The Patent Landscape of Protein Crystals

Felix White, Patent Analyst, IP Australia

Time Series Analysis

Time series analysis can be used to visualise trends in inventive activity in a field. In the field of protein 3D structures, the first PCT applications were published in 1994 but it was not until the turn of the century that the volume of applications reached significant levels (more than 10 applications per year), reflecting the increasing availability of high-resolution crystal structures at that time.

The number of PCT applications peaked in 2003 and has gradually declined, with a minor peak in 2009 (Figure 1). The surge of patent applications is likely to have receded once easy-to-crystallise and/or commercially relevant structures were completed. It is also possible that the perception that protein crystal structures are now "obvious" to make has slowed recent patenting activity in the field.

Figure 1: PCT applications for protein 3D structures since 2000, by publication year.
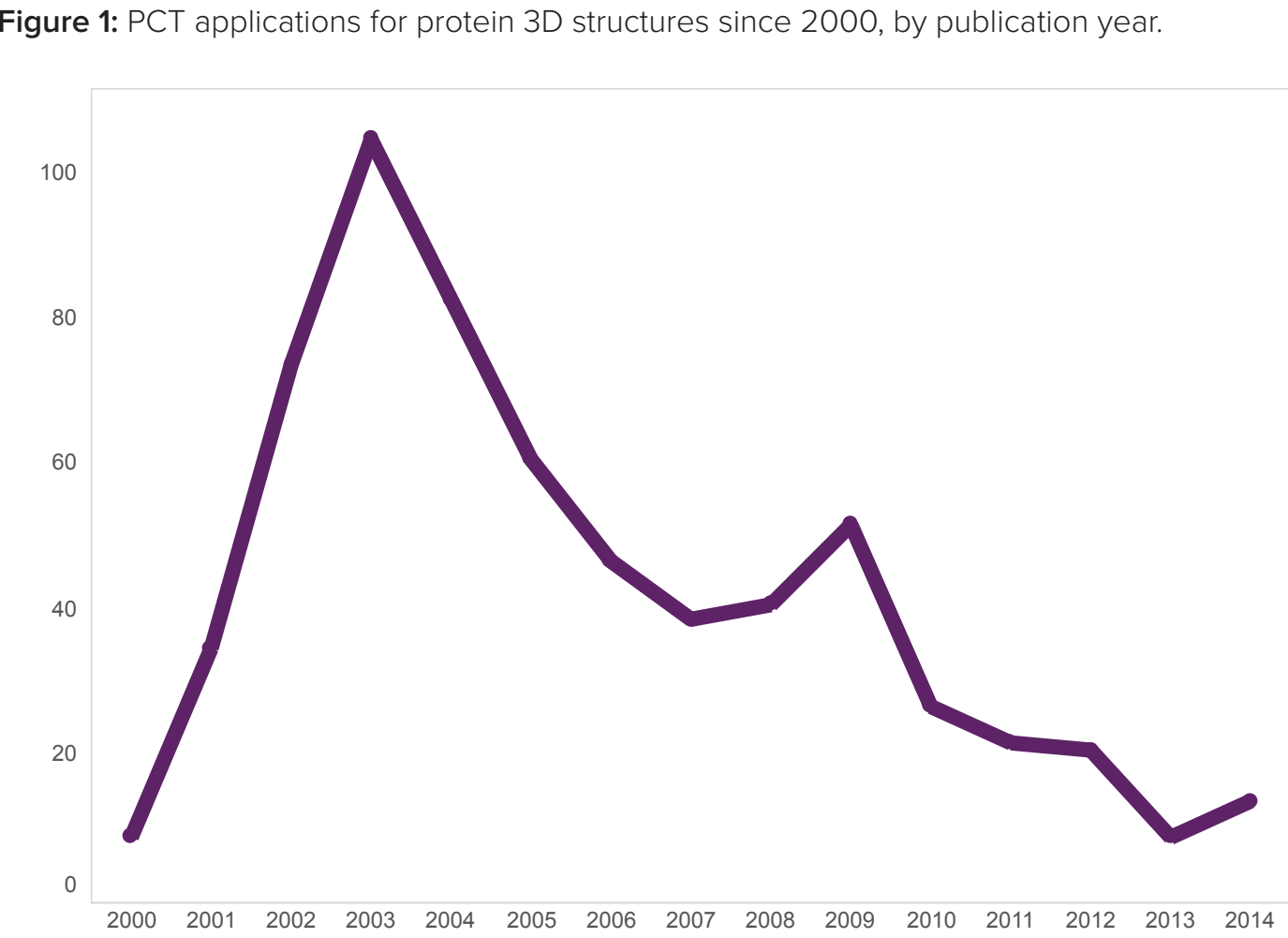
The classification breakdown of the applications in the early vs late 2000s suggests that the majority of the applications in the early peak were related to structural genes, particular bacterial and viral proteins, while the applications in the late 2000s are predominantly related to antibody structures (data not shown).
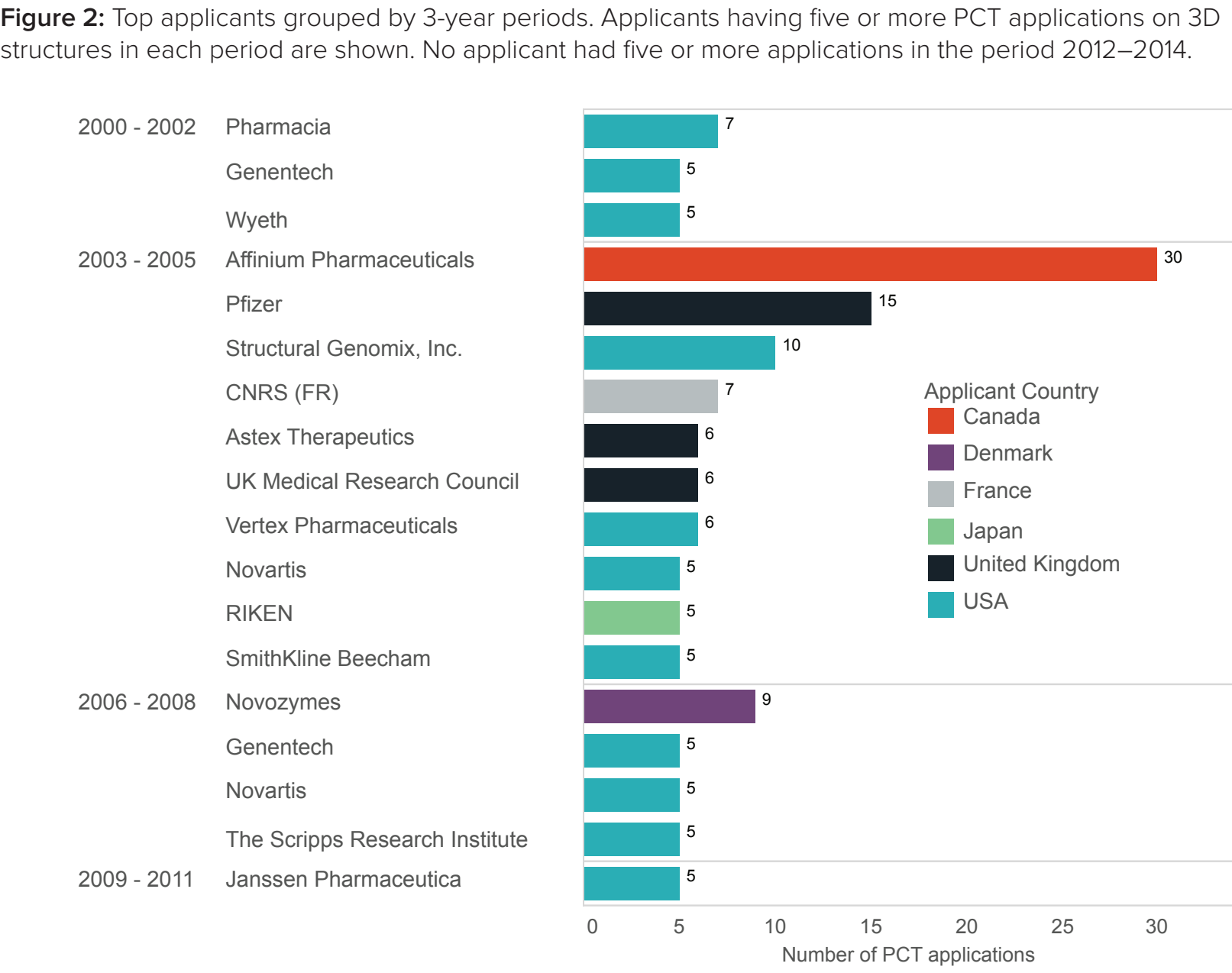
There is a wealth of data available in international patent databases that is often underutilised by the research community. The emerging field of patent analytics provides the ability to extract, analyse, visualise and interpret patent data.

Protein crystallography continues to be essential for understanding proteins in health and disease, and enabling rational design of drug targets. This poster provides a short demonstration of some of our patent analytical techniques, using PCT applications from the OECD REGPAT database related to 3D protein structures.

Figure 2: Top applicants grouped by 3-year periods. Applicants having five or more PCT applications on 3D structures in each period are shown. No applicant had five or more applications in the period 2012–2014.

The major companies filing in each period of the 2000s are shown in Figure 2. The single largest patent filer in the surge period of 2003–2005 was Affinium Pharmaceuticals.
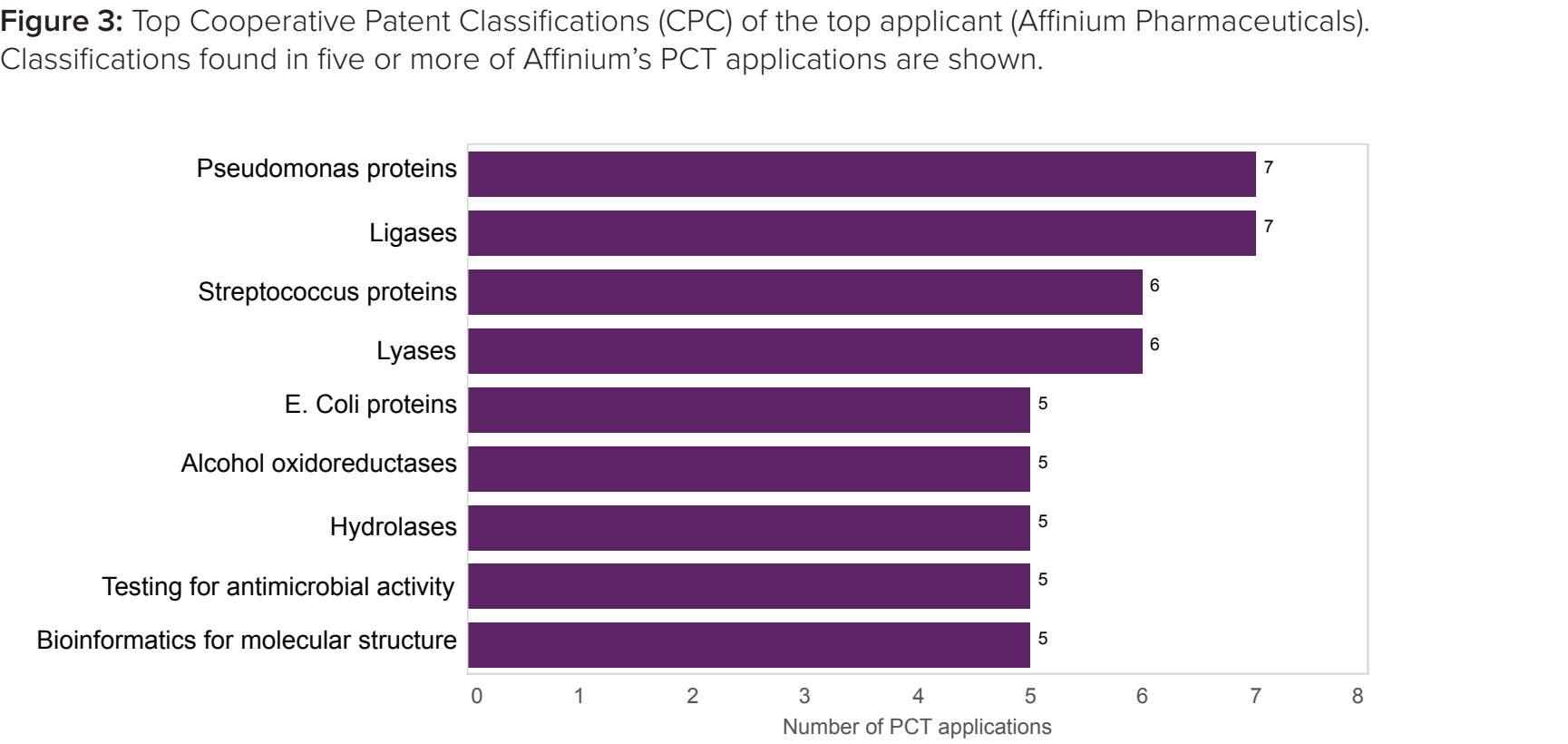
Inventive activity in the field of protein 3D structures measured by PCT applications peaked in 2003. The most active company in terms of patenting activity was Affinium Pharmaceuticals, which developed narrow-spectrum antimicrobials by in silico drug design based on bacterial metabolic enzyme 3D structures.

The most active patenting regions for 3D crystal structures are in the north east and west coast of the United States. Melbourne is the sixth most active region worldwide.

The major Australian patent applicants in the crystallography field are publicly funded research organisations and they have collaborated in a quarter of their PCT applications in the field.

Figure 3: Top Cooperative Patent Classifications (CPC) of the top applicant (Affinium Pharmaceuticals). Classifications found in five or more of Affinium's PCT applications are shown.
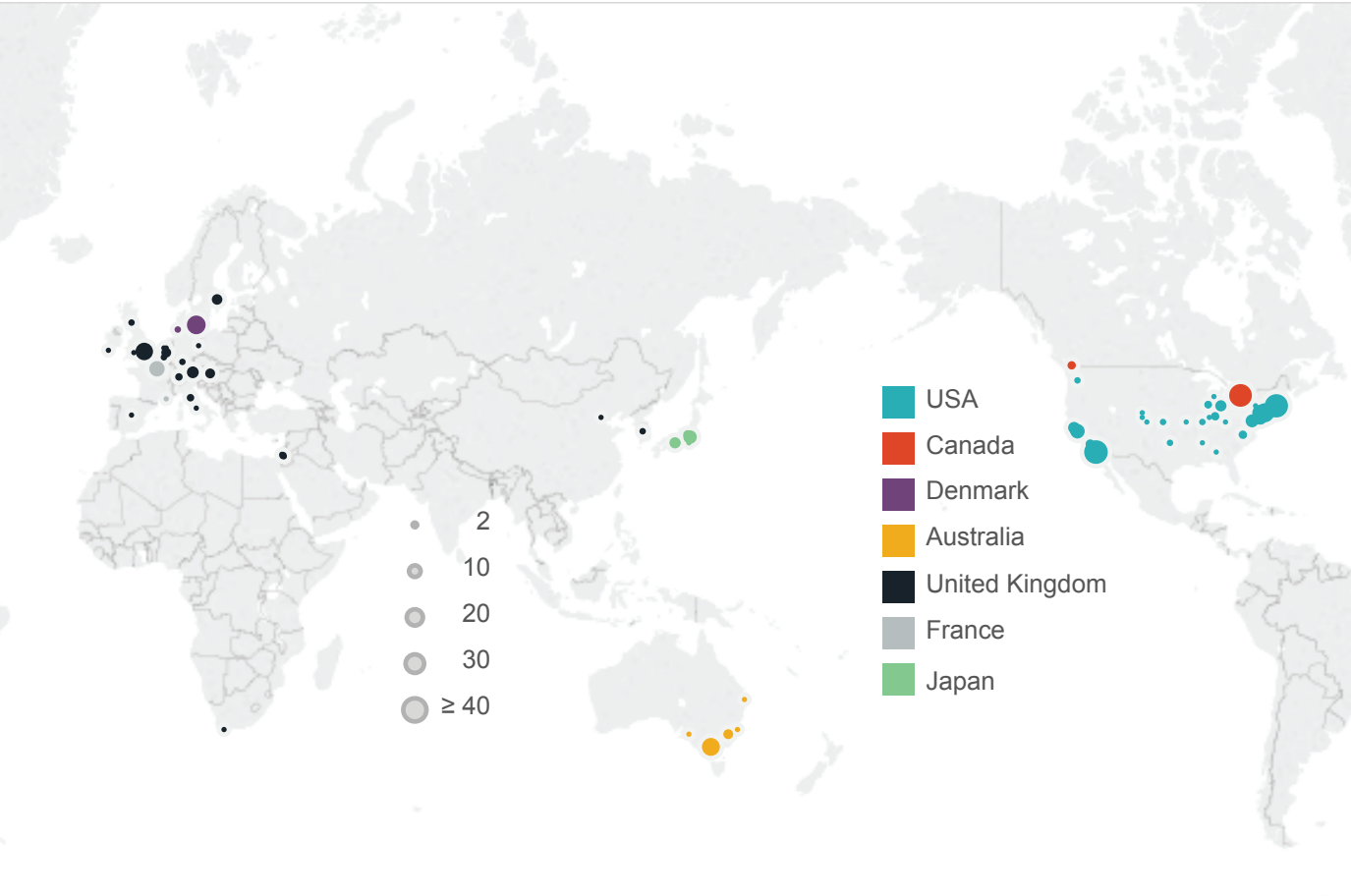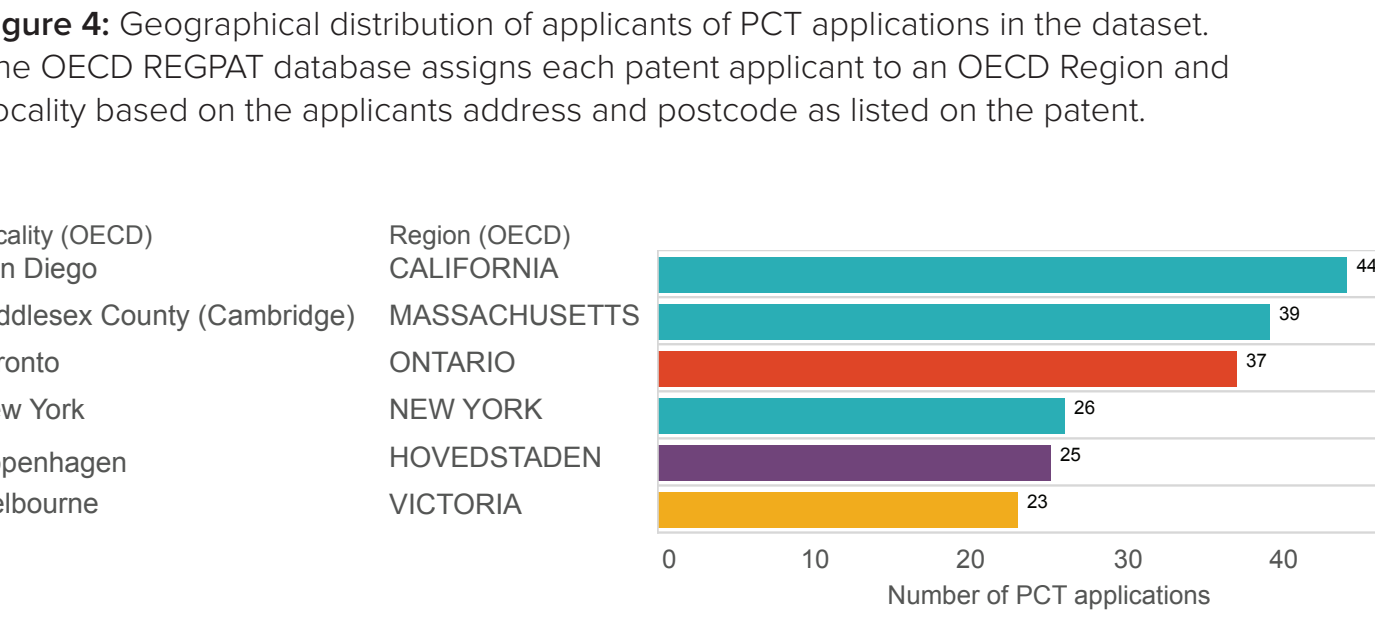

In order to visualise the technology being pursued by Affinium in that time, the CPC classes assigned to those applications are visualised in Figure 3. These indicate the majority of these applications are directed to structures of bacterial proteins, in particular pseudomonas and streptococcus enzymes and also testing for antimicrobial activity.

Geographical Analysis

Geographical analysis can be used to visualise hotspots of inventive activity. The OECD REGPAT database assigns regional locations to all patent applicants of PCT applications, allowing easy visualisation with Tableau software (Figure 3b) The greatest concentration of applications can be seen in the north-east of north America as well as California, and spread throughout western Europe.

The highest ranked cities can be seen in Figure 4. The patent applications arising from San Diego and Cambridge, MA, are shared between a number of different organisations whereas Toronto is dominated by a single company, Affinium Pharmaceuticals (see Figure 2). Copenhagen is the leading European centre for inventive activity in this field, and this is made up of Novozymes and Novo Nordisk (data not shown).

Figure 4: Geographical distribution of applicants of PCT applications in the dataset. The OECD REGPAT database assigns each patent applicant to an OECD Region and Locality based on the applicants address and postcode as listed on the patent.

Australia is the seventh country in terms of overall applications (data not shown), and Melbourne is the sixth OECD region in the field (Figure 4). This is most likely due to the presence of the Australian Synchrotron in Melbourne, although the direct involvement of the synchrotron cannot be determined from this patent data set since it does not generally retain rights to IP generated and therefore is not listed as a patent applicant.

Patenting and Collaboration in Australia

The most active filer in Australia in 3D protein structures is the Walter and Eliza Hall Institute for Medical Research. WEHI's PCT applications in the data set include WO 2007/147213 and WO 2014/047673 relating to the structure of the insulin receptor, WO 2003/014159, WO 2003/025017 and WO 2004/031232 relating to the structure of cytokine receptors, enabling design of anti-cancer agents, and WO2007128080 and WO2007147213 relating to the structure of apoptotic peptides (MCL-1 and BCL-W).

Figure 5: Australian Applicants with more than one PCT application related to 3D crystal structures.
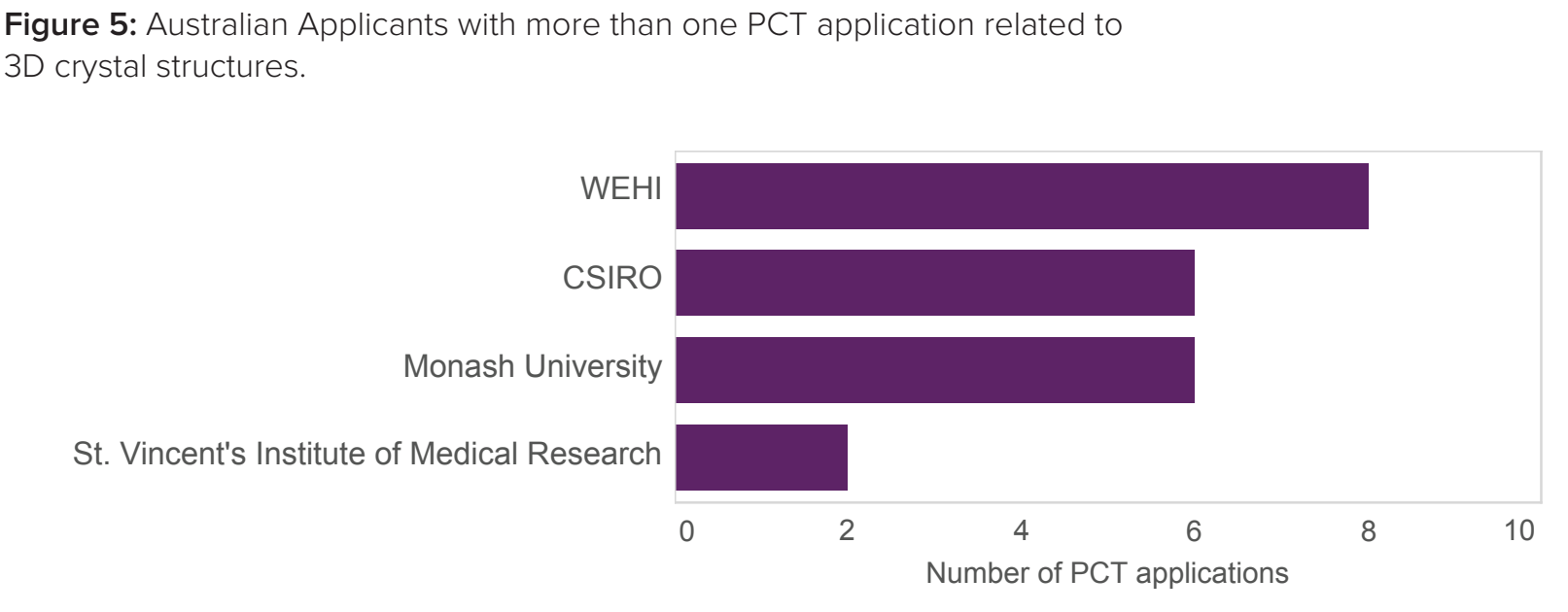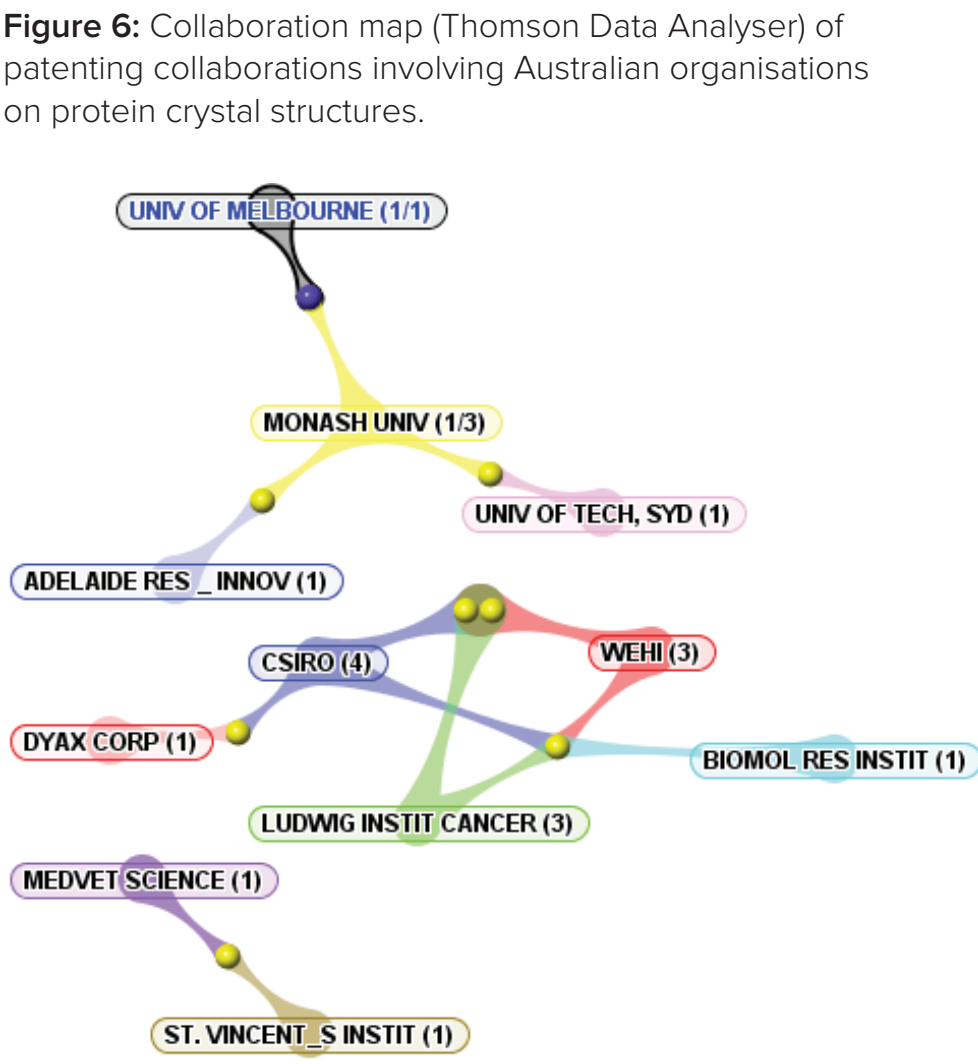

The structural biology division of WEHI is continuing to work with the Australian Synchrotron to obtain high-resolution structural information on proteins and peptides related to apoptosis and necroptosis. This will assist in the development of targeted treatments for conditions including autoimmune disease and cancer.

Source www.wehi.edu.au/about-structure/scientific-divisions/structural-biology

Because patent applications can be filed in the name of more than one entity, analysis of co-applicants can be used to provide evidence of collaborative activity in the dataset.

Consistent with Melbourne's status as one of the hot spots for crystallography patenting activity, there is a high degree of collaborative activity between Australian applicants. Figure 6 shows all collaborations between Australian institutions (which make up 27 per cent of the total Australian applications in this dataset). The three WEHI applications to cytokine receptors mentioned above were made in collaboration with CSIRO and the Ludwig Institute for Cancer Research. CSIRO has also collaborated with the US-based Dyax Corporation on antibodies to IGF-II. Another notable collaboration is between St Vincent's Institute in Melbourne and Medvet Science in Adelaide. This related to a 2008 patent application (WO 2008/052277) on the crystallisation of GM-CSF complexed with its receptor and methods of in vitro and in silico screening using this crystal structure.

Figure 6: Collaboration map (Thomson Data Analyser) of patenting collaborations involving Australian organisations on protein crystal structures.

Classification Analysis

The proportion of different types of technical activity within the field of protein crystal structures can be visualised by classification analysis. CPC classification captures all aspects of technology and so counting the number of applications in each classification allows visualisation of the most commonly patented technologies, for example the type of the protein under study and the species from which the protein is obtained.

Figure 7 shows the top CPC marks found in the data set. The most common mark relates to computational techniques for determination of molecular structure, indicating that a considerable proportion of inventive activity in the field is related to structural assignment algorithms. Computational proteomics is also found among the most common CPCs. Medicinal applications are also seen in the form of protein based therapies and antibody combination therapies. The most commonly patented molecular structures are of antibodies and kinases.

The level of detail in the CPC shows that different applications focus on different properties of antibodies, for example humanised antibodies, high affinity antibodies and structures of linear epitopes. Other well represented CPCs relate to analysis of ligand-target interactions and in silico screening methods. Methods of single crystal growth are the 15th most common classification, featuring in 30 patent applications.

Figure 7: Most frequently assigned Cooperative Patent Classification (CPC) categories. Categories assigned to more than 30 PCT applications in the dataset are shown. The classification C07K 2299/00 (protein three-dimensional structures) was found on all applications in the data set and is not shown.

About Patent Analytics

Patent Analytics is the analysis of patent information to discover insights, relationships and trends to inform business strategy. It comprises searching for and analysing a broad range of information associated with intellectual property. Patent analytics has the ability to visualise historical and geographical trends of inventive activity, assess the impact of public research, decrease risks associated with project selection, and identify collaboration opportunities for research and licensing.

IP Australia, which comprises Australia's patent and trademark offices, has recently founded a Patent Analytics Hub to help Australian innovators make the most of their intellectual property. By engaging with publicly funded research organisations such as universities and medical research institutes, we want to use our knowledge and skills to help you understand technology areas, find collaborators and boost the commercial returns from your research.
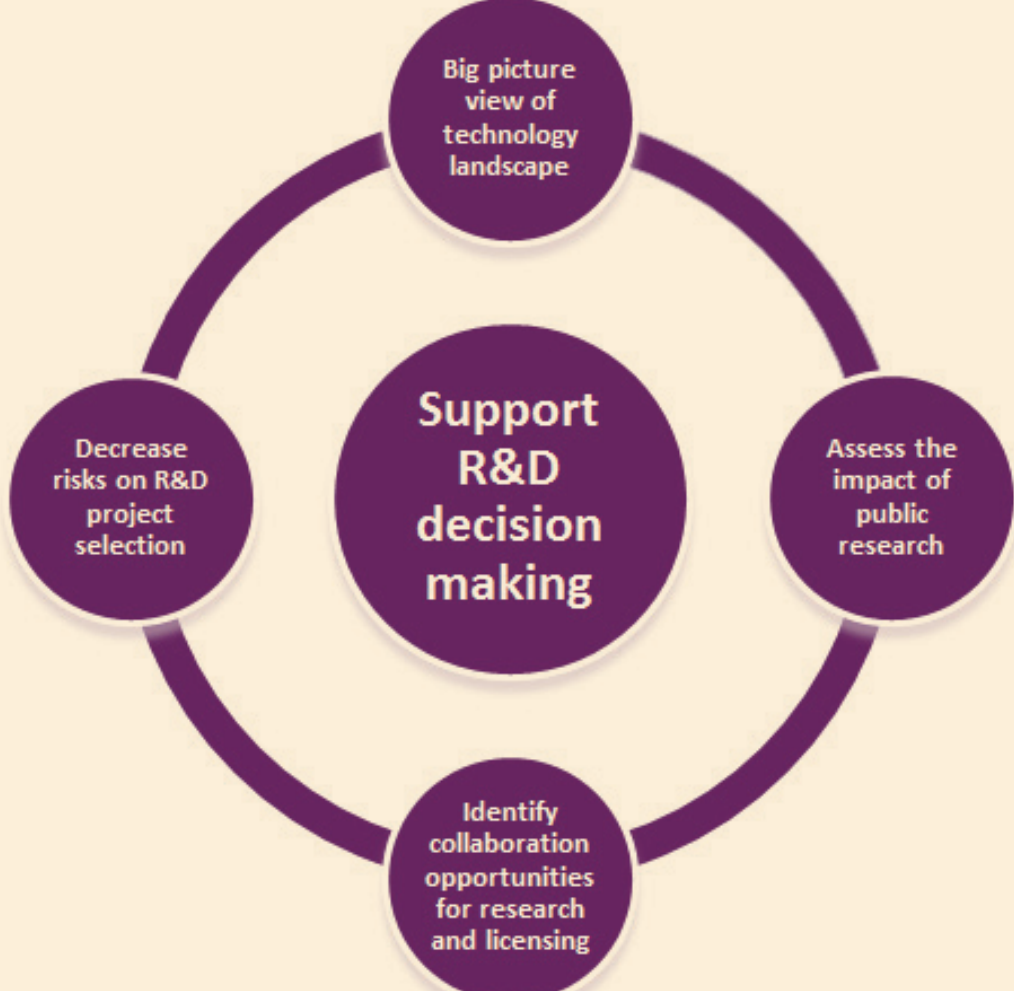
The Patent Analytics Hub aims to support Australian research institution to increase collaboration and strategic use of IP and to improve commercialisation outcomes. We can do that by using the following analyses:

- Technology landscape analyses to identify and map patent documents in an area of interest, to determine which organisations are innovating in that area and to ascertain when they operated.
- Technology assessment and benchmarking to inform patent management and licensing strategies.
- Identification of competitors, licensees, prospective collaborators or partners for complimentary or joint IP development, by identifying companies (or individuals) innovating in a technology area both in Australia and overseas.
- Indicators of patent impact, such as geographic filing breadth, grant success and citation count

If you would like to know more about patent analytics email analytics@ipaustralia.gov.au

See some more of our work at www.ipaustralia.gov.au

Figure 8: Advantages of patent analytics. Patent information can provide:

Big picture view of technology landscape

Decrease risks on R&D project selection

Support R&D decision making

Assess the impact of public research

Identify collaboration opportunities for research and licensing

Glossary of terms

The Patent Cooperation Treaty (PCT) is a mechanism that allows a single patent application to be used as a basis for filing later applications around the world. Each PCT application receives an international (WO) publication and a preliminary examination report. The PCT entered into force in 1978 and WO publications have been issued since 1980.

PCT applications are a useful metric for inventive activity because in general there will be one PCT application for one invention, regardless of how many later applications may be filed around the world to cover that invention.

PCT applications are also considered to represent valuable inventions in the sense that their applicants intend to obtain patent protection overseas.

To assist patent searching and examination, patents are classified into technical categories. This study uses the Cooperative Patent Classification (CPC) which has been developed jointly by the US and European patent offices and combines both offices' systems into a single classification. Each published patent application is assigned one or more CPC marks representing which technical fields to which it relates.

Methods and databases

The data set for analysis was obtained by searching the OECD REGPAT database (September 2015 edition) for all PCT applications that contained the CPC mark C07K 2299/00, which is specific for inventions relating to protein three-dimensional structures. We identified 660 PCT applications having this CPC mark which were used as a data set for further analysis. REGPAT was also used to obtain the applicant names and addresses, publication dates and all CPC marks for the patents in the data set. CPC marks (alphanumeric codes) were replaced with descriptive text in all figures.