I. Introduction

A biological sequence is a single and continuous molecule, either nucleic acid or protein, presented in the form of its structural combination. Sequences are the essential substance of many biological inventions. The functionality of such inventions is predominantly determined by the order of 4 nucleobases- designated as C, G, A, and T (U), in the case of DNA (RNA), or 20 amino acid residues- designated as single or triple alphabet(s) codes in the event of proteins, such as R or Arg for arginine. Changes to these sequences may lead to disparate results: from complete loss-of-function to approximately maintaining identical functions. The phenomena of codon degeneracy (for nucleic acids) and neutral mutation (for proteins) constitute the main basis of shared functions amongst similar biological sequences, or homologous sequences.

With the state-of-the-art biotechnology, mutagenesis[1] proves easier each day at an unprecedented pace. Modification of biological sequences and making variants become relatively simple tasks. The same functionality of a certain biological sequence is assumed to be easily achieved by creating a variant imitating the reference sequence. Moreover, the possible number of variants can easily reach an astronomical figure according to combinatorics. As a consequence, patent protection over only the specific sequences disclosed in a patent could not reward the contribution of its inventor, and thus cannot achieve the *quid pro quo* of the patent system. Inventors, therefore, wish to draft their patent claims in a broader way so as to encompass a wide range of similar sequences, usually by means of a minimal percentage homology to a specific sequence. On the other hand, this practice may bring anti-innovation effects, as some technical progress owing nothing to the teachings of these patents could also fall within the claimed scope of protection, merely because of sequence homology. Therefore, a delicate line should be drawn as to what extent the inventors are allowed to claim homologous sequences.

---

1 See *Mutagenesis* at Wikipedia <https://en.wikipedia.org/wiki/Mutagenesis> accessed 10 September 2017. “Mutagenesis is a process by which the genetic information of an organism is changed, resulting in a mutation”.

Homology claims are generally allowed by patent offices across many jurisdictions, exemplified in examination guidelines by UKIPO.[2] However, the allowable threshold may vary from case to case, and may also change from time to time.[3] It reflects the plight of patent offices in balancing the interest of patent proprietors and the public. A recent case in China, *Novozymes*,[4] has brought homology claims in hot water again. Briefly, a patent relating to one kind of thermostable glucoamylase was invalidated by the Beijing High Court for lack of support,[5] due to the homology language used in the claims. It was argued that the homology claims encompass a large number of variant sequences whose functionality cannot be predicted, and a person skilled in the art cannot reasonably know which particular variant would work the invention. The Supreme Court upheld the patent on the ground that a further "species of origin" limitation in its auxiliary claims narrows down the scope of protection to only a few sequences attributed to a particular species, and a skilled addressee would reasonably predict that the sequences within the same species perform functions similar to each other.

Although the Supreme Court's decision stabilised the patent-in-suit on a seemingly valid ground, it did not touch upon the essence of the issue, and could possibly leave in problems for the future. This thesis aims at discussing the role of homology in biotech patents particularly in relation to

---

2 Intellectual Property Office of the UK, *Examination Guidelines for Patent Applications relating to Biotechnological Inventions in the Intellectual Property Office* (6 May 2010, last updated: 21 October 2016) (UK Biotech Guidelines) 49, Example 5: "A protein / polypeptide having the sequence SEQ ID No. 1 or a variant, homologue, or portion / fragment thereof".

3 Ibid. "There is no general rule for determination of the required agreement, which depends on context, most significantly the stringency conditions. As an example, a low homology sequence may 'pick out' a newly sequenced DNA/RNA, whereas to separate sequences encoding isoenzymes (which have closely related structures), homology of over 95% may be required. Thus the scope of the claim needs to be considered in the context of the specification as a whole".

4 *The Patent Reexamination Board (PRB) & Novozymes A/S v. Jiangsu Boli Bioproducts (Boli) Co., Ltd*, the Supreme People's Court (2016) Zui Gao Fa Xing Zai No.85; *The Patent Reexamination Board & Novozymes A/S v. Shandong Longda Bioproducts Co., Ltd (Longda)*, the Supreme People's Court (2016) Zui Gao Fa Xing Zai No.86.

5 *Novozymes v PRB*, The Beijing High People's Court (2014) Gao Xing (Zhi) Zhong Zi No. 3522; *PRB v Boli, Longda* The Beijing High People's Court (2014) Gao Xing (Zhi) Zhong Zi No. 3523/3524.

polypeptides and proteins, analysing the drawbacks of the *Novozymes* decision, and providing suggestions for future patent practice.

In this thesis, Section II summarises the ins and outs in relation to *Novozymes*. In the same part, reasons for the decision not being satisfactory are presented. Section III discusses the meaning of homology language in relation to its technical background. Next, Section IV explains that *Novozymes* leaves an unjustifiable and unclaimable gap in the technological space under patent law. Finally, Section V presents a more appropriate way to apply the test for support requirement.