001	37 142	7339	8839
C	10 878	238 341	54 866	90 091	2388	31 768	16 907	6693	11 239	944 350	143 876	145 749
CPP	9.7	26.7	19.4	16.9	7.3	22.5	19.8	14.1	11.2	25.4	19.6	16.5
PNC	14.1	9.3	3.6	4.1	18.8	9.1	6.2	6.1	11.7	7.8	11.9	7.9
h	39	146	84	83	26	74	52	38	42	258	124	116
					(b) 2006−2008							
P	948	1737	755	1188	521	771	468	291	1033	8132	2078	1278
C	2961	13 886	5161	7182	1262	4772	2193	1446	2124	49 831	10 520	4719
CPP	3.1	8.0	6.8	6.0	2.4	6.2	4.7	5.0	2.1	6.1	5.1	3.7
PNC	42.7	15.4	13.0	15.8	44.9	19.1	18.8	22.3	48.3	23.1	27.9	28.5
h	19	42	29	31	14	25	18	17	15	57	32	23

[a] The indicators are: number of publications, *P*; number of citations, *C*; mean number of citations per publication, *CPP*; percentage of publications not cited, *PNC*; and Hirsch index, *h*.

half of all of the outputs in each of the three areas, specifically, 45.9, 61.3, and 50.8% for cyclization, MALDI-TOF, and antibiotics, respectively. Given the short period of time, it is not surprising that China has smaller *CPP* and *h* and larger *PNC* values for 2006−08 as compared to pre-2006; however, the difference is much less than is the case for the other three countries. For example, the USA's *h* value for antibiotics for 2006−08 is 3.8 times the Chinese value, whereas this figure is 6.1 times for pre-2006. The relative impact of the Chinese research in the three areas is hence increasing rapidly.

Effect of Collaboration on Research Impact. We have described in Methods how each article was categorized as NO (no collaboration), NAT (national collaboration), or INT (international collaboration). A publication based on INT normally attracts more citations than ones based on NAT or NO;[34,39−42] however, this is not always the case[43,44] and the extent of the difference varies across disciplines.[40] Previous studies of the importance of collaboration in Chinese research have given rather variable results. Haiqi and Hong found that the relative proportions for the three types of collaboration (NO, NAT, and INT) for Chinese scientific publications in 1997 were 73.0, 25.4, and 1.6%, respectively,[45] while data provided by Wang et al. suggest values of 16.8, 81.6, and 1.6%, respectively, for 2001.[46] Both of these studies were very wide-ranging, involving over 101K outputs and over 202K outputs, respectively, in the two studies; however, they were restricted to journals published in China and thus did not include outputs in journals published elsewhere in the world. This would imply under-counting of international collaboration outputs, whereas an analysis of plant science, biophysics, and cell biology publications in WOK suggests a value for INT of 21.8%.[47] The INT values for the three areas studied here—9.1% for cyclization, 15.2% for MALDI-TOF, and 25.2% for antibiotics—are much closer to the latter figure, and we can hence conclude that international collaborations already play an important role in Chinese chemical research.

Table 3. Numbers of Publications Resulting from Chinese International Collaborations

nation	cyclization	MALDI-TOF	antibiotics
USA	70	51	186
Germany	35	18	74
Japan	13	10	34
total (3)	118	79	294
total (all)	203 (25)	141 (26)	742 (65)

This role is explored further in Table 3, which summarizes the research collaboration of China with other nations in the three chosen areas; a broader study of Chinese international collaborations has been reported recently by He.[48] Each column in the main body of Table 3 gives the number of joint outputs involving the USA, Germany, and Japan; involving any of these three nations; and involving any of the nations identified in the search outputs. For example, collaborations with the USA, German, and Japan accounted for 58.1% (118 out of 203) of the joint publications involving China and one or more of 25 other countries. These three countries accounted for 56.0% of the MALDI-TOF collaborations and 39.6% of the antibiotics collaborations.

Table 4. Chronological Distribution of the Three Types of Collaboration in Chinese Cyclisation Research

	type of collaboration								
	NO			NAT			INT		
year	*C*	*P*	*CPP*	*C*	*P*	*CPP*	*C*	*P*	*CPP*
1985	17	1	17.0						
1986									
1987	3	1	3.0						
1988	5	1	5.0						
1989	2	1	2.0	6	2	3.0			
1990									
1991	72	3	24.0				27	2	13.5
1992	31	5	6.2	15	1	15.0			
1993	208	13	16.0	4	1	4.0	13	2	6.5
1994	62	8	7.8	2	1	2.0	7	1	7.0
1995	224	17	13.2	55	8	6.9	8	1	8.0
1996	172	19	9.1	9	3	3.0	71	4	17.8
1997	138	26	5.3	102	9	11.3			
1998	505	30	16.8	70	12	5.8	11	3	3.7
1999	494	34	14.5	102	15	6.8	20	3	6.7
2000	793	60	13.27	179	24	7.4	78	7	11.1
2001	742	57	13.0	223	33	6.8	119	10	11.9
2002	822	71	11.6	512	55	9.3	177	14	12.6
2003	806	72	11.2	767	82	9.4	147	14	10.5
2004	540	72	7.5	619	83	7.5	175	20	8.80
2005	762	102	7.5	831	95	8.7	131	19	6.9
2006	770	131	5.9	621	112	5.5	164	33	5.0
2007	438	147	3.0	601	143	4.2	146	31	4.7
2008	83	156	0.5	121	171	0.7	17	24	0.7
total	7689	1027	7.5	4839	850	5.7	1311	188	7.0

Table 5. Chinese *CPP* Values for No Collaboration (NO), National Collaboration (NAT), and International Collaboration (INT) When There Are at Least 10 Such Outputs in a Year for Each Type of Publication

	cyclization			MALDI-TOF			antibiotics		
year	NO	NAT	INT	NO	NAT	INT	NO	NAT	INT
2000							7.8	10.0	18.5
2001	13.0	6.8	11.9				9.0	13.8	16.0
2002	11.6	9.3	12.6				9.6	12.5	16.3
2003	11.2	9.4	10.5				6.0	9.2	23.2
2004	7.5	7.5	8.80	5.8	8.8	10.8	7.3	8.1	14.7
2005	7.5	8.7	6.9	9.0	6.2	9.2	4.4	7.7	10.2
2006	5.9	5.5	5.0	5.1	4.3	7.9	3.8	4.2	5.6
2007	3.0	4.2	4.7	2.4	1.8	2.5	2.1	1.7	3.2
2008	0.5	0.7	0.7	0.6	0.4	0.3	0.3	0.3	0.7

Table 3 summarizes the INT outputs; Table 4 provides the full details not just for the INT outputs but also for the NAT and NO outputs. It is evident that there are far fewer INT outputs than NAT or NO outputs and that INT has grown much more slowly than the other two categories. As noted above, it is normally the case that INT outputs attract more citations than other types of output, and we have hence considered the data in Table 4 for 2001−08 to ascertain whether this is in fact the case here, using the *CPP* data; 2001 is taken as a starting point for the cyclization search, as this is the first year for which $P \geq 10$ for all three categories of output. This subset of the *CPP* data is shown in the left-hand part of Table 5, with the corresponding subsets for the other two subjects completing the table.

Inspection of Table 5 will show that international collaboration is rather less consistently beneficial in terms of attracting citations than might have been expected from the many previous studies in the literature.[34,39−42] INT research has achieved a greater impact in antibiotics than in the other two areas, and it is possible that this is related to the fact that international collaborations are better established here than in MALDI-TOF and cyclization research. Specifically, there are far more INT antibiotics outputs (both in terms of the actual numbers of publications and as a percentage of the total number of publications for all three types of collaboration) than for the other two research areas; moreover, the first INT collaborations were as early as 1983 (as against 1991 and 1998 for cyclisation and MALDI-TOF, respectively) so that at least some of the collaborations have had a chance to become widely recognized in the relevant research community.

Role of the Chinese Academy of Sciences. In addition to suggesting that INT outputs would be highly cited, Leeuwen also suggested that NAT outputs would be cited more than NO outputs.[34] We ascribe the surprisingly good performance of NO here (as reflected in the *CPP* values) to the way that science is organized in China, with much of the research being carried out in the many institutes that collectively comprise the Chinese Academy of Sciences. These various institutes—there are about 90 of them covering all branches of science—are all described in WOK as CHINESE ACAD SCI. This means that this multicomponent organization produces very large numbers of outputs that are regarded by WOK as coming from a single organization, even through many of them may, in fact, may be NAT outputs because they involve collaborations between different parts of the Academy. Moreover, many of the country's leading scientists work for the Academy, rather than for a conventional university or research institution, and it is hence hardly surprising that Chinese NO outputs can attract proportionally far more citations than might be the case in other countries. If we take cyclization, where a total of 297 different Chinese institutions figure in the search outputs, no less than 373 of the 1027 NO outputs (36.3% of them) come from the Academy, these attracting 4721 of the 7689 NO citations (61.2% of them). This corresponds to an institutional *CPP* value of 12.7 for NO outputs involving the Academy, versus a value of 6.6 for NAT outputs involving the Academy and 4.9 for NAT outputs not involving the Academy. Thus, NAT collaboration is good for non-Academy institutions in terms of research impact (which is the expected situation), but the converse applies for the Academy itself; as the Academy contributes so much research it is not really surprising that we found NO outputs to be significantly more cited than NAT outputs. The Academy makes a rather smaller, but still substantial, contribution to Chinese research in MALDI-TOF—70 out of 302 NO (23.2%) outputs, involving 277 different institutions—and this is sufficient to ensure that NAT is not superior to NO. The Academy makes a much smaller contribution to antibiotics research, only 85 out of 737 (11.5%) NO outputs, these 737 involving no less than 560 different institutions and this, coupled with the state of INT collaboration noted above, means that the expected superiority of INT to NO is observed for this data set.

The important role played by the Chinese Academy of Sciences in Chinese chemical research is also apparent if we focus on the publications that contribute to an institution's *h*-index, since these are, by definition,[38] the publications with most impact. Tables 6−8 list the ten most productive institutions in each of the three subject areas (12 institutions

Table 6. The 10 Most Productive of the 297 Chinese Institutions Listed in WOK as Publishing on Cyclization and the Composition of Their *h*-Indices

institution	*P*	*h*-index	type of collaboration NO	NAT	INT
Chinese Academy of Sciences	807	35	30	4	1
Zhejiang University	193	18	4	14	0
Lanzhou University	154	13	3	10	0
Peking University	108	18	7	6	5
Nankai University	107	12	3	9	0
Fudan University	67	12	0	12	0
University of Hong Kong	64	19	13	6	0
Nanjing University	56	10	1	5	4
Central China Normal University	46	8	6	2	0
East China University of Science and Technology	46	8	1	5	2

Table 7. The 12 Most Productive of the 277 Chinese Institutions Listed in WOK as Publishing on MALDI-TOF and the Composition of Their *h*-Indices

institution	*P*	*h*-index	type of collaboration NO	NAT	INT
Chinese Academy of Sciences	194	17	8	3	6
Fudan University	98	12	1	9	2
Peking University	63	11	7	2	2
Sun Yat Sen University	40	7	3	3	1
Zhejiang University	31	6	2	3	1
Tsing Hua University	30	7	0	2	5
Jilin University	29	6	2	2	2
Xiamen University	29	6	1	5	0
Beijing Institute of Radiation Medicine	20	6	1	5	0
Chinese University of Hong Kong	20	6	2	2	2
National Center of Biomedical Analysis	20	6	0	5	1
Shandong University	20	6	1	5	0

Table 8. The 10 Most Productive of the 560 Chinese Institutions Listed in WOK as Publishing on Antibiotics and the Composition of Their *h*-Indices

institution	*P*	*h*-index	type of collaboration NO	NAT	INT
Chinese Academy of Sciences	312	19	6	5	8
University of Hong Kong	129	21	2	14	5
Chinese University of Hong Kong	124	19	6	9	4
Zhejiang University	105	10	2	6	2
Chinese Academy of Medical Sciences	79	15	2	3	10
Peking University	75	11	4	3	4
Shanghai Jiao Tong University	62	7	2	0	5
Fudan University	58	10	1	2	7
China Agricultural University	51	8	1	2	5
Hong Kong University of Science and Technology	42	14	4	5	5

for MALDI-TOF where four of them have the same *P* value). The Chinese Academy of Sciences is the most productive Chinese institution in all three subject areas (it is also the most productive worldwide in cyclization and MALDI-TOF, but only 55th worldwide in antibiotics). Tables 6 and 7 further demonstrate the strong internal focus of the Chinese Academy of Sciences, since 30/35 (cyclization) and 8/17 (MALDI-TOF) of its *h*-index publications do not involve collaboration (the 6/19 fraction for antibiotics is less notable). Conversely, in Table 5, the remaining nine institutions have 37 of their 69 NAT *h*-index publications involving the Academy; the corresponding figures for Tables 7 and 8 are 9 of 43 and 3 of 44.

CONCLUSIONS

China is playing an increasingly large role in scientific research, and this paper has charted the development of its research in three important areas of chemistry. Analysis of Chinese publications demonstrates the increasing productivity of Chinese researchers; however, the impact of this work (as denoted by citation-based indicators (such as *C*, *CPP*, *PNC*, and the *h*-index) is still noticeably less than that of major international competitors such as Germany, Japan, and the USA, although the gap is decreasing rapidly.

Studies of the extent to which Chinese institutions collaborate with each other and with institutions in other countries show that the level of collaboration is growing. However, international collaborations have, to date, not had as much impact as might have been expected from previous studies of academic collaboration; and national collaborations tend to have less impact, on average, than single-institution research owing to the crucial role played in research by the Chinese Academy of Sciences.

ACKNOWLEDGMENT

We thank the China Scholarship Council for funding this research and Dr. Beining Chen for suggesting the three subject areas studied here.

REFERENCES AND NOTES

(1) Pritchard, A. Statistical Bibliography or Bibliometrics. *J. Doc.* **1969**, *25*, 348–349.
(2) Hood, W. W.; Wilson, C. S. The Literature of Bibliometrics, Scientometrics, and Informetrics. *Scientometrics* **2001**, *52*, 291–314.
(3) Borgman, C. L.; Furner, J. Scholarly Communication and Bibliometrics. *Ann. Rev. Inf. Sci. Technol.* **2002**, *36*, 3–72.
(4) Thelwall, M.; Vaughan, L.; Björneborn, L. Webometrics. *Ann. Rev. Inf. Sci. Technol.* **2005**, *39*, 81–135.
(5) Bar-Ilan, J. Informetrics at the Start of the 21st Century—A Review. *J. Informetrics* **2008**, *2* (1), 1–52.
(6) Hawkins, D. T. Electochemistry Journals. *J. Chem. Inf. Comput. Sci.* **1977**, *17*, 41–45.
(7) Windsor, D. A. Bibliometric Problems Associated with the Patent Literature. *J. Chem. Inf. Comput. Sci.* **1978**, *18*, 55–56.
(8) Hawkins, D. T. The Literature of Noble Gas Compounds. *J. Chem. Inf. Comput. Sci.* **1978**, *18*, 190–199.
(9) Windsor, D. A. Using Bibliometric Analyses of Patent Literature for Predicting the Clinical Fates of Developing Drugs. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 218–221.
(10) Hawkins, D. T. Crystallographic Literature: A Bibliometric and Citation Analysis. *Acta Crystallogr.* **1980**, *A36*, 475–482.
(11) Tsay, M. Y. Bibliometric Study of the Application of Computers in Synthetic Organic, Physical, Inorganic, and Anlytical Chemistry Literature Abstracted by Chemical Abstracts in 1986. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 156–158.
(12) Karki, M. M. S.; Garg, K. C. Bibliometrics of Alkaloid Chemistry Research in India. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 157–161.
(13) Pohl, L.; Quoniam, L.; Quazzotti, S.; Dou, H. Microlevel Informetric Investigation in Phytochemistry. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 821–827.
(14) Kostoff, R. N.; Braun, T.; Schubert, A.; Toothman, D. R.; Humenik, J. A. Fullerene Data Mining Using Bibliometrics and Database Tomography. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 19–39.
(15) Neuhaus, C.; Daniel, H.-D. A New Reference Standard for Citation Analysis in Chemistry and Related Fields Based on the Sections of *Chemical Abstracts*. *Scientometrics* **2009**, *78*, 219–229.
(16) Bishop, N.; Gillet, V. J.; Holliday, J. D.; Willett, P. Chemoinformatics Research at the University of Sheffield: A History and Citation Analysis. *J. Inf. Sci.* **2003**, *29*, 249–267.
(17) Kim, M.-J.; Kim, B.-J. A Bibliometric Analysis of Publications by the Chemistry Department, Seoul National University, Korea, 1992—1998. *J. Inf. Sci.* **2000**, *26*, 111–119.
(18) Bornmann, L.; Daniel, H.-D. Selecting Manuscripts for a High Impact Jouurnal through Peer Review: A Citation Analysis of Communications That Were Accepted by *Angewandte Chemie International Edition*, or Rejected but Published Elsewhere. *J. Am. Soc. Inf. Sci. Technol.* **2008**, *59*, 1841–1852.
(19) Onodera, N. A Bibliometric Study on Chemical Information and Computer Sciences Focusing on Literature of JCICS. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 878–888.

(20) Warr, W. A., Twenty Five Years of Progress in Cheminformatics. In *229th American Chemical Society National Meeting*; San Diego, USA, 2005 (at http://www.warr.com/25years.html, accessed 13th November, 2009).

(21) Willett, P. A Bibliometric Analysis of the *Journal of Molecular Graphics and Modelling*. *J. Mol. Graph. Modell.* **2007**, *26*, 602–606.

(22) Behrens, H.; Luksch, P. A Bibliometric Study in Crystallography. *Acta Crystallogr.* **2006**, *B62*, 993–1001.

(23) Redman, J.; Willett, P.; Allen, F. H.; Taylor, R. A Citation Analysis of the Cambridge Crystallographic Data Centre. *J. Appl. Crystallogr.* **2001**, *34*, 375–380.

(24) Willett, P. A Bibliometric Analysis of Chemoinformatics. *Aslib Proc.* **2008**, *60*, 4–17.

(25) Agarwal, P.; Searls, D. B. Can Literature Analysis Identify Innovation Drivers in Drug Discovery? *Nature Rev. Drug Discov.* **2009**, *8*, 865–878.

(26) Leydesdorff, L.; Zhou, P. Are the Contributions of China and Korea Upsetting the World System of Science. *Scientometrics* **2005**, *63*, 617–630.

(27) Zhou, P.; Leydesdorf, L. The Emergence of China as a Leading Nation in Science. *Res. Policy* **2006**, *35*, 83–104.

(28) Zhou, P.; Thijs, B.; Glanzel, W. Is China Also Becoming a Giant in Social Sciences. *Scientometrics* **2009**, *79*, 593–621.

(29) Cyranoski, D. China Increases Share of Global Scientific Publications. *Nature* **2004**, *431*, 116.

(30) Glänzel, W.; Schubert, A.; Braun, T. A Relational Charting Approach to the World of Basic Research in Twelve Science Fields at the End of the Second Millennium. *Scientometrics* **2002**, *55*, 335–348.

(31) National Science Board. *Science and Engineering Indicators 2008*; National Science Foundation: Arlington, VA, 2008.

(32) Kumari, G. L. Synthetic Organic Chemistry Research: Analysis by Scientometric Indicators. *Scientometrics* **2009**, *80*, 559–570.

(33) Bornmann, L.; Daniel, H. What Do Citation Counts Measure? A Review of Studies on Citing Behavior. *J. Doc.* **2006**, *64*, 45–80.

(34) van Leeuwen, T. N. Strength and Weakness of National Science Systems: A Bibliometric Analysis through Cooperation Patterns. *Scientometrics* **2009**, *79*, 389–408.

(35) Schubert, A.; Glanzel, W.; Braun, T. Scientometric Datafiles. A Comprehensive Set of Indicators on 2649 Journals and 96 Countries in All Major Science Fields and Subfields 1981−1985. *Scientometrics* **1989**, *16*, 3–478.

(36) Moed, H. F.; De Bruin, R. E.; Van Leeuwen, T. N. New Bibliometric Tools for the Assessment of National Research Performance: Database Description, Overview of Indicators and First Applications. *Scientometrics* **1995**, *33*, 381–422.

(37) Van Raan, A. F. J. Comparison of the Hirsch Index with Standard Bibliometric Indicators and with Peer Judgement for 147 Chemistry Research Groups. *Scientometrics* **2006**, *67*, 391–502.

(38) Hirsch, J. E. An Index to Quantify an Individual's Scientific Research Output. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16569–16572.

(39) Glänzel, W.; Schubert, A. Double Effort−Double Impact? A Critical Review at International Co-Authorship in Chemistry. *Scientometrics* **2001**, *50*, 199–214.

(40) Glänzel, W.; De Lange, C. A Distributional Approach to Multinationality Measures of International Scientific Collaboration. *Scientometrics* **2002**, *54*, 75–89.

(41) Roberts, G. *International Partnerships of Research Excellence. UK−USA Academic Collaboration*; University of Oxford: Oxford, 2006 (at http://www.wolfson.ox.ac.uk/UK-US-Academic-Collaboration/GarethRobertsIPoREx.pdf, accessed 13th November, 2009).

(42) Adams, J.; Gurney, K.; Marshall, S. *Patterns of International Collaboration for the UK and Leading Partners*; UK Office of Science and Innovation: Leeds, 2007.

(43) Glänzel, W. National Characteristics in International and Scientific Co-Authorship Relations. *Scientometrics* **2001**, *51*, 69–115.

(44) Schmoch, U.; Schubert, T. Are International Co-Publications an Indicator for Quality of Scientific Research. *Scientometrics* **2008**, *74*, 361–377.

(45) Haiqi, Z.; Hong, G. Scientific Research Collaboration in China. *Scientometrics* **1997**, *38*, 309–319.

(46) Wang, Y.; Wu, Y.; Pan, Y.; Ma, Z.; Rousseau, R. Scientific Collaboration in China as Reflected in Co-Authorship. *Scientometrics* **2005**, *62*, 183–198.

(47) Jonkers, K. Emerging Ties: Factors Underlying China's Co-Publication Patterns with Western European and North American Research Systems in Three Molecular Life Science Subfields. *Scientometrics* **2009**, *80*, 775–795.

(48) He, T. International Scientific Collaboration of China with the G7 Countries. *Scientometrics* **2009**, *80*, 571–582.

CI9003199