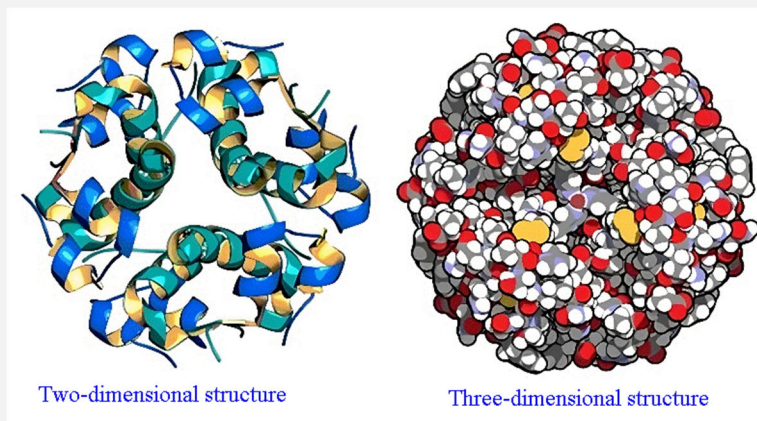


Featured Article

Biological Sequence Claims in CNIPA: Can Claims Use Open-Ended Definitions?



Biotechnology and biomedicine are among the most strategically significant frontier technology fields of the 21st century. The 2026 Report on the Work of the Government of China for the first time explicitly incorporated biomedicine into the layout of emerging pillar industries. Among the 76 innovative drugs approved domestically in 2025, biologics accounted for 23 (1), highlighting the importance of this industry. In the biomedical field, proteinaceous drugs such as monoclonal antibodies, bispecific/multispecific antibodies, antibody-drug conjugates, chimeric antigen receptors, and biological sequence-type applications such as gapmers and (modified) small interfering RNAs (siRNAs) are among the core forms of protection.

However, biological sequences, especially amino acid sequences, may have their spatial structures and/or biological activities significantly affected by the primary structures thereof (i.e., the sequences *per se*). Accordingly, it has always been a key concern to the Examiners before CNIPA when examining such claims whether claims relating to a biological sequence are supported by the description. When

examining the support issue, open-ended definitions are almost one of the primary reasons the Examiners cite to conclude that claims are not supported by the description. This article will briefly discuss the open-ended definitions of claims relating to a biological sequence, the common forms of the Examiners' opinions that such claims are not supported by the description, and the author's thoughts and suggestions on such issues, with a view to providing some reference for the industry.

I. Principal categories of biological molecules and key information thereof

The principal categories of biological molecules that the author encounters in practical work include the following:

1. Monoclonal antibodies, bispecific/multispecific antibodies, and antigen-binding fragments (Fab/scFv).

Main group: C07K 16/00: Immunoglobulins [IG], e.g., monoclonal or polyclonal antibodies.

Subgroup: C07K 16/18: Against material from animals or humans, targeting a specific target such as PD-1, or HER2.

Subgroup: C07K 16/28: Against receptors, cell surface antigens or cell surface determinants, targeting the core positions of tumor immunology and autoimmune disease antibodies.

Subgroup: C07K 16/46: Hybrid immunoglobulins, which is the exclusive classification position for bispecific antibodies, covering bispecific, trispecific and multispecific constructs.

2. Fusion proteins & chimeric antigen receptors (CARs).

Main group: C07K 19/00: Hybrid peptides, which is the core classification number for CARs, covering fusion structures of scFv-hinge region-transmembrane domain-intracellular co-stimulatory domain.

3. Active proteins, active peptides and antigenic peptides.

Main group: C07K 5/00: Peptides having up to four amino acids linked entirely or partially by peptide bonds (applicable to small active peptides).

Main group: C07K 7/00: Peptides having 5 to 20 amino acids, which is the core position for antigenic peptides, peptide drugs such as GLP-1 analogs.

Main group: C07K 14/00 and a large number of subgroups thereof: Peptides having more than 20 amino acids and derivatives thereof, covering naturally occurring or recombinant therapeutic proteins/enzymes (if the active moiety of a fusion protein is a specific receptor ligand or cytokine, auxiliary classification into lower subgroups is needed, e.g., C07K 14/705 (Receptors; Cell surface antigens; Cell surface determinants)).

4. Nucleotide coding sequences (DNA/RNA), antisense oligonucleotides.

Main group: C12N 15/00: Mutation or genetic engineering; DNA or RNA concerning genetic engineering, vectors, e.g., plasmids.

Subgroup: C12N 15/11: DNA or RNA fragments; Modified forms thereof (applicable to general coding sequences).

Subgroup: C12N 15/113: Non-coding nucleic acids modulating the expression of genes, e.g., antisense oligonucleotides, siRNAs, miRNAs, nucleic acid aptamers, which is the mandatory

classification number for antisense drugs.

Subgroup: **C12N 15/62**: DNA sequences coding for fusion proteins, which is strictly corresponding to the subgroup C07K 19/00, serving as the auxiliary classification at the nucleic acid level.

Among the above biological molecules, monoclonal antibodies, bispecific antibodies, Fab/scFv and other antigen-binding fragments falling within C07K 16/00; CAR structures, cytokine fusion proteins, dual-target fusion molecules and the like falling within C07K 19/00; and non-coding nucleic acid modulatory molecules such as gapmers, (modified) siRNAs, (modified) miRNAs and nucleic acid aptamers falling within C12N 15/113, are all among the biological-molecule categories with the substantial commercial value in the current biomedical industry.

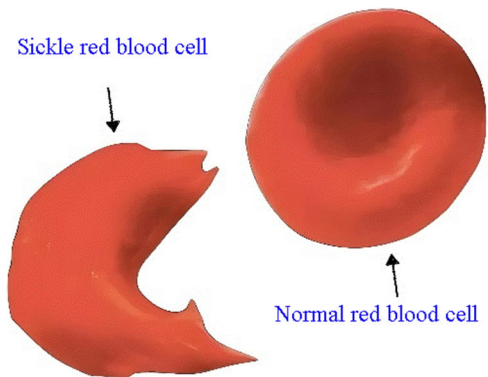
II. Unique issues encountered during CNIPA substantive examination of claims relating to biological sequences and corresponding countermeasures

1. Particularities of biological sequences.

As can be seen from the above classifications, biological sequences can be divided into proteinaceous biological sequences and nucleic acid-based biological sequences. Among them, for polypeptides and proteins requiring correct folding to exert their biological activities, and for antisense oligonucleotides requiring correct binding to the surface-exposed portions of DNA or RNA to exert their biological activities, the amino acid sequences or nucleotide sequences themselves may indeed have an important impact on their spatial structures or steric hindrance.

Moreover, the structure-activity relationship of biological molecules has not yet been elucidated to date. Therefore, it is currently impossible to reasonably infer the spatial structures and/or the biological activities or steric hindrance from the sequences *per se* as a primary structure. A classic example frequently cited by Examiners before CNIPA to show that the field has not yet established a sound structure-activity relationship such that the biological activity of a certain sequence cannot be predicted is the single point mutation (GAG→GTG) in the hemoglobin subunit beta (HBB) coding gene. This mutation results in the substitution of glutamic acid with valine at position 6 of the hemoglobin subunit beta (Glu→Val), leading to a dramatic change in the structure of the hemoglobin formed (abnormal hemoglobin S), which can no longer carry oxygen, thereby causing sickle cell anemia.

Therefore, in the examination of claims directed to biological sequences, CNIPA Examiners generally apply a relatively strict standard to the support requirement. In particular, where a claim uses an open-ended sequence definition, or defines a sequence with identity percentage definitions or substitutions / deletions / additions of one or more amino acid residues or nucleotide residues, the Examiners may reject the claim if mutants covered were not experimentally verified. This article will focus on the issue of claims using open-ended sequence definitions.



2. Common expressions of open-ended and closed-ended definitions.

2.1 Common expressions of open-ended definitions.

Comprise/contain/include/have (and their grammatical variations)

2.2 Common expressions of closed-ended definitions.

Consist of/is/of (and their grammatical variations)

3. Practice regarding claims defining biological sequences using open-ended language.

3.1 Common examination opinions on claims relating to a biological sequence defined with open-ended language.

Legal basis: Article 26, paragraph 4 of the Patent Law: The claims shall be based on the description and shall define the scope of the patent protection sought for in a clear and concise manner.

Core conclusion: Claims seeking protection of a biological sequence *per se* will encounter support objections during substantive examination when open-ended language is used. In practice, CNIPA Examiners will generally require that the applicant shall amend the open-ended definition to a closed-ended one, sometimes applying the requirement quite strictly.

For example, claim 1 of an application has already been defined in a closed-ended manner: "A XX polypeptide, which is a XX polypeptide variant occurring a D73K or Y85A mutation on the basis of SEQ ID NO: Y," and original claim 2 reads: "The XX polypeptide according to claim 1, wherein the XX polypeptide **has** an N-terminal truncation of residues 1 to 20 or 1 to 45." The Examiner still insisted that the applicant shall amend it to avoid using open-ended wording: "The XX polypeptide according to claim 1, wherein the XX polypeptide **occurs** an N-terminal truncation of residues 1 to 20 or 1 to 45."

The examination opinions raised by the Examiners that claims relating to a biological sequence defined with open-ended language are not supported by the description are generally as follows.

For example, the subject matter of a certain claim is: "1. An isolated enzyme having glucoamylase activity, **comprising** the full-length sequence of SEQ ID NO: 7."

The corresponding examination opinion of the Examiner is: "Addition of some propeptides or signal peptides before SEQ ID NO: 7, or addition of some histidines for purification convenience at its C-end, generally does not affect its function; however, the definitions of claim 1 do not only include these cases. The present description only provides that the SEQ ID NO: 33 or 7 from a specific source has the claimed enzyme activity, but **does not provide other examples and rules for addition that the resultant variants still retain the claimed function after what kind and how many nucleotide residues or amino acid residues were added at one or both ends thereof**. Consequentially, the generalization of claim 1 contains speculative contents, effects of which are difficult to predetermine and evaluate. Such a generalization goes beyond the scope disclosed in the description." (From (2014) Gao Xing (Zhi) Zhong Zi No. 3522 and Decision on Request for Invalidation No. WX17956)

That is, the Examiners will generally point out that since the claim adopts an open-ended language, it means that **any number** and **any kind** of amino acid residues or nucleotide residues can be added to **one or both ends** of the listed sequence, thereby leading to countless mutants. Since the biological activity of a biological sequence is substantively affected by the three-dimensional structures or spatial hindrance thereof and cannot be predicted, it is unforeseeable whether these mutants will have the claimed biological activity without specific experimental data.

3.2 Current examination practice for major types of biological sequence claims defined with open-ended language and corresponding countermeasures.

3.2.1 Types of biological sequence claims in which the use of open-ended definitions may currently be permitted.

If a claim seeks protection of a biological sequence *per se*, the use of open-ended language to define the biological sequence is not permitted. However, if the biological sequence is part of a larger **common biological molecule**, then open-ended definitions may be permitted when defining that larger common biological molecule.

For example, if a claim seeks protection of a recombinant expression vector such as a recombinant Adeno-Associated Virus (rAAV) and the inventive finding thereof lies in the biological sequence comprised therein (the payload), then the following expression would generally not be considered not being supported by the description due to the use of an open-ended language: "A recombinant adeno-associated virus, which **comprises** the nucleotide sequence set forth in SEQ ID NO: X."

As another example, if a claim seeks protection of an antibody, then defining the antibody as comprising all six complementarity-determining regions (CDRs) with determined sequences is generally accepted: "A monoclonal antibody against X, which **comprises** heavy chain CDRs and light chain CDRs of the amino acid sequences set forth in SEQ ID NOs: 1-6, respectively."

One possible reason why open-ended language may be permitted in the above cases may lie in that it is well known in the art that a recombinant expression vector or an antibody is a known molecule with relatively fixed components consisting of framework regions with very low variability and a payload/six CDRs, which obviously cannot consist solely of the payload/six CDRs, and the other portions comprised therein are known in the art and vary very little.

3.2.2 Current examination practice for major types of biological sequence claims defined with open-ended language and corresponding countermeasures.

3.2.2.1 Immunoglobulins

At present, the broadest scope of protection for a claim relating to an immunoglobulin acceptable by the CNIPA is an immunoglobulin against a certain antigen defined by all six CDR sequences, i.e., the Examiners require the applicant to use closed-ended language to define the sequences of the six CDRs.

Possible countermeasures: Since there are multiple numbering systems for defining CDRs (e.g., Kabat, Chothia, IMGT) and the CDR sequences defined by these numbering systems differ considerably, we recommend the applicants defining the six CDR sequences using these numbering systems respectively, in order to protect the antibody as broadly as possible.

3.2.2.2 Active peptides/proteins having a specific biological activity

Where a claim seeks protection of an active peptide/protein *per se* with a specific activity, and the shortest core sequence determining the activity or the core mutation mode leading to the activity has been identified, it is recommended to define the peptide/protein with the activity thereof (a functional definition), the core sequence comprised therein, and a definition that the length thereof is the length of the core sequence plus a few additional amino acid residues at one or both ends or additionally comprised functional amino acid residues or sequence fragments.

In preferred solutions, rules for adding these additional amino acid residues or fragments shall be provided, such as a signal peptide sequence, a tag for affinity purification, a fluorescent protein sequence for localization, linking amino acid residues for fusion expression, and/or amino acid residues added to avoid terminal amino acid hydrolysis or oxidation or amidation or deamidation or cyclodehydration (for support, please see (2014) Gao Xing (Zhi) Zhong Zi No. 3522 and Decision on Request for Invalidation No. WX17956).

If a peptide/protein *per se* is a protein of a known activity with a limited length (tens of amino acid residues, e.g., 20-30 amino acid residues in length) and is characterized by a certain core

mutation, then defining the peptide/protein as comprising said mutation may be accepted by the Examiners as being supported by the description. For example, in Decision on Request for Re-examination No. FS77448, the expression relating to the biological sequence is "a DNA sequence encoding an insulin precursor having the following amino acid sequence: B(1-30)-X-X-Z-Y-Y-A(1-21), wherein B(1-30) is an analog of B-chain of human insulin, A(1-21) is an analog of A-chain of human insulin, X-X and Y-Y are both Lys-Arg, and Z is AspGlyLeuGly, wherein the analogs of A-and/or B-chain of human insulin comprise three mutations including the mutation at position A14, Glu." The panel in charge pointed out, in substance: "when the description has demonstrated that the mutation A14E is the key site affecting expression yield and the solutions involving the three-site mutations A14E+B10E+A8H have better technical effects, and since said A-chain and B-chains of human insulin are only 21 and 30 amino acids in length, and it also belongs to a routine technical means in the art to select the other two mutations in the A-chain and B-chains of human insulin with limited amino acid length, a person skilled in the art can reasonably select the other two mutations in addition to A14E according to the description or in combination with known knowledge on the structure-activity relationship of human insulin analogs through routine experiments, and such a selection does not require undue experimentation."

This case illustrates that the acceptability of an open-ended definition may depend on the factors such as the length of the sequence, the extent of the claimed variation, the predictability of the relevant technical field, the disclosed structure-activity relationship, and the amount of experimental support provided in the specification.

3.2.2.3 Antisense oligonucleotides (ASOs) such as gapmers and (modified) siRNAs.

As described above, ASOs such as gapmers and (modified) siRNAs require the antisense strand to bind to the surface-exposed portion of the corresponding target nucleic acid molecule *in vivo* to exert gene silencing activity thereof. Hence, the Examiners will certainly require the applicant to define the sequence of the ASO to a closed-ended sequence.

Possible countermeasures: Applicant may consider to define the ASO as an ASO for inhibiting the expression of a certain gene (a functional definition), define the ASO as comprising the shortest core sequence, and/or define the length of the ASO to a length common for similar ASOs.

For example, in Decision on Request for Invalidation No. WX58530, the claim reads: "A double-stranded RNAi agent **for inhibiting hepatitis B virus (HBV) expression in a cell**, wherein **the double-stranded RNAi agent comprises a sense strand and an antisense strand forming a double-stranded region**, wherein **the sense strand comprises** 5'-GUGUGCACUUCGCUUCACA-3' (SEQ ID NO:39) **and the length of the sense strand does not exceed 21 nucleotides**, and **the antisense strand comprises** 5'-UGUGAAGCGAAGUGCACACUU-3' (SEQ ID NO:40) **and the length of the antisense strand does not exceed 23 nucleotides**, wherein substantially all nucleotides of the sense strand and substantially all nucleotides of the antisense strand are modified nucleotides, wherein the sense strand is conjugated with a ligand attached to the 3'-terminus, and wherein the ligand is one or more GalNAc derivatives attached using a divalent or trivalent branched linker." The invalidation petitioner argued that the claim is not supported by the description due to the open-ended definitions used therein, but the panel in charge held that "the complementarity between the core sequence of the dsRNA and the target RNA sequence is key to achieving mRNA degradation and secondary siRNA generation. A person skilled in the art, when understanding the term 'comprising' in claims 1 and 2, also does so on the premise that the complementarity between the core sequence of the dsRNA and the target RNA sequence is not impaired. Additionally, the specification of this patent specifically discloses the meaning of 'a chain comprising a sequence', regions that can tolerate mismatches to a relative extent, structures forming overhangs at the sense or antisense strand, definitions and conditions of sequence complementarity, etc. ... A person skilled in the art, based on the systematic records of the specification of this patent, in combination with routine technical means for selection and verification, is capable of adding one to two nucleotides at both ends of the core sequence of the

dsRNA to form overhang structures and the like, while maintaining the corresponding essential functional activity."

III. Suggestions for drafting

Although CNIPA currently adopts a very strict examination standard for open-ended definitions of claims relating to a biological sequence, the author still recommends using open-ended language to define the composition of biological sequences during drafting, particularly where broader protection may be desirable in jurisdictions that apply different standards.

In addition, the author suggests that during the drafting phase, if the applicant can provide more examples of mutants comprising the core sequence and relevant activity data, they should be included in the examples as much as possible, so as to provide a more sufficient basis for the claims to be supported.

References

(1) <https://baijiahao.baidu.com/s?id=1861151240170735479&wfr=spider&for=pc>

The "Featured article" is not equal to legal opinions.

If you need special legal opinions, please consult our professional consultants and lawyers.

Email address : ltbj@lungtin.com Website www.lungtin.com

For more information, please contact the author of this article.



FU,
Wenchuan

Senior Patent
Attorney

Mr. Wenchuan Fu has expertise in patent invalidation, patent administrative litigation, freedom to operate investigation, patent prosecution, patent strategy design and portfolio development. Mr. Fu focuses primarily on patent cases in technical fields of antibody-based drugs, vaccine development, protein- or peptide-based drugs, diagnostic reagent, RNAi, gene editing, biological therapy, stem cell therapy, tumor immunity, biomedicine, biochip, development of microbial strains, enzyme engineering, food technology, biological cosmetology, biochemical engineering, biological equipment, plant breeding, etc. Since January 2008, Mr. Fu has represented many famous life science companies to handle thousands of patent cases.

Mr. Fu has rich experience in patent agency practice in China, and cooperates with local attorneys to handle patent cases in U.S., Europe, India, Brazil, Argentina, Indonesia, Philippines, Vietnam, etc. Mr. Fu focuses on patent matters, primarily on patent application drafting, translation, proofreading, rewriting, advice for notification of OA and OA responses.



隆天知識產權
LUNG TIN IP ATTORNEYS

18th Floor, Tower B, Grand Place, No 5, Huizhong Road,
Chaoyang District, Beijing 100101, P. R. China

Tel: 0086-10-84891188 Email: LTBJ@lungtin.com
Fax: 0086-10-84891189 Web: www.lungtin.com