

## Beschluß des Präsidenten des Europäischen Patentamts vom 11. Dezember 1992 über die Darstellung von Nucleotid- und Aminosäuresequenzen in Patentanmeldungen und die Einreichung von Sequenzprotokollen\*

Der Präsident des Europäischen Patentamts, gestützt auf Artikel 10 EPÜ, Regeln 27a (1) und (2) und 49 (1) EPÜ sowie Regel 13<sup>ter</sup> PCT, beschließt:

### Artikel 1

*Standardisierte Darstellung von Nucleotid- und Aminosäuresequenzen in europäischen Patentanmeldungen*

Hat die in einer europäischen Patentanmeldung offenbarte Erfindung Nucleotid- oder Aminosäuresequenzen zum Gegenstand, so hat die Beschreibung ein Sequenzprotokoll zu enthalten, das dem in **Anhang I** wiedergegebenen WIPO-Standard ST. 23 entspricht

### Artikel 2

*Einreichung von Sequenzprotokollen auf elektronischem Datenträger*

(1) Zusätzlich ist ein Datenträger mit dem Sequenzprotokoll gemäß Artikel 1 einzureichen, der den in **Anhang II** enthaltenen Vorschriften entspricht.

(2) Wird ein Sequenzprotokoll nach seiner Einreichung beim Europäischen Patentamt berichtigt, so ist ein neuer Datenträger mit dem Sequenzprotokoll in der geänderten Fassung einzureichen

(3) Jedem Datenträger ist eine Erklärung des Anmelders beizufügen, daß die auf dem Datenträger gespeicherte Information mit dem schriftlichen Sequenzprotokoll übereinstimmt

(4) Ist ein Datenträger beschädigt oder aus einem anderen Grund nicht lesbar, so hat der Anmelder einen neuen, einwandfrei lesbaren Datenträger einzureichen

### Artikel 3

*Beseitigung von Mängeln*

Wird den in Regel 27a EPÜ und in den Artikeln 1 und 2 dieses Beschlusses genannten Erfordernissen nicht entsprochen, so teilt die Eingangsstelle dies dem Anmelder mit und fordert ihn auf, die festgestellten Mängel innerhalb einer nicht verlängerbaren Frist von zwei Monaten zu beseitigen (Regel 41 (1) EPÜ).

## Decision of the President of the European Patent Office dated 11 December 1992 concerning the representation of nucleotide and amino acid sequences in patent applications and the filing of sequence listings\*

The President of the European Patent Office, having regard to Article 10 EPC, Rules 27a(1) and (2) and 49(1) EPC and Rule 13<sup>ter</sup> PCT, has decided as follows:

### Article 1

*Standardised representation of nucleotide and amino acid sequences in European patent applications*

If the subject-matter of the invention disclosed in a European patent application is nucleotide or amino acid sequences, the description shall contain a sequence listing complying with WIPO Standard ST.23 reproduced in **Annex I**.

### Article 2

*Filing of sequence listings on electronic data carriers*

(1) In addition, a data carrier shall be filed containing the sequence listing under Article 1 and complying with the rules laid down in **Annex II**

(2) If a sequence listing is corrected after being filed with the European Patent Office, a new data carrier containing the amended sequence listing shall be filed.

(3) Each data carrier shall be accompanied by a statement by the applicant that the information recorded on the data carrier is identical to the written sequence listing.

(4) If a data carrier is damaged or unreadable for some other reason, the applicant shall submit a new, perfectly readable data carrier.

### Article 3

*Rectification of deficiencies*

If the requirements laid down in Rule 27a EPC and Articles 1 and 2 of this Decision are not met, the Receiving Section shall inform the applicant accordingly and invite him to remedy the disclosed deficiencies within a non-extendable period of two months (Rule 41(1) EPC)

## Décision du Président de l'Office européen des brevets, en date du 11 décembre 1992, relative à la représentation de séquences de nucléotides et d'acides aminés dans les demandes de brevet et au dépôt de listes de séquences\*

Le Président de l'Office européen des brevets, vu l'article 10 CBE, les règles 27bis(1) et (2) et 49 (1) CBE ainsi que la règle 13<sup>ter</sup> PCT, décide:

### Article premier

*Représentation normalisée de séquences de nucléotides et d'acides aminés dans les demandes de brevet européen*

Si l'invention exposée dans une demande de brevet européen a pour objet des séquences de nucléotides ou d'acides aminés, la description doit contenir une liste de séquences conforme à la norme ST.23 de l'OMPI, reproduite en **Annexe I**.

### Article 2

*Dépôt de listes de séquences sur support électronique de données*

(1) En outre, il y a lieu de déposer un support de données contenant la liste de séquences visée à l'article premier et satisfaisant aux prescriptions figurant à l'**Annexe II**.

(2) Si une liste de séquences est rectifiée après avoir été déposée auprès de l'Office européen des brevets, il y a lieu de déposer un nouveau support de données contenant la liste de séquences modifiée.

(3) Tout support de données doit être accompagné d'une déclaration du demandeur selon laquelle l'information figurant sur ce support est identique à celle que contient la liste écrite

(4) Si un support de données est endommagé ou rendu illisible pour toute autre raison, le demandeur doit déposer un nouveau support de données parfaitement lisible

### Article 3

*Correction d'irrégularités*

S'il n'est pas satisfait aux exigences visées à la règle 27bis CBE et aux articles premier et 2 de la présente décision, la section de dépôt le signale au demandeur et l'invite à remédier aux irrégularités constatées dans un délai non prorogeable de deux mois (règle 41(1) CBE).

\* Siehe dazu den Beschluß des Verwaltungsrats vom 5 Juni 1992 zur Änderung der Ausführungsordnung zum EPÜ, ABI EPA 1992. 342. sowie die Mitteilung des EPA vom 11. Dezember 1992 in dieser Beilage zum Amtsblatt auf Seite 3

\* See the Decision of the Administrative Council of 5 June 1992 amending the Implementing Regulations to the EPC, OJ EPO 1992. 342. and the Notice from the EPO dated 11 December 1992 in this Supplement to the Official Journal, page 3

\* Voir à ce propos la décision du Conseil d'administration du 5 juin 1992 modifiant le règlement d'exécution de la CBE, JO OEB 1992. 342. et le communiqué de l'OEB, en date du 11 décembre 1992, publié dans le présent supplément au Journal officiel, page 3

**Artikel 4**

*Einreichung von Sequenzprotokollen beim EPA als Internationaler Recherchenbehörde*

In bezug auf internationale Anmeldungen nach dem PCT, für die nach Regel 5.2 PCT ein Sequenzprotokoll vorgeschrieben ist, verfährt das Europäische Patentamt als Internationale Recherchenbehörde nach Regel 13<sup>ter</sup> 1 PCT und den Abschnitten 208 und 513 der Verwaltungsrichtlinien zum PCT. Artikel 2 (2) bis (4) dieses Beschlusses ist entsprechend anzuwenden

**Artikel 5**

*Veröffentlichung von Sequenzprotokollen*

Ein am Anmeldetag der europäischen Patentanmeldung eingereichtes Sequenzprotokoll wird als Bestandteil der Beschreibung zusammen mit den Anmeldungsunterlagen veröffentlicht. Ein nach dem Anmeldetag, jedoch innerhalb der vom Amt bestimmten Frist eingereichtes oder berichtiges Sequenzprotokoll wird, wenn es vor dem Zeitpunkt vorliegt, zu dem die technischen Vorbereitungen für die Veröffentlichung als abgeschlossen gelten, als Anlage zu den Anmeldungsunterlagen, andernfalls, bei Patenterteilung, als Anlage zur europäischen Patentschrift veröffentlicht.

**Artikel 6**

*Inkrafttreten*

Dieser Beschluß tritt am **1. Januar 1993** in Kraft

Geschehen zu München, am 11 Dezember 1992

PAUL BRAENDLI  
Präsident

**Article 4**

*Filing of sequence listings with the EPO as International Searching Authority*

With regard to international applications under the PCT, for which a sequence listing is prescribed under Rule 5.2 PCT, the European Patent Office shall act as International Searching Authority in accordance with Rule 13<sup>ter</sup>.1 PCT and Sections 208 and 513 of the PCT Administrative Instructions. Article 2(2) to (4) of this Decision shall apply *mutatis mutandis*.

**Article 5**

*Publication of sequence listings*

A sequence listing filed on the date of filing of the European patent application shall be published together with the application documents as part of the description. A sequence listing filed or corrected after the date of filing but within the period specified by the Office shall be published as an annex to the application documents, provided it is received before the date on which the technical preparations for publication are deemed to have been completed; otherwise, it shall be published as an annex to the European patent specification when the patent is granted

**Article 6**

*Entry into force*

This decision shall enter into force on **1 January 1993**.

Done at Munich, 11 December 1992.

PAUL BRAENDLI  
President

**Article 4**

*Dépôt de listes de séquences auprès de l'OEB agissant en qualité d'administration chargée de la recherche internationale*

En ce qui concerne les demandes internationales déposées au titre du PCT pour lesquelles la règle 5.2 PCT prescrit une liste de séquences, l'Office européen des brevets, agissant en qualité d'administration chargée de la recherche internationale, procède conformément aux dispositions de la règle 13<sup>ter</sup>.1 PCT et des instructions administratives 208 et 513 du PCT L'article 2(2) à (4) de la présente décision s'applique.

**Article 5**

*Publication de listes de séquences*

Une liste de séquences déposée à la date de dépôt de la demande de brevet européen est publiée comme élément de la description avec les documents de la demande. Une liste de séquences déposée ou rectifiée après la date de dépôt, mais dans le délai imparti par l'Office, est, pour autant qu'elle soit disponible avant la date à laquelle les préparatifs techniques entrepris en vue de la publication de la demande sont réputés achevés, publiée en annexe aux documents de la demande, ou, si ce n'est pas le cas, au moment de la délivrance du brevet, en annexe au fascicule du brevet européen.

**Article 6**

*Entrée en vigueur*

La présente décision entre en vigueur le **1<sup>er</sup> janvier 1993**.

Fait à Munich, le 11 décembre 1992

PAUL BRAENDLI  
Président

## Mitteilung des Europäischen Patentamts vom 11. Dezember 1992 über die Darstellung von Nucleotid- und Aminosäuresequenzen in Patentanmeldungen und die Einreichung von Sequenzprotokollen

Im Anschluß an die 1991 eingefügten neuen Regeln 5.2 und 13<sup>ter</sup> PCT<sup>1</sup> hat der Verwaltungsrat der Europäischen Patentorganisation mit Beschluß vom 5. Juni 1992<sup>2</sup> die neuen Bestimmungen der Regeln 27a und 104b (3a) in das Europäische Patentübereinkommen mit Wirkung vom 1. Januar 1993 aufgenommen

Mit Beschluß des Präsidenten des EPA vom 11. Dezember 1992<sup>3</sup> sind die notwendigen Ausführungsbestimmungen zu den Regeln 27a EPÜ und 13<sup>ter</sup> PCT erlassen worden

Diese Mitteilung ergänzt diese Beschlüsse und soll die neuen Vorschriften sowie das Verfahren nach dem EPÜ und PCT erläutern

### I. Allgemeines

Mit der Änderung des EPÜ wird die bisher nur empfohlene Darstellung von Nucleotid- und Aminosäuresequenzen in einer international abgestimmten Form<sup>4</sup> ab **1. Januar 1993 zwingend vorgeschrieben**. Darüber hinaus sind Sequenzprotokolle auf einem vorgeschriebenen Datenträger einzureichen

Die neuen Bestimmungen sollen die einheitliche Darstellung von Nucleotid- und Aminosäuresequenzen in europäischen und internationalen Anmeldungen gewährleisten und eine systematische EDV-gestützte Recherche ermöglichen, um die Qualität der europäischen und internationalen Recherchenberichte auch in Zukunft zu sichern. Schließlich wird damit der Aufbau einer internationalen Patent-Sequenzdatenbank in Zusammenarbeit mit dem Patent- und Markenamt der Vereinigten Staaten (USPTO) und dem japanischen Patentamt (JPO) ermöglicht, die der Öffentlichkeit zugänglich sein wird

### II. Europäisches Patenterteilungsverfahren

#### 1 Sequenzprotokoll

1.1 Die Darstellung von Nucleotid- und Aminosäuresequenzen im vorgeschriebenen Sequenzprotokoll muß dem **WIPO-Standard ST.23** (dem Be-

## Notice from the European Patent Office dated 11 December 1992 concerning the representation of nucleotide and amino acid sequences in patent applications and the filing of sequence listings

Following the insertion of the new Rules 5.2 and 13<sup>ter</sup> into the PCT in 1991<sup>1</sup>, the Administrative Council of the European Patent Organisation in its decision of 5 June 1992<sup>2</sup> incorporated the new provisions of Rules 27a and 104b(3)(a) into the European Patent Convention with effect from 1 January 1993.

The necessary implementing provisions for Rules 27a EPC and 13<sup>ter</sup> PCT were laid down in the Decision of the President of the EPO dated 11 December 1992<sup>3</sup>.

This notice supplements the above decisions and is intended to clarify the new rules and the EPC and PCT procedures.

### I. General remarks

The presentation of nucleotide and amino acid sequences in an internationally agreed form, which was hitherto merely recommended<sup>4</sup>, will become **mandatory as from 1 January 1993** when the amendment to the EPC enters into force. In addition, sequence listings must be filed on a prescribed data carrier.

The new provisions are intended to ensure that nucleotide and amino acid sequences are presented in a standardised form in European and international applications and to permit systematic EDP-based searches so that the quality of European and international search reports can continue to be guaranteed in the future. Finally, they will make it possible to set up, in co-operation with the United States Patent and Trademark Office (USPTO) and the Japanese Patent Office (JPO), an international patent sequence database which will be accessible to the public

### II. European patent grant procedure

#### 1. Sequence listing

1.1 The representation of nucleotide and amino acid sequences in the prescribed sequence listing must comply with **WIPO Standard ST.23** (Annex I

## Communiqué de l'Office européen des brevets, en date du 11 décembre 1992, relatif à la représentation de séquences de nucléotides et d'acides aminés dans les demandes de brevet et au dépôt de listes de séquences

A la suite de l'adoption en 1991 des nouvelles règles 5.2 et 13<sup>ter</sup> PCT<sup>1</sup>, le Conseil d'administration de l'Organisation européenne des brevets a, par sa décision du 5 juin 1992<sup>2</sup>, inclus dans la Convention sur le brevet européen, avec effet à compter du 1<sup>er</sup> janvier 1993, les nouvelles dispositions des règles 27bis et 104<sup>ter</sup> (3bis).

La décision du Président de l'Office européen des brevets, en date du 11 décembre 1992<sup>3</sup>, arrête les modalités d'application des règles 27bis CBE et 13<sup>ter</sup> PCT.

Le présent communiqué complète ces décisions et donne des explications sur les nouvelles dispositions et la procédure prévue par la CBE et le PCT.

### I. Généralités

Du fait de la modification de la CBE, la représentation de nucléotides et d'acides aminés sous une forme harmonisée au niveau international, qui ne faisait jusqu'à présent que l'objet d'une recommandation<sup>4</sup>, deviendra **obligatoire à compter du 1<sup>er</sup> janvier 1993**. En outre, des listes de séquences devront être déposées sur des supports de données conformes à une norme imposée.

Les nouvelles dispositions doivent garantir une représentation uniforme des séquences de nucléotides et d'acides aminés dans les demandes de brevet européen et les demandes internationales et permettre une recherche systématique assistée par ordinateur, afin d'assurer également à l'avenir la qualité des rapports de recherche européenne et internationale. Enfin, elles contribuent à la constitution, avec la collaboration de l'Office des brevets et des marques des Etats-Unis (USPTO) et de l'Office japonais des brevets (JPO), d'une banque internationale de données relatives aux séquences contenues dans les brevets qui sera accessible **au public**.

### II. Procédure de délivrance des brevets européens

#### 1 Liste de séquences

1.1 La représentation de séquences de nucléotides et d'acides aminés dans la liste de séquences prescrite doit être conforme à la **norme ST.23 de l'OM-**

<sup>1</sup> Abgedruckt in der PCT Gazette Nr. 20/1991 vom 5. September 1991

<sup>2</sup> S. ABI EPA 1992 342

<sup>3</sup> Abgedruckt in dieser Beilage zum Amtsblatt auf Seite 1

<sup>4</sup> S. Mitteilung des EPA in der Beilage zum ABI EPA 12/1989

<sup>1</sup> Printed in PCT Gazette No. 20/1991 of 5 September 1991

<sup>2</sup> See OJ EPO 1992 342.

<sup>3</sup> Printed in this Supplement to the Official Journal, page 1

<sup>4</sup> See Notice from the EPO in the Supplement to OJ EPO 12/1989

<sup>1</sup> Cf. Gazette du PCT n° 20/1991 du 5 septembre 1991

<sup>2</sup> Cf. JO OEB 1992, 342

<sup>3</sup> Publiée dans le présent supplément au Journal officiel, page 1

<sup>4</sup> Cf. communiqué de l'OEB publié au supplément au JO OEB 12/1989

schluß des Präsidenten des EPA vom 11. Dezember 1992 als Anhang I beige-fügt) entsprechen. Dieser Standard ist das Ergebnis der Aktualisierung der international abgestimmten Regeln, die in der Beilage zum Amtsblatt EPA 12/1989 veröffentlicht worden waren.

1.2 Ein Sequenzprotokoll ist nur für solche Sequenzen einzureichen, die der Definition laut WIPO-Standard ST. 23 Nummer 3 iii) entsprechen und sich auf die in der Anmeldung offenbarte Erfindung beziehen.

Ein Sequenzprotokoll ist nicht erforderlich, soweit in der Beschreibung auf Sequenzen Bezug genommen wird, die zum Stand der Technik gehören. In diesem Fall ist die Sequenz jedoch hinreichend zu identifizieren, gegebenenfalls unter Angabe der Datenbank-Zugangsnummer, so daß sie leicht auffindbar ist.

## 2. Elektronische Datenträger

2.1 Mit seinem Beschluß vom 11. Dezember 1992 hat der Präsident des EPA die Einreichung des Sequenzprotokolls auf einem Datenträger zusätzlich zur schriftlichen Form vorgeschrieben. Die näheren Einzelheiten über die zugelassenen Datenträger, ihre Kennzeichnung, das Format der darauf gespeicherten Information und die erforderliche Hardware sind den **Vorschriften über die Einreichung von Sequenzprotokollen auf elektronischem Datenträger** in Anhang II dieses Beschlusses zu entnehmen. Diese Vorschriften entsprechen dem WIPO-Standard ST 24.

2.2 Das EPA empfiehlt die Verwendung des **Patentin-Programms**, das den Anmeldern kostenlos zur Verfügung steht<sup>5</sup>. Durch Verwendung dieses Programms in der jeweils gültigen Fassung kann der Anmelder sicherstellen, daß die Formerfordernisse nach Regel 27a (1) EPÜ erfüllt sind.

### 3 Erklärung gemäß Regel 27a (2) EPÜ

Der vorgeschriebene Datenträger wird für die EDV-gestützte europäische oder internationale Recherche sowie, nach der Veröffentlichung der Anmeldung, für die Aktualisierung der im Aufbau befindlichen internationalen Patent-Sequenzdatenbank benötigt.

Die auf dem Datenträger gespeicherte Information muß mit dem schriftlichen Sequenzprotokoll übereinstimmen. Der Anmelder oder sein Vertreter hat daher eine entsprechende Erklärung gemäß Regel 27a (2) EPÜ zusammen mit dem Datenträger einzureichen. Eine solche Erklärung ist im Erteilungs-

to the Decision of the President of the EPO dated 11 December 1992). This standard is the product of updating the internationally agreed rules published in the Supplement to EPO Official Journal 12/1989.

1.2 A sequence listing need only be filed for those sequences matching the definition given in WIPO Standard ST.23 paragraph 3(iii) and relating to the invention disclosed in the application.

A sequence listing is not required if the description refers to sequences which belong to the prior art. In this case, however, the sequence must be identified adequately—where applicable, by stating the database accession number—so that it can be found easily.

## 2. Electronic data carriers

2.1 In his Decision of 11 December 1992 the President of the EPO laid down that the sequence listing must be submitted on a data carrier as well as in written form. Further details with regard to authorised data carriers, their identification, the format for the information recorded on them and the necessary hardware are set out in the **Rules for the filing of sequence listings on electronic data carriers** in Annex II to this Decision. These rules conform to WIPO Standard ST.24.

2.2 The EPO recommends the use of the **Patentin program** which is available to applicants free of charge<sup>5</sup>. By using the applicable version of this program the applicant can ensure that the physical requirements under Rule 27a(1) EPC are met.

### 3 Statement under Rule 27a(2) EPC

The prescribed data carrier is required for EDP-based European or international searches and, after publication of the application, for updating the international patent sequence database now in the process of being set up.

The information recorded on the data carrier must be identical to the written sequence listing. The applicant or his representative must therefore file together with the data carrier a statement in accordance with Rule 27a(2) EPC. The statement is contained in the Request for Grant (EPO Form 1001,

PI (jointe en Annexe I à la décision du Président de l'OEB en date du 11 décembre 1992). Cette norme résulte de l'actualisation des règles harmonisées au niveau international, qui ont été publiées dans le supplément au Journal officiel de l'OEB n° 12/1989.

1.2 Il ne faut déposer une liste de séquences que pour les séquences correspondant à la définition figurant au paragraphe 3 iii) de la norme ST.23 de l'OMPI et se rapportant à l'invention exposée dans la demande.

Une liste de séquences n'est pas nécessaire s'il est fait référence dans la description à des séquences comprises dans l'état de la technique. Il faut cependant dans ce cas identifier suffisamment la séquence, en indiquant, le cas échéant, son numéro d'accès à la banque de données, afin que l'on puisse la retrouver facilement.

## 2. Supports électroniques de données

2.1 Par sa décision en date du 11 décembre 1992, le Président de l'OEB a imposé la production de la liste de séquences sur un support de données, en plus de la liste écrite. Les détails concernant les supports de données admis, leur identification, le format des informations y figurant et le matériel nécessaire sont contenus dans les **prescriptions relatives au dépôt de listes de séquences sur des supports électroniques de données** exposées à l'Annexe II de cette décision. Ces prescriptions correspondent à la norme ST.24 de l'OMPI.

2.2 L'OEB recommande l'utilisation du **programme Patentin**, qui est mis gratuitement à la disposition des demandeurs<sup>5</sup>. En utilisant ce programme dans sa version en vigueur, le demandeur peut s'assurer que les conditions de forme de la règle 27bis (1) CBE sont respectées.

### 3. Déclaration prévue à la règle 27bis (2) CBE

Le support de données prescrit est nécessaire pour la recherche européenne ou internationale assistée par ordinateur ainsi que, après la publication de la demande, pour la mise à jour de la banque internationale de données relatives aux séquences contenues dans les brevets qui est en cours de création.

L'information figurant sur ce support doit être identique à celle que contient la liste écrite de séquences. C'est pourquoi le demandeur ou son mandataire est tenu de joindre au support de données une déclaration en ce sens, conformément à la règle 27bis(2) CBE. Une telle déclaration figure dans

<sup>5</sup> S. Mitteilung des EPA in ABI EPA 1992, 312

<sup>5</sup> See Notice from the EPO in OJ EPO 1992, 312.

<sup>5</sup> Cf. communiqué de l'OEB publié au JO OEB 1992 312

antrag enthalten (EPA Form 1001, Feld 31)<sup>6</sup>. Diese Erklärung kann als Muster auch bei Nachreichung des Datenträgers verwendet werden.

Stellt die Prüfungsabteilung fest, daß der eingereichte Datenträger das Identitätserfordernis nicht erfüllt und somit die Erklärung gemäß Regel 27a (2) EPÜ unrichtig ist, so kann dies zur Zurückweisung der Anmeldung gemäß Artikel 97 (1) EPÜ führen. Darüber hinaus kann ein anderer Stand der Technik einschlägig sein als er der Prüfung zugrunde gelegt wurde. Dies kann zu einer anderen Beurteilung der Patentierbarkeitsvoraussetzungen auch nach Patenterteilung führen.

#### 4. Nachreichung eines Sequenzprotokolls

4.1 Ist ein Sequenzprotokoll einzureichen, so ist es grundsätzlich den Anmeldeunterlagen als Teil der Beschreibung beizufügen (vgl. WIPO-Standard ST.23 Nummer 4).

4.2 Ein nach dem Anmeldetag eingereichtes Sequenzprotokoll ist nicht Bestandteil der Beschreibung und kann daher für die Bestimmung des ursprünglichen Offenbarungsgehalts der Anmeldung nicht herangezogen werden (Regel 27a (4) EPÜ).

4.3 Gleichwohl nimmt das EPA auch nach dem Anmeldetag eingereichte Sequenzprotokolle entgegen, sofern der vorgeschriebene Datenträger und die Erklärungen gemäß Regel 27a (2) und (3) EPÜ vorliegen, weil dies für die Durchführung der Recherche und die Fortschreibung der internationalen Patent-Sequenzdatenbank unerlässlich ist.

4.4 Ist der Anmelder nicht imstande, das Sequenzprotokoll als Teil der Beschreibung einzureichen, sollte er es von sich aus baldmöglichst nachreichen und die weiteren Erfordernisse nach Regel 27a (2) und (3) EPÜ erfüllen (Einreichung des vorgeschriebenen Datenträgers und der vorgesehenen Erklärungen, siehe auch Punkt II.5). Andernfalls fordert ihn die Eingangsstelle auf, die festgestellten Mängel innerhalb einer **nicht verlängerbaren Frist von 2 Monaten** (Art. 91 (2), Regel 41 (1) EPÜ und Artikel 3 des Beschlusses des Präsidenten des EPA vom 11. Dezember 1992) zu beheben.

#### 5. Erklärung gemäß Regel 27a (3) EPÜ

Ein nachgereichtes Sequenzprotokoll darf nichts enthalten, was über den Inhalt der Anmeldung in der ursprünglich eingereichten Fassung hinausgeht, da die Recherche auf der Grundlage des nachgereichten Sequenzprotokolls durchgeführt werden soll. Mit anderen Worten soll das nachgereichte Sequenzprotokoll nur diejenige Se-

Section 31)<sup>6</sup>. This statement may also be used as a model when the data carrier is filed subsequently.

If the Examining Division finds that the data carrier submitted does not comply with the requirement that it be identical and the statement under Rule 27a(2) EPC is therefore incorrect, this may lead to refusal of the application under Article 97(1) EPC. Furthermore, a different prior art than that on which the examination was based may be relevant. This may lead to a different assessment of the conditions for patentability even after a patent has been granted.

#### 4. Subsequent filing of a sequence listing

4.1 If a sequence listing has to be filed, it should normally be annexed to the application documents as part of the description (see WIPO Standard ST.23, paragraph 4).

4.2 A sequence listing filed after the date of filing does not form part of the description and cannot therefore be used in determining the original disclosed content of the application (Rule 27a(4) EPC).

4.3 Nevertheless, the EPO accepts sequence listings filed after the date of filing so long as the prescribed data carrier and the statements under Rule 27a(2) and (3) EPC have been received because they are absolutely essential for carrying out the search and updating the international patent sequence database.

4.4 If the applicant is unable to file the sequence listing as part of the description, he should file it without prompting as soon as possible thereafter and comply with the other requirements under Rule 27a(2) and (3) EPC (see also point II.5 for submission of the prescribed data carrier and statements). Otherwise, the Receiving Section will invite him to remedy the deficiencies within a **non-extendable period of two months** (Art. 91(2), Rule 41(1) EPC and Article 3 of the Decision of the President of the EPO dated 11 December 1992).

#### 5. Statement under Rule 27a(3) EPC

A subsequently filed sequence listing may not include any matter which goes beyond the content of the application as filed, since the search is to be carried out on the basis of the subsequently filed sequence listing. In other words, the latter must only contain the sequence information — in a different form — that was already in the

la requête en délivrance (OEB Form 1001, rubrique 31)<sup>6</sup>. Cette déclaration peut également être utilisée comme modèle en cas de production ultérieure du support de données

Si la division d'examen constate que le support de données ne satisfait pas à l'exigence d'identité avec la liste écrite et que la déclaration selon la règle 27bis (2) CBE est donc inexacte, la demande peut être rejetée conformément à l'article 97(1) CBE. Il se peut en outre qu'un état de la technique autre que celui qui a été pris en compte pour l'examen soit pertinent et qu'il en résulte une appréciation divergente des conditions de brevetabilité même après la délivrance du brevet.

#### 4. Dépôt ultérieur d'une liste de séquences

4.1 Toute liste de séquences doit en principe être jointe aux pièces de la demande comme faisant partie de la description (cf. norme OMPI ST.23, paragraphe 4).

4.2 Une liste de séquences déposée après la date de dépôt ne fait pas partie de la description et ne peut donc servir à déterminer le contenu initial de la demande (règle 27bis (4) CBE).

4.3 L'OEB accepte néanmoins les listes déposées après la date de dépôt, dans la mesure où le support de données prescrit et les déclarations selon la règle 27bis (2) et (3) CBE sont disponibles, étant donné que ces informations sont indispensables pour effectuer la recherche et actualiser la banque internationale de données relatives aux séquences contenues dans les brevets.

4.4 Si le demandeur n'est pas en mesure de déposer la liste de séquences comme partie de la description, il a intérêt à en effectuer spontanément le dépôt ultérieurement dans les meilleurs délais, en veillant à remplir les autres conditions énoncées à la règle 27bis (2) et (3) (production du support de données prescrit et des déclarations prévues; cf. également point II.5). Sinon, la section de dépôt l'invite à remédier dans un **délai non prorogable de deux mois** aux irrégularités constatées (art. 91 (2), règle 41 (1) CBE et art. 3 de la décision du Président de l'OEB en date du 11 décembre 1992).

#### 5. Déclaration prévue à la règle 27bis (3) CBE

Une liste de séquences déposée ultérieurement ne doit contenir aucun élément s'étendant au-delà du contenu de la demande telle qu'elle a été déposée, puisque la recherche est censée être exécutée sur la base de cette liste. En d'autres termes, la liste de séquences déposée ultérieurement ne doit contenir — sous une autre forme —

<sup>6</sup> Abgedruckt im ABI EPA 1992. 758

<sup>6</sup> Printed in OJ EPO 1992, 758.

<sup>6</sup> Reproduite au JO OEB 1992, 758.

quenzinformation — in einer anderen Form — enthalten, die in den ursprünglichen Anmeldungsunterlagen bereits vorhanden war. Der Anmelder hat eine entsprechende Erklärung gemäß Regel 27a (3) EPÜ einzureichen. Es wird empfohlen, hierfür die in der Fußnote 7 abgedruckte Mustererklärung<sup>7</sup> zu verwenden. Für den ebenfalls einzureichenden Datenträger gilt Punkt II.3.

## 6. Beschädigte oder nicht lesbare Datenträger

Ist ein beim EPA eingereichter Datenträger beschädigt oder aus einem anderen Grund nicht lesbar, wird der Anmelder aufgefordert, einen neuen, einwandfrei lesbaren Datenträger innerhalb einer **nicht verlängerbaren Frist von zwei Monaten** einzureichen.

### 7 Rechtsverlust

Werden die Erfordernisse der Regel 27a EPÜ in Verbindung mit dem Beschluß des Präsidenten des EPA vom 11. Dezember 1992 nicht rechtzeitig, gegebenenfalls nach einer Aufforderung der Eingangsstelle gemäß Regel 41 (1) EPÜ erfüllt, so wird die Anmeldung zurückgewiesen (Art. 91 (1) b) und (3) EPÜ).

Der Anmelder kann die Weiterbehandlung der Anmeldung nach Artikel 121 EPU beantragen.

## III. Verfahren vor dem EPA als Internationaler Recherchenbehörde oder als mit der internationalen vorläufigen Prüfung beauftragter Behörde

1 Sequenzprotokolle in maschinenlesbarer Form werden für die Durchführung der internationalen Recherche ebenso benötigt wie im Rahmen der europäischen Recherche.

Liegt ein dem WIPO-Standard ST.23 entsprechendes Sequenzprotokoll (Regel 5.2 PCT, Abschnitt 208 der Verwaltungsrichtlinien zum PCT; Anhang I zum Beschluß des Präsidenten des EPA vom 11. Dezember 1992) zu Beginn der internationalen Recherche nicht oder nicht auf elektronischem Datenträger (Regel 13<sup>ter</sup>.1 a) ii) PCT, Abschnitt 208 der Verwaltungsrichtlinien zum PCT, Anhang II zum Beschluß des Präsidenten des EPA vom 11. Dezember 1992) vor, so wird der Anmelder vom EPA aufgefordert, das Sequenzprotokoll und/oder den vorgeschriebenen Datenträger innerhalb einer **nicht verlängerbaren Frist von zwei Monaten** nachzureichen (Regel 13<sup>ter</sup>.1 a) PCT).

2 Bei Nachreichung des Sequenzprotokolls und/oder des Datenträgers hat der Anmelder zu erklären, daß das nachgereichte Sequenzprotokoll nicht über den Inhalt der Anmeldung in der ursprünglich eingereichten Fassung

<sup>7</sup> "Es wird hiermit erklärt, daß das Sequenzprotokoll nicht über den Inhalt der Anmeldung in der ursprünglich eingereichten Fassung hinausgeht."

original application documents. The applicant must file a statement to that effect in accordance with Rule 27a(3) EPC. It is recommended that the specimen statement<sup>7</sup> printed under footnote 7 be used for this purpose. Point II.3 is applicable for the data carrier which must also be submitted.

## 6. Damaged or unreadable data carriers

If a data carrier submitted to the EPO is damaged or unreadable for some other reason, the applicant is invited to submit a new, perfectly readable data carrier within a **non-extendable period of two months**.

### 7 Loss of rights

If the requirements laid down in Rule 27a EPC in conjunction with the Decision of the President of the EPO dated 11 December 1992 are not met in good time — if applicable, after an invitation from the Receiving Section under Rule 41(1) EPC — the application is refused (Art. 91(1)(b) and (3) EPC).

The applicant may request further processing of the application under Article 121 EPC.

## III. Proceedings before the EPO as International Searching Authority or International Preliminary Examining Authority

1 Sequence listings in machine readable form are just as necessary in carrying out the international search as they are for the European search.

If at the start of the international search a sequence listing complying with WIPO Standard ST.23 (Rule 5.2 PCT, Section 208 of the PCT Administrative Instructions; Annex I to the Decision of the President of the EPO dated 11 December 1992) has not been received or has not been received on an electronic data carrier (Rule 13<sup>ter</sup>.1(a)(ii) PCT, Section 208 of the PCT Administrative Instructions; Annex II to the Decision of the President of the EPO dated 11 December 1992), the EPO invites the applicant to file the sequence listing and/or the prescribed data carrier within a **non-extendable period of two months** (Rule 13<sup>ter</sup>.1(a) PCT).

2. If the sequence listing and/or data carrier is filed subsequently, the applicant must state that the sequence listing so filed does not include matter which goes beyond the content of the application as filed and/or that the

<sup>7</sup> "It is hereby stated that the sequence listing does not include matter which goes beyond the content of the application as filed."

que les informations concernant les séquences qui figuraient déjà dans les pièces initiales de la demande. Le demandeur est tenu de produire une déclaration en ce sens conformément à la règle 27bis (3) CBE. Il est recommandé d'utiliser à cette fin la déclaration type<sup>7</sup> reproduite à la note 7 en bas de page. Le point II.3 est applicable au support de données que le demandeur doit en outre produire.

## 6. Supports de données endommagés ou illisibles

Si un support de données reçu par l'OEB est endommagé ou rendu illisible pour toute autre raison, le demandeur est invité à produire un nouveau support parfaitement lisible dans un **délai non prorogable de deux mois**.

### 7 Perte de droits

Si le demandeur ne satisfait pas en temps utile aux conditions énoncées à la règle 27bis CBE en liaison avec la décision du Président de l'OEB en date du 11 décembre 1992, le cas échéant après y avoir été invité par la section de dépôt conformément à la règle 41(1) CBE, la demande est rejetée (art. 91 (1) b) et (3) CBE).

Le demandeur peut demander la poursuite de la procédure conformément à l'article 121 CBE.

## III. Procédure devant l'OEB agissant en qualité d'administration chargée de la recherche internationale ou d'administration chargée de l'examen préliminaire international

1. Les listes de séquences sous forme déchiffable par machine sont nécessaires tant pour la recherche internationale que pour la recherche européenne.

Si une liste de séquences conforme à la norme ST.23 de l'OMPI (règle 5.2 PCT, instruction administrative 208 du PCT; Annexe I de la décision du Président de l'OEB en date du 11 décembre 1992) n'est pas parvenue à l'OEB ou bien si elle n'est pas disponible sur support électronique de données (règle 13<sup>ter</sup>.1 a) ii) PCT, instruction administrative 208 du PCT; Annexe II de la décision du Président de l'OEB en date du 11 décembre 1992) lorsque la recherche internationale est entreprise, le demandeur est invité par l'OEB à fournir dans un **délai non prorogable de deux mois** la liste et/ou le support de données prescrit (règle 13<sup>ter</sup>.1 a) PCT).

2. Lorsqu'il fournit ultérieurement la liste de séquences et/ou le support de données, le demandeur est tenu de déclarer que cette liste ne s'étend pas au-delà du contenu de la demande telle qu'elle a été déposée et/ou que

<sup>7</sup> "Je déclare que la liste de séquences ne s'étend pas au-delà du contenu de la demande telle qu'elle a été déposée."

hinausgeht und/oder die auf dem Datenträger gespeicherte Information mit dem schriftlichen Sequenzprotokoll übereinstimmt (Regel 13<sup>ter</sup>.1 b) PCT, Art. 4 und 2 (3) des Beschlusses des Präsidenten des EPA vom 11. Dezember 1992).

3. Das EPA beabsichtigt nicht, einen Umschreibedienst im Sinne von Regel 13<sup>ter</sup>.1 a) ii) PCT zur Konvertierung der schriftlichen Sequenzprotokolle in eine maschinenlesbare Form gegen Gebühr einzurichten

4. Werden diese Erfordernisse nicht oder nicht rechtzeitig erfüllt, so führt das EPA keine internationale Recherche durch, soweit eine sinnvolle Recherche nicht möglich ist (Regel 13<sup>ter</sup>.1 c) PCT). In diesem Fall wird der internationale Recherchenbericht ganz oder teilweise durch die Erklärung gemäß Artikel 17 (2) a) ii) PCT ersetzt

5. Ist der internationale Recherchenbericht aus den vorgenannten Gründen ganz oder teilweise nicht erstellt worden, so führt das EPA als mit der internationalen vorläufigen Prüfung beauftragte Behörde keine internationale vorläufige Prüfung durch, soweit eine sinnvolle Prüfung nicht möglich ist (Regel 66.1 e) PCT). Dies wird dem Anmelder mitgeteilt und im internationalen vorläufigen Prüfungsbericht vermerkt (Regeln 66.2 a) vii) und 70.12 iv) PCT). Eine Aufforderung zur Einreichung eines Sequenzprotokolls gemäß Abschnitt 610 der Verwaltungsrichtlinien zum PCT ergeht nicht.

#### IV. Eintritt der Euro-PCT-Anmeldungen in die regionale Phase vor dem EPA

1. Liegt ein nach Regel 5.2 PCT vorgeschriebenes Sequenzprotokoll bei Ablauf der internationalen Phase dem EPA nicht vor, entspricht es nicht dem WIPO-Standard ST.23, ist es nicht auf dem vorgeschriebenen Datenträger (Anhang II zum Beschlusse des Präsidenten des EPA vom 11. Dezember 1992) eingereicht worden oder nicht in einer Amtssprache (vgl. Regel 1 (1) EPÜ) des EPA abgefaßt, so wird der Anmelder aufgefordert, die festgestellten Mängel innerhalb einer **nicht verlängerbaren Frist von zwei Monaten** zu beheben (Regel 104b (3a) EPÜ). Gegebenenfalls wird der Anmelder aufgefordert, gleichzeitig zu erklären, daß das nachgereichte Sequenzprotokoll nicht über den Inhalt der Anmeldung in der ursprünglich eingereichten Form hinausgeht und/oder die auf dem Datenträger gespeicherte Information mit dem schriftlichen Sequenzprotokoll übereinstimmt

2. Werden bestehende Mängel nach einer solchen Aufforderung nicht oder nicht rechtzeitig beseitigt, so wird die Anmeldung zurückgewiesen (Art. 91 (3) EPÜ).

information recorded on the data carrier is identical to the written sequence listing (Rule 13<sup>ter</sup>.1(b) PCT, Art. 4 and 2(3) of the Decision of the President of the EPO dated 11 December 1992).

3. The EPO does not intend to set up a transcription service within the meaning of Rule 13<sup>ter</sup>.1(a)(ii) PCT for converting the written sequence listings into machine readable form on payment of a fee.

4. If these requirements are not met or not met in due time, the EPO does not perform the international search where a meaningful search cannot be carried out (Rule 13<sup>ter</sup>.1 (c) PCT). In this case the international search report is replaced in full or in part by the statement under Article 17 (2)(a)(ii) PCT.

5. If the international search report has not been drawn up in full or in part for the aforementioned reasons, the EPO as International Preliminary Examining Authority does not perform the international preliminary examination where a meaningful examination cannot be carried out (Rule 66.1(e) PCT). This is notified to the applicant and stated in the international preliminary examination report (Rules 66.2(a)(vii) and 70.12(iv) PCT). An invitation under Section 610 of the PCT Administrative Instructions to file a sequence listing is not issued.

#### IV. Entry of Euro-PCT applications into the regional phase before the EPO

1. If on expiry of the international phase the sequence listing prescribed in Rule 5.2 PCT is not available to the EPO, does not comply with WIPO Standard ST.23, has not been filed on the prescribed data carrier (Annex II to the Decision of the President of the EPO dated 11 December 1992) or has not been drawn up in one of the official languages of the EPO (cf. Rule 1(1) EPC), the applicant is invited to remedy such deficiencies within a **non-extendable period of two months** (Rule 104b(3)(a) EPC). Where appropriate, the applicant is invited at the same time to state that the subsequently filed sequence listing does not include matter which goes beyond the content of the application as filed and/or that the information recorded on the data carrier is identical to the written sequence listing.

2. If existing deficiencies are not remedied or not remedied in due time after such an invitation, the application is refused (Art. 91(3) EPC).

l'information figurant sur ce support est identique à celle que contient la liste écrite (règle 13<sup>ter</sup>.1 b) PCT, art. 4 et 2(3) de la décision du Président de l'OEB, en date du 11 décembre 1992).

3. L'OEB n'envisage pas de créer un service de transcription au sens de la règle 13<sup>ter</sup>.1 a)ii) PCT pour convertir contre versement d'une taxe les listes écrites de séquences en une forme déchiffable par machine.

4. Si ces conditions ne sont pas remplies ou ne le sont pas dans les délais, l'OEB ne procède pas à la recherche internationale dans la mesure où une recherche significative ne peut pas être effectuée (règle 13<sup>ter</sup>.1 c) PCT). Dans ce cas, le rapport de recherche internationale est remplacé en tout ou en partie par la déclaration prévue à l'article 17(2)a)ii) PCT.

5. Si le rapport de recherche internationale n'a pas été établi, en tout ou en partie, pour les raisons qui viennent d'être énoncées, l'OEB agissant en qualité d'administration chargée de l'examen préliminaire international ne procède pas à l'examen préliminaire international dans la mesure où un examen significatif ne peut pas être effectué (règle 66.1 e) PCT). L'OEB en informe le demandeur et l'indique dans le rapport d'examen préliminaire international (règles 66.2 a) vii) et 70.12 iv) PCT). Il n'est pas envoyé d'invitation à déposer une liste de séquences, telle que prévue par l'instruction administrative 610 du PCT.

#### IV. Entrée des demandes euro-PCT dans la phase régionale devant l'OEB

1. Si, au terme de la phase internationale, la liste de séquences visée à la règle 5.2 PCT n'est pas parvenue à l'OEB, n'a pas été établie conformément à la norme ST.23 de l'OMPI, n'a pas été produite sur le support de données prescrit (Annexe II de la décision du Président de l'OEB en date du 11 décembre 1992) ou n'est pas rédigée dans une langue officielle (cf. règle 1(1) CBE) de l'OEB, le demandeur est invité à remédier dans un **délai non prorogeable de deux mois** aux irrégularités constatées (règle 104ter (3bis) CBE). Le cas échéant, le demandeur est invité à déclarer en même temps que la liste de séquences déposées ultérieurement ne s'étend pas au-delà du contenu de la demande telle qu'elle a été déposée et/ou que l'information figurant sur le support de données est identique à celle que contient la liste écrite

2. Si, après une telle invitation, le demandeur ne remédie pas aux irrégularités constatées ou n'y remédie pas en temps utile, la demande est rejetée (art. 91(3) CBE).

Der Anmelder kann die Weiterbehandlung der Anmeldung gemäß Artikel 121 EPÜ beantragen.

#### V. Veröffentlichung von Sequenzprotokollen

1. Ist das Sequenzprotokoll **zusammen** mit den Unterlagen der europäischen oder internationalen Anmeldung als Teil der Beschreibung eingereicht worden, so wird es zusammen mit der Anmeldung und, im Fall des Artikels 158 (3) EPU, zusammen mit der nach Artikel 158 (2) EPÜ eingereichten Übersetzung der Anmeldung veröffentlicht.

2 Wird ein Sequenzprotokoll **nachgereicht**, so hängt seine Veröffentlichung von dem Zeitpunkt der Einreichung ab:

Wird es rechtzeitig vor Abschluß der technischen Vorbereitungen für die Veröffentlichung der Anmeldung<sup>8</sup> eingereicht, so wird es als Anlage zu den Anmeldeunterlagen (A-Schrift) oder im Fall einer Euro-PCT-Anmeldung, die nicht in einer Amtssprache des EPA international veröffentlicht worden ist, als Anlage zu der nach Artikel 158 (3) EPÜ veröffentlichten Übersetzung mitveröffentlicht.

In allen anderen Fällen besteht nur die Möglichkeit, bei Patenterteilung das Sequenzprotokoll zusammen mit der Patentschrift zu veröffentlichen. Die Veröffentlichung des nachgereichten Sequenzprotokolls zusammen mit dem gesonderten Recherchenbericht (sogenannte A3-Schrift) ist nicht vorgesehen, da dieser nicht druckschriftlich verbreitet wird

3 Ein Sequenzprotokoll, das nicht zusammen mit den Anmeldeunterlagen oder später veröffentlicht wird, ist nur durch Akteneinsicht, wenn die Voraussetzungen hierfür erfüllt sind, und gegebenenfalls durch Abfragen der Patent-Sequenzdatenbank der Öffentlichkeit zugänglich.

The applicant may request further processing of the application under Article 121 EPC.

#### V. Publication of sequence listings

1. If the sequence listing has been filed **together** with the documents making up the European or international application as part of the description, it is published together with the application and, where Article 158(3) EPC applies, together with the translation of the application filed under Article 158(2) EPC.

2. If a sequence listing is **subsequently filed**, its publication depends on the time of filing:

If it is filed in due time before completion of the technical preparations for publication of the application<sup>8</sup>, it is published as an annex to the application documents (A document) or, in the case of a Euro-PCT application which has not been published internationally in one of the official languages of the EPO, as an annex to the translation published under Article 158(3) EPC.

In all other cases it is only possible to publish the sequence listing together with the patent specification when the patent is granted. There is no provision for publishing the subsequently filed sequence listing together with the separate search report ("A3 document"), as the latter is not disseminated as a printed document.

3. A sequence listing which is not published together with the application documents or later is only accessible to the public by way of file inspection, if the conditions therefor are met, and, where applicable, by interrogating the patent sequence database.

Le demandeur peut demander la poursuite de la procédure conformément à l'article 121 CBE.

#### V. Publication des listes de séquences

1. Si la liste de séquences a été déposée **en même temps** que les pièces de la demande européenne ou internationale comme partie de la description, elle est publiée avec la demande ou, dans le cas visé à l'article 158(3) CBE, avec la traduction de la demande remise conformément à l'article 158(2) CBE.

2 Si la liste de séquences a été **déposée ultérieurement**, sa publication dépend de la date de son dépôt:

Si elle est déposée en temps utile avant la fin des préparatifs techniques entrepris en vue de la publication de la demande<sup>8</sup>, elle est publiée en annexe aux pièces de la demande (document A) ou, s'il s'agit d'une demande euro-PCT qui n'a pas fait l'objet d'une publication internationale dans une langue officielle de l'OEB, en annexe à la traduction publiée conformément à l'article 158(3) CBE

Dans tous les autres cas, la seule possibilité consiste, si la demande donne lieu à la délivrance d'un brevet, à publier la liste de séquences avec le fascicule de brevet. Il n'est pas prévu qu'une liste de séquences déposée ultérieurement soit publiée avec le rapport de recherche séparé ("document A3"), étant donné que ce rapport n'est pas diffusé sous forme imprimée.

3. Une liste de séquences qui n'est publiée ni avec les pièces de la demande ni ultérieurement n'est accessible au public que dans le cadre de l'inspection publique, lorsque les conditions requises pour cela sont réunies, et, le cas échéant, par consultation de la banque de données relatives aux séquences contenues dans les brevets.

<sup>8</sup> Ab 1 Januar 1993 7 Wochen vor Ablauf von 18 Monaten nach dem Anmelde- oder Prioritätstag; Beschluß des Präsidenten des EPA vom 14. Dezember 1992, der im ABI EPA 1-2/1993 veröffentlicht wird

<sup>8</sup> As from 1 January 1993 7 weeks before the expiry of 18 months from the date of filing or priority; Decision of the President of the EPO dated 14 December 1992, to be published in OJ EPO 1-2/1993.

<sup>8</sup> A compter du 1<sup>er</sup> janvier 1993: 7 semaines avant l'expiration d'un délai de 18 mois suivant la date de dépôt ou de priorité, décision du Président de l'OEB en date du 14 décembre 1992, qui sera publiée au JO OEB 1-2/1993



## ANHANG I zum Beschluß des Präsidenten des EPA vom 11. Dezember 1992

### WIPO-STANDARD ST. 23

**Empfehlung für die Darstellung von Nucleotid- und Aminosäuresequenzen in Patentanmeldungen und veröffentlichten Patentedokumenten**

#### EINFÜHRUNG

1 Mit der vorliegenden Empfehlung soll auf eine standardisierte Darstellung von Nucleotid- und Aminosäuresequenzen in Patentanmeldungen und veröffentlichten Patentedokumenten hingewirkt werden. Ziel der Empfehlung ist es, die Kosten für die Konvertierung der zur Veröffentlichung bestimmten Anmeldungen in ein standardisiertes digitales Format zu senken, eine genauere und qualitativ bessere Wiedergabe von Nucleotid- und Aminosäuresequenzen in Patentanmeldungen und Patentedokumenten zu erreichen, den Anmeldern, der Öffentlichkeit und den Prüfern die Auswertung von Sequenzen zu erleichtern und den Einsatz computergestützter Datenbanken auf dem Gebiet der Biotechnologie zu fördern.

2 Der Empfehlung liegen entsprechende Bestimmungen zugrunde, die einige der Behörden für den gewerblichen Rechtsschutz bereits eingeführt haben. Die Empfehlung enthält keine spezifische Regelung zur Wiedergabe von Nucleotid- und Aminosäuresequenzen auf vom Anmelder selbst erstellten elektronischen Datenträgern. Anmeldern, die keine elektronischen Datenträger zur Offenbarung dieser Sequenzen einreichen, wird empfohlen, sich bei der Einreichung schriftlicher Unterlagen eng an den WIPO-Standard ST. 22 (Empfehlung für die Wiedergabe von Patentanmeldungen in einem für die optische Zeichenerkennung (OCR) geeigneten Format) zu halten, damit Nucleotid- und Aminosäuresequenzen richtig übertragen werden können.

#### DEFINITIONEN

3 Im Rahmen der vorliegenden Empfehlung gelten folgende Definitionen:

i) Unter einer "Patentanmeldung" ist ein von einem Anmelder oder einem Vertreter in seinem Auftrag eingereichtes Dokument zu verstehen, mit dem um ein Patent nachgesucht wird. Es enthält in der Regel eine detaillierte Beschreibung der Erfindung, die Ansprüche und gegebenenfalls Zeichnungen, falls diese für das Verständnis der Erfindung erforderlich sind.

ii) Unter einem "Patentedokument" sind Erfindungspatente, Erfinderscheine, Gebrauchszertifikate, Gebrauchsmuster, Zusatzpatente oder -zertifikate,

## ANNEX I to the Decision of the President of the EPO dated 11 December 1992

### WIPO STANDARD ST.23

**Recommendation for the presentation of nucleotide and amino acid sequences in patent applications and in published patent documents**

#### INTRODUCTION

1 This Recommendation has been elaborated so as to provide standardization of the presentation of nucleotide and amino acid sequences in patent applications and in published patent documents. The Recommendation is intended to reduce the cost of converting applications into a standardized digitized form for publication, to enhance the accuracy and quality of presentations of nucleotide and amino acid sequences given in patent applications and in patent documents, to make for easier interpretation of sequences by applicants, the public and examiners and to improve the use of computerized data bases in the field of biotechnology.

2 This Recommendation has been based upon corresponding regulations already introduced by some industrial property offices. The Recommendation makes no specific provision for the presentation by applicants of nucleotide and amino acid sequences on electronic data carriers. Failing the filing of disclosures of those sequences on electronic data carriers applicants are recommended to submit typed documents that closely follow WIPO Standard ST.22 (Recommendation for the presentation of patent applications typed in Optical Character Recognition (OCR) format) so as to permit the accurate transcription of nucleotide and amino acid sequences.

#### DEFINITIONS

3 For the purposes of this Recommendation:

(i) the expression "patent application" means a document filed by an applicant, or by an agent (representative) on his (her) behalf, requesting the grant of a patent. It usually contains a detailed description of the invention, the claims, and drawings, when necessary for the understanding of the invention;

(ii) the expression "patent document" means patents for invention, inventors' certificates, utility certificates, utility models, patents or certificates of addi-

## ANNEXE I à la Décision du Président de l'OEB du 11 décembre 1992

### NORME OMPI ST.23

**Recommandation relative à la présentation des séquences de nucléotides et d'acides aminés dans les demandes de brevet et les documents de brevet publiés**

#### INTRODUCTION

1. La présente recommandation a été élaborée en vue de normaliser la présentation des séquences de nucléotides et d'acides aminés dans les demandes de brevet et les documents de brevet publiés. Elle vise à réduire le coût de transformation des demandes en demandes numérisées et normalisées aux fins de publication, à accroître la précision et la qualité des présentations de séquences de nucléotides et d'acides aminés dans les demandes de brevet et les documents de brevet, à faciliter l'interprétation des séquences par les déposants, le public et les examinateurs et à améliorer l'utilisation des bases de données informatisées dans le domaine de la biotechnologie.

2. Cette recommandation tient compte des réglementations correspondantes déjà introduites par quelques offices de propriété industrielle. Elle ne comporte aucune disposition particulière concernant la présentation, par les déposants, des séquences de nucléotides et d'acides aminés sur des supports électroniques de données. Si les divulgations de ces séquences ne sont pas déposées sur un support électronique de données, il est recommandé aux déposants de remettre des documents dactylographiés conformes aux dispositions de la norme ST.22 de l'OMPI (Recommandation relative à la présentation de demandes de brevet dactylographiées sous une forme permettant une reconnaissance optique des caractères (ROC)) afin de permettre la transcription exacte des séquences de nucléotides et d'acides aminés.

#### DEFINITIONS

3. Aux fins de la présente recommandation:

i) l'expression "demande de brevet" désigne un document dans lequel le déposant — ou, pour le compte de celui-ci, un mandataire (représentant) — demande la délivrance d'un brevet. Ce document contient habituellement une description détaillée de l'invention, les revendications, et des dessins lorsque ceux-ci sont nécessaires à la compréhension de l'invention;

ii) l'expression "document de brevet" recouvre les brevets d'invention, les certificats d'auteur d'invention, les certificats d'utilité, les modèles d'utilité,

Zusatzerfinderscheine, Zusatzgebrauchszertifikate und veröffentlichte Anmeldungen für solche Schutzrechte zu verstehen.

iii) Unter den Begriffen "Nucleotidsequenz" und "Aminosäuresequenz" sind unverzweigte Sequenzen von mindestens 10 benachbarten Nucleotiden bzw. unverzweigte Sequenzen von mindestens 4 benachbarten Aminosäuren zu verstehen. Verzweigte Sequenzen sind aus dieser Definition ausdrücklich ausgeschlossen. Für Nucleotide und Aminosäuren gilt folgende weitere Definition:

a) Unter "Nucleotiden" sind nur die Nucleotide zu verstehen, die mittels der unter Nummer 8 aufgeführten Symbole wiedergegeben werden können. Modifikationen wie z. B. methylierte Basen können nach der Anleitung unter Nummer 9 beschrieben werden, sind aber in der Nucleotidsequenz nicht explizit auszuweisen.

b) Unter "Aminosäuren" sind die unter Nummer 11 aufgeführten gängigen L-Aminosäuren aus den in der Natur vorkommenden Proteinen zu verstehen. Aminosäuresequenzen, die D-Aminosäuren enthalten, fallen nicht unter diese Definition. Aminosäuresequenzen, die posttranslational modifizierte Aminosäuren enthalten, können mittels der unter Nummer 11 aufgeführten Symbole wie die zunächst translatierte Aminosäure beschrieben werden, wobei die modifizierten Positionen wie z. B. Hydroxylierungen oder Glycosilierungen nach der Anleitung unter Nummer 12 beschrieben werden, diese Modifikationen aber in der Aminosäuresequenz nicht explizit auszuweisen sind. Unter die vorstehende Definition fallen auch Peptide oder Proteine, die anhand der unter Nummer 11 aufgeführten Symbole sowie einer an anderer Stelle aufgenommenen Beschreibung, die beispielsweise Aufschluß über ungewöhnliche Bindungen, Quervernetzungen und "end caps", Nicht-peptidbindungen usw. gibt, als Sequenz wiedergegeben werden können.

#### SEQUENZPROTOKOLL

4 Die Sequenzen sind in der Patentanmeldung oder dem veröffentlichten Patentdokument als letzter, gesonderter Teil der Beschreibung anzugeben, der die Überschrift "Sequenzprotokoll" trägt und auf einer neuen Seite beginnt. Jede offenbare Sequenz ist im "Sequenzprotokoll" gesondert auszuweisen. Jeder Sequenz ist eine eigene Sequenzkennzahl in der Schreibweise SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 usw. zuzuweisen. Es wird empfohlen, die Anzahl der im Sequenzprotokoll wiedergegebenen Sequenzen anzugeben.

5 In der Beschreibung oder den Ansprüchen der Anmeldung ist auf die im Sequenzprotokoll dargestellten Se-

tion, inventors' certificates of addition, utility certificates of addition, and published applications therefor;

(iii) the expressions "nucleotide sequence" and "amino acid sequence" mean an unbranched sequence of ten or more contiguous nucleotides and an unbranched sequence of four or more contiguous amino acids, respectively. Branched sequences are specifically excluded from this definition. Nucleotides and amino acids are further defined as follows:

(a) "Nucleotides" embrace only those nucleotides that can be represented using the symbols set forth in paragraph 8. Modifications, e.g. methylated bases, may be described as set forth in paragraph 9, but shall not be shown explicitly in the nucleotide sequence.

(b) "Amino acids" are those L-amino acids commonly found in naturally occurring proteins and are listed in paragraph 11. Those amino acid sequences containing D-amino acids are not intended to be embraced by this definition. Any amino acid sequence that contains post-translationally modified amino acids may be described as the amino acid sequence that is initially translated using the symbols shown in paragraph 11 with the modified positions, e.g. hydroxylations or glycosylations, being described as set forth in paragraph 12, but these modifications shall not be shown explicitly in the amino acid sequence. Any peptide or protein that can be expressed as a sequence using the symbols in paragraph 11 in conjunction with a description elsewhere to describe, for example, abnormal linkages, cross links and end caps, non-peptidyl bonds, etc., is embraced by this definition.

#### SEQUENCE LISTING

4 The sequences should be given as a last, separate part of the description in the patent application or published patent document, entitled "Sequence Listing", and should begin on a new page. Each sequence disclosed should appear separately in the "Sequence Listing". Each sequence should be assigned a separate identifier number, written as SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, etc. It is recommended that the number of sequences presented in the Sequence Listing be indicated.

5 In the description or claims of the application, the sequences represented in the Sequence Listing should

les brevets ou les certificats d'addition, les certificats d'auteur d'invention additionnels, les certificats d'utilité additionnels et les demandes publiées visant tous ces titres;

iii) les expressions "séquence de nucléotides" et "séquence d'acides aminés" désignent respectivement une séquence linéaire d'au moins dix nucléotides contigus et une séquence linéaire d'au moins quatre acides aminés contigus. Les séquences non linéaires sont expressément exclues de cette définition. De plus, on entend par:

a) "nucléotides", uniquement les nucléotides qui peuvent être représentés à l'aide des symboles indiqués au paragraphe 8. Les modifications (p.ex. les bases méthylées) peuvent être décrites de la manière indiquée au paragraphe 9, mais elles ne doivent pas être représentées de façon explicite dans la séquence de nucléotides.

b) "acides aminés", les acides aminés L que l'on rencontre généralement dans des protéines naturelles et qui sont énumérés au paragraphe 11. N'entrent pas dans cette définition les séquences d'acides aminés qui contiennent des acides aminés D. Toute séquence d'acides aminés qui contient des acides aminés modifiés après traduction peut être représentée sous la forme de la séquence initialement traduite à l'aide des symboles indiqués au paragraphe 11, les positions modifiées (p.ex. hydroxylations ou glycosylations) étant elles-mêmes décrites de la manière indiquée au paragraphe 12; ces modifications ne doivent toutefois pas être représentées de façon explicite dans la séquence d'acides aminés. Entre dans cette définition tout peptide ou toute protéine qui peut être exprimé(e) sous forme de séquence à l'aide des symboles énumérés au paragraphe 11 et accompagné(e), par exemple, d'une description des liaisons anormales, des liaisons croisées et des coiffes terminales, des liaisons non peptidiques, etc.

#### LISTE DE SEQUENCES

4. Les séquences doivent figurer dans une partie distincte à la fin de la description contenue dans la demande de brevet ou dans le document de brevet publié, intitulée "Liste des séquences". La liste doit commencer sur une nouvelle page. Chaque séquence divulguée doit être indiquée séparément dans cette liste et porter un numéro d'identification tel que SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, etc. Il est recommandé que le nombre des séquences présentées dans la liste soit indiqué.

5 Dans la description ou les revendications de la demande, toute séquence de la liste à laquelle il est fait

quenzen anhand der Kennzahl zu verweisen, auch wenn die Sequenz selbst oder weitere oder abgewandelte Darstellungen der Sequenz im Text oder in den zur Beschreibung gehörigen Zeichnungen enthalten sind.

6 Für die Darstellung von Nucleotid- und Aminosäuresequenzen ist zumindest eine der folgenden drei Möglichkeiten zu wählen:

- i) bloße Nucleotidsequenz
- ii) Aminosäuresequenz
- iii) Nucleotidsequenz zusammen mit der entsprechenden Aminosäuresequenz

#### ZU VERWENDEnde SYMBOLE

7 Nucleotidsequenzen sind nur anhand eines Einzelstrangs in Richtung vom 5'-Ende zum 3'-Ende von links nach rechts wiederzugeben.

8 Die Basen einer Nucleotidsequenz sind anhand des einbuchstabigen Codes für Nucleotidsequenzzeichen darzustellen. Es wird empfohlen, die in der nachstehenden Liste aufgeführten Großbuchstaben zu verwenden.

Symbol	Bedeutung	Ableitung der Bezeichnung
A	A	Adenin
G	G	Guanin
C	C	Cytosin
T	T	Thymin
U	U	Uridin*
R	G oder A	PuRin
Y	T/U oder C	pYrimidin
M	A oder C	aMino
K	G oder T/U	Keto
S	G oder C	Starke Bindungen 3 H-Brücken
W	A oder T/U	schwache (E: Weak) Bindungen 2 H-Brücken
B	G oder C oder T/U	nicht A
D	A oder G oder T/U	nicht C
H	A oder C oder T/U	nicht G
V	A oder G oder C	nicht T, nicht U
N	(A oder G oder C oder T/U) oder (unbekannt oder sonstige)	beliebig (E: aNy)

9 Modifizierte Basen können in der Sequenz selbst wie die entsprechenden unmodifizierten Basen wiedergegeben werden, wenn es sich um eine der nachstehend aufgeführten modifizierten Basen handelt und die Modifikation an anderer Stelle des Sequenzprotokolls näher beschrieben wird. Die

be referred to by the identifier number, even if the sequence or other additional or modified representations of the sequence are embedded in the text or in the drawings accompanying the description.

6. Nucleotide and amino acid sequences should be represented by at least one of the following three possibilities:

- (i) a pure nucleotide sequence;
- (ii) an amino acid sequence;
- (iii) a nucleotide sequence together with its corresponding amino acid sequence

#### SYMBOLS TO BE USED

7. A nucleotide sequence should be presented only by a single strand, in the 5'-end to 3'-end direction from left to right.

8 The bases of a nucleotide sequence should be represented using the one-letter code for nucleotide sequence characters. It is recommended to use upper case letters in conformity with the following list:

Symbol	Meaning	Origin of designation
A	A	Adenine
G	G	Guanine
C	C	Cytosine
T	T	Thymine
U	U	Uridine*
R	G or A	puRine
Y	T/U or C	pYrimidine
M	A or C	aMino
K	G or T/U	Keto
S	G or C	Strong interactions 3H-bonds
W	A or T/U	Weak interactions 2H-bonds
B	G or C or T/U	not A
D	A or G or T/U	not C
H	A or C or T/U	not G
V	A or G or C	not T, not U
N	(A or G or C or T/U) or (unknown or other)	aNy

9. Modified bases may be represented as the corresponding unmodified bases in the sequence itself if the modified base is one of those listed below and the modification is further described elsewhere in the Sequence Listing. The codes from the list below may be used in the description or the

référence doit être désignée par son numéro d'identification, même si cette séquence ou des représentations supplémentaires ou modifiées de celle-ci sont insérées dans le texte ou dans les dessins accompagnant la description.

6. Les séquences de nucléotides et d'acides aminés doivent être représentées au moins sous l'une des trois formes suivantes:

- i) une séquence de nucléotides pure;
- ii) une séquence d'acides aminés;
- iii) une séquence de nucléotides ainsi que sa séquence d'acides aminés correspondante.

#### SYMBOLES A UTILISER

7. Toute séquence de nucléotides doit être représentée par un seul brin de codage, dans le sens 5'-3' et de gauche à droite.

8. Les bases d'une séquence de nucléotides doivent être représentées au moyen du code à une lettre utilisé pour les séquences de nucléotides. Il est recommandé d'employer des majuscules conformément à la liste ci-après;

Symbole	Signification	Origine de la désignation
A	A	Adénine
G	G	Guanine
C	C	Cytosine
T	T	Thymine
U	U	Uridine*
R	G ou A	puRine
Y	T/U ou C	pYrimidine
M	A ou C	aMino
K	G ou T/U	K eto (Ceto)
S	G ou C	Interactions fortes (3 liaisons d'hydrogène)
W	A ou T/U	Interactions faibles (2 liaisons d'hydrogène)
B	G ou C ou T/U	autre que A
D	A ou G ou T/U	autre que C
H	A ou C ou T/U	autre que G
V	A ou G ou C	autre que T et U
N	(A ou G ou C ou T/U) ou (non connu ou autre)	(n'importe lequel)

9. Les bases modifiées peuvent être représentées par les bases non modifiées correspondantes contenues dans la séquence elle-même si la base modifiée figure parmi celles qui sont énumérées ci-après et si la modification fait l'objet d'une description plus détaillée dans la liste de séquences. Les

\* Dem zuständigen Ausschuss der WIPO liegt ein Antrag vor, "Uridin" durch "Uracil" zu ersetzen

\* The competent WIPO committee has been asked to substitute "Uridine" by "Uracil"

\* Le comité compétent de l'OMPI a été saisi d'une demande de remplacement du terme "Uridine" par "Uracile".

Codes aus der nachstehenden Liste dürfen in der Beschreibung oder im Sequenzprotokoll, nicht aber in der Sequenz selbst verwendet werden.

Sequence Listing but not in the sequence itself.

codes de la liste qui suit peuvent être utilisés dans la description ou la liste de séquences mais non dans la séquence proprement dite:

<i>Symbol</i>	<i>Bedeutung</i>	<i>Symbol</i>	<i>Meaning</i>	<i>Symbole</i>	<i>Signification</i>
ac4c	4-Acetylcytidin	ac4c	4-acetylcytidine	ac4c	4-acétylcytidine
chm5u	5-(Carboxyhydroxymethyl)uridin	chm5u	5-(carboxyhydroxymethyl)uridine	chm5u	5-(carboxyhydroxyméthyl)uridine
cm	2'-O-Methylcytidin	cm	2'-O-methylcytidine	cm	2'-O-méthylcytidine
cmnm5s2u	5-Carboxymethylaminomethyl-2-thiouridin	cmnm5s2u	5-carboxymethylaminomethyl-2-thiouridine	cmnm5s2u	5-carboxyméthylaminométhyl-2-thio-uridine
cmnm5u	5-Carboxymethylaminomethyluridin	cmnm5u	5-carboxymethylaminomethyluridine	cmnm5u	5-carboxyméthylaminométhyluridine
d	Dihydrouridin	d	dihydrouridine	d	dihydro-uridine
fm	2'-O-Methylpseudo-uridin	fm	2'-O-methylpseudo-uridine	fm	2'-O-méthylpseudo-uridine
gal q	beta, D-Galactosyl-queuosin	gal q	beta, D-galactosylqueuosine	gal q	bêta, D-galactosyl-queuosine
gm	2'-O-Methylguanosin	gm	2'-O-methylguanosine	gm	2'-O-méthylguanosine
i	Inosin	i	inosine	i	inosine
i6a	N6-Isopentenyladenosin	i6a	N6-isopentenyladenosine	i6a	N6-isopentényladénosine
m1a	1-Methyladenosin	m1a	1-methyladenosine	m1a	1-méthyladénosine
m1f	1-Methylpseudouridin	m1f	1-methylpseudouridine	m1f	1-méthylpseudo-uridine
m1g	1-Methylguanosin	m1g	1-methylguanosine	m1g	1-méthylguanosine
m1i	1-Methylinosin	m1i	1-methylinosine	m1i	1-méthylinosine
m22g	2,2-Dimethylguanosin	m22g	2,2-dimethylguanosine	m22g	2,2-diméthylguanosine
m2a	2-Methyladenosin	m2a	2-methyladenosine	m2a	2-méthyladénosine
m2g	2-Methylguanosin	m2g	2-methylguanosine	m2g	2-méthylguanosine
m3c	3-Methylcytidin	m3c	3-methylcytidine	m3c	3-méthylcytidine
m5c	5-Methylcytidin	m5c	5-methylcytidine	m5c	5-méthylcytidine
m6a	N6-Methyladenosin	m6a	N6-methyladenosine	m6a	N6-méthyladénosine
m7g	7-Methylguanosin	m7g	7-methylguanosine	m7g	7-méthylguanosine
mam5u	5-Methylaminomethyl-uridin	mam5u	5-methylaminomethyl-uridine	mam5u	5-méthylaminométhyl-uridine
mam5s2u	5-Methoxyaminomethyl-2-thiouridin	mam5s2u	5-methoxyaminomethyl-2-thiouridine	mam5s2u	5-méthoxyaminométhyl-2-thio-uridine
man q	beta, D-Mannosequeuosin	man q	beta, D-mannosylqueuosine	man q	bêta, D-mannosyl-queuosine
mcm5s2u	5-Methoxycarbonylmethyl-2-thiouridin	mcm5s2u	5-methoxycarbonylmethyl-2-thiouridine	mcm5s2u	5-méthoxycarbonylméthyl-2-thio-uridine
mcm5u	5-Methoxycarbonylmethyluridin	mcm5u	5-methoxycarbonylmethyluridine	mcm5u	5-méthoxycarbonylméthyluridine
mo5u	5-Methoxyuridin	mo5u	5-methoxyuridine	mo5u	5-méthoxyuridine
ms2i6a	2-Methylthio-N6-isopentenyladenosin	ms2i6a	2-methylthio-N6-isopentenyladenosine	ms2i6a	2-méthylthio-N6-isopentényladénosine
ms2t6a	N-((9-beta-D-Ribofuranosyl-2-methylthiopurin-6-yl) carbamoyl)threonin	ms2t6a	N-((9-beta-D-ribofuranosyl-2-methylthiopurine-6-yl) carbamoyl)threonine	ms2t6a	N-((9-bêta-D-ribofuranosyl-2-méthylthiopurine-6-yl) carbamoyl)thréonine
mt6a	N-((9-beta-D-Ribofuranosylpurin-6-yl)N-methylcarbamoyl)threonin	mt6a	N-((9-beta-D-ribofuranosylpurine-6-yl)N-methylcarbamoyl)threonine	mt6a	N-((9-bêta-D-ribofuranosylpurine-6-yl)N-méthylcarbamoyl)thréonine
mv	Uridin-5-oxyessigsäuremethylester	mv	uridine-5-oxyacetic acid-methylester	mv	ester méthylé d'uridine 5 oxyacétique acide
o5u	Uridin-5-oxyessigsäure (v)	o5u	uridine-5-oxyacetic acid (v)	o5u	acide d'uridine 5 oxyacétique
osyw	Wybutoxosin	osyw	wybutoxosine	osyw	wybutoxosine
p	Pseudouridin	p	pseudouridine	p	pseudo-uridine
q	Queuosin	q	queuosine	q	queuosine
s2c	2-Thiocytidin	s2c	2-thiocytidine	s2c	2-thiocytidine
s2t	5-Methyl-2-thiouridin	s2t	5-methyl-2-thiouridine	s2t	5-méthyl-2-thio-uridine
s2u	2-Thiouridin	s2u	2-thiouridine	s2u	2-thio-uridine
s4u	4-Thiouridin	s4u	4-thiouridine	s4u	4-thio-uridine
t	5-Methyluridin	t	5-methyluridine	t	5-méthyluridine
t6a	N-((9-beta-D-Ribofuranosylpurin-6-yl) carbamoyl)threonin	t6a	N-((9-beta-D-ribofuranosylpurine-6-yl) carbamoyl)threonine	t6a	N-((9-bêta-D-ribofuranosylpurine-6-yl) carbamoyl)thréonine
tm	2'-O-Methyl-5-methyluridin	tm	2'-O-methyl-5-methyluridine	tm	2'-O-méthyl-5-méthyl-uridine
um	2'-O-Methyluridin	um	2'-O-methyluridine	um	2'-O-méthyluridine
yw	Wybutosin	yw	wybutosine	yw	wybutosine
x	3-(3-Amino-3-carboxypropyl)uridin, (acp3)u	x	3-(3-amino-3-carboxypropyl)uridine, (acp3)u	x	3-(3-amino-3-carboxypropyl)uridine, (acp3)u

Ansonsten sind modifizierte Basen in der Sequenz mit "N" aufzuführen und an anderer Stelle des Sequenzprotokolls nähere Angaben zu machen

Otherwise, modified bases should be listed in the sequence as "N", with further information given elsewhere in the Sequence Listing

Dans le cas contraire, les bases modifiées doivent être désignées par "N" dans la séquence et assorties de renseignements complémentaires dans la liste de séquences.

10 Die Aminosäuren einer Protein- oder Peptidsequenz sind in Richtung von der Amino- zur Carboxylgruppe von links nach rechts aufzuführen, wobei die Amino- und Carboxylgruppen in der Sequenz nicht darzustellen sind.

10. The amino acids in a protein or peptide sequence should be listed in the amino to carboxy direction from left to right, and the amino and carboxy groups should not be represented in the sequence.

10. Les acides aminés d'une séquence de protéines ou de peptides doivent être énumérés dans le sens amino-carboxy et de gauche à droite, mais les groupes amino et carboxy ne doivent pas être représentés dans la séquence

11 Die Aminosäuren sind anhand des dreibuchstabigen Codes mit großem Anfangsbuchstaben entsprechend der nachstehenden Liste darzustellen:

11 The amino acids should be represented using the three-letter code with the first letter as a capital and should conform to the following:

11 Les acides aminés doivent être représentés au moyen du code à trois lettres (la première lettre étant en majuscule), conformément à la liste suivante:

<i>Symbol</i>	<i>Bedeutung</i>	<i>Symbol</i>	<i>Meaning</i>	<i>Symbole</i>	<i>Signification</i>
Ala	Alanin	Ala	Alanine	Ala	Alanine
Cys	Cystein	Cys	Cysteine	Cys	Cystéine
Asp	Asparaginsäure	Asp	Aspartic Acid	Asp	Acide aspartique
Glu	Glutaminsäure	Glu	Glutamic Acid	Glu	Acide glutamique
Phe	Phenylalanin	Phe	Phenylalanine	Phe	Phénylalanine
Gly	Glycin	Gly	Glycine	Gly	Glycine
His	Histidin	His	Histidine	His	Histidine
Ile	Isoleucin	Ile	Isoleucine	Ile	Isoleucine
Lys	Lysin	Lys	Lysine	Lys	Lysine
Leu	Leucin	Leu	Leucine	Leu	Leucine
Met	Methionin	Met	Methionine	Met	Méthionine
Asn	Asparagin	Asn	Asparagine	Asn	Asparagine
Pro	Prolin	Pro	Proline	Pro	Proline
Gln	Glutamin	Gln	Glutamine	Gln	Glutamine
Arg	Arginin	Arg	Arginine	Arg	Arginine
Ser	Serin	Ser	Serine	Ser	Sérine
Thr	Threonin	Thr	Threonine	Thr	Thréonine
Val	Valin	Val	Valine	Val	Valine
Trp	Tryptophan	Trp	Tryptophan	Trp	Tryptophane
Tyr	Tyrosin	Tyr	Tyrosine	Tyr	Tyrosine
Asx	Asp oder Asn	Asx	Asp or Asn	Asx	Asp ou Asn
Glx	Glu oder Gln	Glx	Glu or Gln	Glx	Glu ou Gln
Xaa	unbekannt oder sonstige	Xaa	unknown or other	Xaa	indéterminé ou autre

12 Modifizierte und seltene Aminosäuren können in der Sequenz selbst wie die entsprechenden unmodifizierten Aminosäuren wiedergegeben werden, wenn es sich um eine der nachstehend aufgeführten modifizierten Aminosäuren handelt und die Modifikation an anderer Stelle des Sequenzprotokolls näher beschrieben wird. Die Codes aus der nachstehenden Liste dürfen in der Beschreibung oder im Sequenzprotokoll, nicht jedoch in der Sequenz selbst verwendet werden

12 Modified and unusual amino acids may be represented as the corresponding unmodified amino acids in the sequence itself if the modified amino acid is one of those listed below and the modification is also further described elsewhere in the Sequence Listing. The codes from the list below may be used in the description or the Sequence Listing but not in the sequence itself

12. Les acides aminés modifiés et peu connus peuvent être représentés par les acides aminés non modifiés correspondants dans la séquence elle-même si l'acide aminé modifié figure parmi ceux qui sont énumérés ci-après et si la modification fait également l'objet d'une description plus détaillée dans la liste de séquences. Les codes de la liste qui suit peuvent être utilisés dans la description ou la liste de séquences mais non dans la séquence proprement dite:

<i>Symbol</i>	<i>Bedeutung</i>	<i>Symbol</i>	<i>Meaning</i>	<i>Symbole</i>	<i>Signification</i>
Aad	2-Amino adipinsäure	Aad	2-Amino adipic acid	Aad	acide 2-amino-adipique
bAad	3-Amino adipinsäure	bAad	3-Amino adipic acid	bAad	acide 3-amino-adipique
bAla	beta-Alanin, beta-Aminopropionsäure	bAla	beta-Alanine, beta-Aminopropionic acid	bAla	bêta-alanine, acide bêta-amino-propionique
Abu	2-Aminobuttersäure	Abu	2-Aminobutyric acid	Abu	acide 2-amino-butyrique
4Abu	4-Aminobuttersäure, Piperidinsäure	4Abu	4-Aminobutyric acid, piperidinic acid	4Abu	acide 4-amino-butyrique, acide pipéridinique
Acp	6-Aminocaprinsäure	Acp	6-Aminocaproic acid	Acp	acide 6-amino-caproïque
Ahe	2-Aminoheptansäure	Ahe	2-Aminoheptanoic acid	Ahe	acide 2-amino-heptanoïque
Aib	2-Aminoisobuttersäure	Aib	2-Aminoisobutyric acid	Aib	acide 2-amino-isobutyrique
bAib	3-Aminoisobuttersäure	bAib	3-Aminoisobutyric acid	bAib	acide 3-amino-isobutyrique
Apm	2-Aminopimelinsäure	Apm	2-Aminopimelic acid	Apm	acide 2-amino-pimélique

<i>Symbol</i>	<i>Bedeutung</i>	<i>Symbol</i>	<i>Meaning</i>	<i>Symbole</i>	<i>Signification</i>
Dbu	2,4-Diaminobuttersäure	Dbu	2,4 Diaminobutyric acid	Dbu	acide 2,4-diamino-butyrrique
Des	Desmosin	Des	Desmosine	Des	desmosine
Dpm	2,2'-Diaminopimelinsäure	Dpm	2,2'-Diaminopimelic acid	Dpm	acide 2,2'-diaminopimélique
Dpr	2,3-Diaminopropionsäure	Dpr	2,3-Diaminopropionic acid	Dpr	acide 2,3-diaminopropionique
EtGly	N-Ethylglycin	EtGly	N-Ethylglycine	EtGly	N-éthylglycine
EtAsn	N-Ethylasparagin	EtAsn	N-Ethylasparagine	EtAsn	N-éthylasparagine
Hyl	Hydroxylysin	Hyl	Hydroxylysine	Hyl	hydroxylysine
aHyl	allo-Hydroxylysin	aHyl	allo-Hydroxylysine	aHyl	allo-hydroxylysine
3Hyp	3-Hydroxyprolin	3Hyp	3-Hydroxyproline	3Hyp	3-hydroxyproline
4Hyp	4-Hydroxyprolin	4Hyp	4-Hydroxyproline	4Hyp	4-hydroxyproline
Ide	Isodesmosin	Ide	Isodesmosine	Ide	isodesmosine
alle	allo-Isoleucin	alle	allo-Isoleucine	alle	allo-isoleucine
MeGly	N-Methylglycin, Sarkosin	MeGly	N-Methylglycine, sarcosine	MeGly	N-méthylglycine, sarcosine
Melle	N-Methylisoleucin	Melle	N-Methylisoleucine	Melle	N-méthylisoleucine
MeLys	6-N-Methyllysin	MeLys	6-N-Methyllysine	MeLys	6-N-méthyllysine
MeVal	N-Methylvalin	MeVal	N-Methylvaline	MeVal	N-méthylvaline
Nva	Norvalin	Nva	Norvaline	Nva	norvaline
Nle	Norleucin	Nle	Norleucine	Nle	norleucine
Orn	Ornithin	Orn	Ornithine	Orn	ornithine

Ansonsten sind modifizierte oder seltene Aminosäuren in der Sequenz mit "Xaa" aufzuführen und an anderer Stelle des Sequenzprotokolls nähere Angaben zu machen

Otherwise, modified or unusual amino acids should be listed in the sequence as "Xaa", with further information given elsewhere in the Sequence Listing.

Dans le cas contraire, les acides aminés modifiés ou peu connus doivent être désignés par "Xaa" dans la séquence et assortis de renseignements complémentaires dans la liste de séquences

#### ZU VERWENDENDEN FORMAT

13 Bei Nucleotidsequenzen sind höchstens 16 Codonen oder 60 Basen pro Zeile— mit einem Leerraum zwischen den einzelnen Codonen oder Gruppen von jeweils 10 Basen—aufzuführen.

14 Bei Protein- oder Peptidsequenzen sind höchstens 16 Aminosäuren pro Zeile mit einem Leerraum zwischen den einzelnen Aminosäuren aufzuführen

15. Die Basen einer Nucleotidsequenz (einschließlich Intronen) sind jeweils in 10er-Gruppen aufzuführen; dies gilt nicht für die codierenden Teile der Sequenz. Bleiben am Ende nichtcodierende Teile einer Sequenz weniger als 10 Basen als Rest, so sind diese zu einer Gruppe zusammenzufassen und durch einen Leerraum von angrenzenden Gruppen zu trennen (s. Anlage 1, Beispiel 10, Nucleotide 356–360)

16 Die Basen der codierenden Teile einer Nucleotidsequenz sind als Triplets (Codonen) aufzuführen

17. Die den Codonen der codierenden Teile einer Nucleotidsequenz entsprechenden Aminosäuren sind unmittelbar unter die jeweiligen Codonen zu schreiben. Wird ein Codon durch ein Intron aufgespalten, so ist das Aminosäuresymbol unter den Teil des Codons zu schreiben, der zwei Nucleotide enthält (s. Anlage 1, Beispiel 11, Position 120)

18 Die Zählung der Nucleotidbasen beginnt bei der ersten Base der Sequenz mit 1. Von hier aus ist die gesamte Sequenz in 5'-3'- Richtung fort-

#### FORMAT TO BE USED

13. A nucleotide sequence should be listed with a maximum of 16 codons or 60 bases per line, with a space between each codon or group of 10 bases.

14 A protein or peptide sequence should be listed with a maximum of 16 amino acids per line, with a space provided between each amino acid.

15 The bases of a nucleotide sequence (including introns) should be listed in groups of 10 bases, except in the coding parts of the sequence. Left-over bases, fewer than 10 in number at the end of noncoding parts of a sequence, shall be grouped together and separated from adjacent groups by a space (see Appendix 1, Example 10, nucleotides 356-360).

16 The bases of the coding parts of a nucleotide sequence should be listed as triplets (codons).

17 Amino acids corresponding to the codons in the coding parts of a nucleotide sequence should be typed immediately under the corresponding codons. Where a codon is split by an intron, the amino acid symbol should be typed below the portion of the codon containing two nucleotides (see Appendix 1, Example 11, position 120).

18. The enumeration of the nucleotide bases should start at the first base of the sequence with number 1. It should be continuous through the whole

#### MODE DE PRESENTATION A SUIVRE

13 Une séquence de nucléotides doit comporter 16 codons ou 60 bases par ligne au maximum, avec un espace entre chaque codon ou groupe de 10 bases.

14. Une séquence de protéines ou de peptides doit comporter 16 acides aminés par ligne au maximum, avec un espace entre chaque acide aminé.

15. Les bases d'une séquence de nucléotides (y compris les introns) doivent figurer sur la liste par groupes de 10 bases, sauf celles qui se situent dans les régions codantes de la séquence. Les bases (moins de 10) qui restent à l'extrémité des régions non codantes d'une séquence doivent être regroupées et séparées des groupes voisins par un espace (voir, dans l'exemple 10 de l'appendice I, les nucléotides 356 à 360).

16. Les bases des régions codantes d'une séquence de nucléotides doivent figurer sur la liste sous forme de triplets (codons).

17. Les acides aminés correspondant aux codons dans les régions codantes d'une séquence de nucléotides doivent figurer immédiatement sous les codons correspondants. Lorsqu'un codon est coupé par un intron, le symbole d'acide aminé doit figurer sous la partie du codon contenant deux nucléotides (voir, dans l'exemple 11 de l'appendice I, la position 120).

18. L'énumération des bases nucléotidiques doit commencer par la première base de la séquence, qui portera le numéro 1. Elle doit être continue

laufend durchzuzählen. Am rechten Rand ist jeweils neben der Zeile mit den einbuchstabigen Codes für die Basen die Nummer der letzten Base dieser Zeile anzugeben. Die vorstehend beschriebene Zählweise für Nucleotidsequenzen gilt auch für Nucleotidsequenzen mit ringförmiger Konfiguration, wobei allerdings die Bestimmung der ersten Base der Nucleotidsequenz dem Anmelder überlassen bleibt.

19 Die Zählung der Aminosäuren beginnt bei der ersten Aminosäure des reifen Proteins mit 1. Die dem reifen Protein vorausgehenden Aminosäuren wie beispielsweise Prä-Sequenzen, Pro-Sequenzen, Prä-Pro-Sequenzen und Signalsequenzen sind, soweit vorhanden, mit negativem Vorzeichen zu numerieren, wobei die rückläufige Zählung mit der Aminosäure vor Nummer 1 beginnt. Null (0) wird nicht verwendet, wenn Aminosäuren zur Abgrenzung gegen das reife Protein mit negativem Vorzeichen numeriert werden. Ansonsten beginnt die Zählung der Aminosäuren bei der ersten Aminosäure am Aminoterminus mit 1. Die Nummern sind im Abstand von jeweils 5 Aminosäuren unter der Sequenz anzugeben. Die vorstehend beschriebene Zählweise für Aminosäuresequenzen gilt auch für Aminosäuresequenzen mit ringförmiger Konfiguration, wobei allerdings die Bestimmung der ersten Aminosäure der Aminosäuresequenz dem Anmelder überlassen bleibt (siehe Anlage 1, Beispiele 4 und 5)

20 Teilsequenzen – die aus einem oder mehreren nichtbenachbarten Segmenten einer größeren Sequenz oder aus Segmenten verschiedener Sequenzen bestehen – sind als gesonderte Sequenzen mit eigener Sequenzkennzahl zu numerieren. Sequenzen mit einer oder mehreren Lücken sind als mehrere gesonderte Sequenzen mit eigenen Sequenzkennzahlen zu numerieren, wobei die Zahl der gesonderten Sequenzen der Zahl der jeweils zusammenhängenden Sequenzdatenreihen entspricht (siehe Anlage 1, Beispiele 13 und 15)

#### SONSTIGE VERFÜGBARE ANGABEN IM SEQUENZPROTOKOLL

21 Es wird empfohlen, die Datenelemente im Sequenzprotokoll in der Reihenfolge und Form, in der sie hier aufgeführt sind, mit den dazugehörigen Rubriknummern und in Großbuchstaben abgefaßten Rubriktiteln ohne die in Klammern gesetzten Erläuterungen anzugeben. Beansprucht der auf einen Rubriktitel folgende Text mehr als eine Zeile, so empfiehlt sich eine Einrück-

sequence in the direction 5' to 3'. It should be marked in the right margin, next to the line containing the one-letter codes for the bases, and giving the number of the last base of that line. The enumeration method for nucleotide sequences set forth above remains applicable to nucleotide sequences that are circular in configuration, with the exception that the designation of the first base of the nucleotide sequence may be made at the option of the applicant.

19. The enumeration of amino acids should start at the first amino acid of the mature protein, with number 1. The amino acids preceding the mature protein, for example pre-sequences, pro-sequences, pre-pro-sequences and signal sequences, when present, should have negative numbers, counting backwards starting with the amino acid next to number 1. Zero (0) is not used when the numbering of amino acids uses negative numbers to distinguish the mature protein. Otherwise, the enumeration of amino acids should start at the first amino acid at the amino terminal as number 1. It should be marked under the sequence every 5 amino acids. The enumeration method for amino acid sequences set forth above remains applicable for amino acid sequences that are circular in configuration, with the exception that the designation of the first amino acid of the amino acid sequence may be made at the option of the applicant. (See Appendix 1, Examples 4 and 5).

20. A partial sequence – made up of one or more non-contiguous segments of a larger sequence or of segments from different sequences – should be numbered as a separate sequence, with a separate sequence identifier. A sequence with a gap or gaps should be numbered as a plurality of separate sequences with separate sequence identifiers, with the number of separate sequences being equal in number to the number of continuous strings of sequence data. (See Appendix 1. Examples 13 and 15).

#### OTHER AVAILABLE INFORMATION IN THE SEQUENCE LISTING

21. It is recommended that the order and presentation of the items of information in the Sequence Listing follow the order in which those items are listed herein with appropriately numbered and lettered headings, wherein the headings are designated in upper case characters, not including any parenthetical explanatory information. When more than one line is necessary

dans toute la séquence dans le sens 5'-3'. Elle doit figurer dans la marge de droite sur la ligne contenant les codes à une lettre correspondant aux bases et indiquer le numéro de la dernière base de cette ligne. La méthode présentée ci-dessus pour énumérer des séquences de nucléotides s'applique aussi aux séquences de nucléotides de configuration circulaire, à cette différence près que la désignation de la première base de la séquence de nucléotides peut être laissée au choix du déposant.

19. L'énumération des acides aminés doit commencer par le premier acide aminé de la protéine mature, qui portera le numéro 1. Les acides aminés précédant la protéine mature, par exemple les préséquences, les proséquences et les pré-proséquences ainsi que les séquences signal lorsqu'elles existent, doivent porter des nombres négatifs numérotés à rebours, en commençant par l'acide aminé voisin de l'acide portant le numéro 1. Le numéro 0 (zéro) n'est pas utilisé lorsque, dans la numérotation des acides aminés, des valeurs négatives servent à distinguer la protéine mature. Sinon, l'énumération des acides aminés doit commencer par le premier acide aminé au niveau du terminal amino et porter le numéro 1. Le nombre doit figurer dans la séquence tous les cinq acides. La méthode présentée ci-dessus pour énumérer les séquences d'acides aminés s'applique aussi aux séquences d'acides aminés de configuration circulaire à l'exception du fait que la désignation du premier acide aminé de la séquence d'acides aminés peut être faite au choix du déposant. (Voir les exemples 4 et 5 de l'appendice I).

20. Une séquence partielle, composée d'un segment ou de plusieurs segments non contigus d'une séquence plus grande ou de segments provenant de différentes séquences, doit être numérotée comme une séquence distincte et porter un numéro d'identification de séquence distinct. Une séquence comportant un ou des espaces doit être numérotée comme une série de séquences distinctes et porter des numéros d'identification distincts, le nombre de séquences distinctes étant égal au nombre de chaînes continues (voir les exemples 13 et 15 de l'appendice I).

#### AUTRES RENSEIGNEMENTS DEVANT FIGURER DANS LA LISTE DES SEQUENCES

21. Il est recommandé que ces éléments d'information soient présentés dans l'ordre où ils sont indiqués ci-après, sous des rubriques en majuscules (non compris les explications éventuelles données entre parenthèses) et adéquatement numérotées. Quand le texte suivant une rubrique occupe plus d'une ligne, il est recommandé de disposer les lignes supplé-

kung der zusätzlichen Zeile(n), um sie vom Rubrikittel am linken Rand abzuheben.

22 Das Sequenzprotokoll hat zusätzlich zu der eigentlichen Nucleotid- und/oder Aminosäuresequenz unmittelbar vor dieser die nachstehend aufgeführten Datenelemente zu enthalten, soweit diese zutreffen und dem Anmelder bekannt sind.

(1) ANGABEN<sup>1</sup> ZU SEQ ID NO:X:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE (Länge der Sequenz, ausgedrückt als Zahl der Basenpaare oder Aminosäurereste):

(B) ART (Art der Sequenz, d. h. Angabe, ob es sich um eine Nucleotid- oder Aminosäuresequenz handelt):

(C) STRANGFORM (bei Nucleinsäuren Zahl der Stränge des Moleküls des Herkunftsorganismus, d. h. Angabe, ob es sich um einen Einzelstrang, einen Doppelstrang oder beides handelt oder aber die Strangform dem Anmelder nicht bekannt ist):

(D) TOPOLOGIE (Angabe, ob das Molekül des Herkunftsorganismus ringförmig, linear oder beides oder aber die Topologie dem Anmelder nicht bekannt ist):

(ii) ART DES MOLEKÜLS (Art des in SEQ ID NO:X sequenzierten Moleküls):

(In das Sequenzprotokoll ist als Unter rubrik mindestens eine der folgenden Angaben aufzunehmen):

- Genom-RNA
- Genom-DNA
- m-RNA
- t-RNA
- r-RNA
- sn-RNA
- sc-RNA
- prä-RNA
- c-DNA zu Genom-RNA
- c-DNA zu m-RNA
- c-DNA zu t-RNA
- c-DNA zu r-RNA
- c-DNA zu sn-RNA
- c-DNA zu sc-RNA
- sonstige Nucleinsäure (angeben):
- Protein
- Peptid

(iii) HYPOTHETISCH: (ja/nein)

(iv) ANTISENSE: (ja/nein)

(v) ART DES FRAGMENTS (nur bei Proteinen und Peptiden ist in das Sequenzprotokoll mindestens eine der folgenden Angaben aufzunehmen):

for the text following a heading, indentation of the additional line(s) is recommended for use to distinguish it from the heading at the left margin.

22. The Sequence Listing should include, in addition to and immediately preceding the actual nucleotide and/or amino acid sequence, the following items of information, if applicable and when available to the applicant.

(1) INFORMATION FOR SEQ ID NO:X:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH (sequence length, expressed as number of base pairs or amino acid residues):

(B) TYPE (sequence type, i.e. whether nucleotide or amino acid):

(C) STRANDEDNESS (if nucleic acid, number of strands of source organism molecule, i.e., whether single stranded, double stranded, both, or unknown to applicant):

(D) TOPOLOGY (whether source organism molecule is circular, linear, both, or unknown to applicant):

(ii) MOLECULE TYPE (type of molecule sequenced in SEQ ID NO:X):

(at least one of the following should be included with subheadings in Sequence Listing):

- Genomic RNA
- Genomic DNA
- mRNA
- tRNA
- rRNA
- snRNA
- scRNA
- preRNA
- cDNA to genomic RNA
- cDNA to mRNA
- cDNA to tRNA
- cDNA to rRNA
- cDNA to snRNA
- cDNA to scRNA
- Other nucleic acid (specify):
- Protein
- Peptide

(iii) HYPOTHETICAL: (yes/no)

(iv) ANTI-SENSE: (yes/no)

(v) FRAGMENT TYPE (for proteins and peptides only, at least one of the following should be included in the Sequence Listing):

mentaires en retrait pour les distinguer de la rubrique dans la marge de gauche.

22. La liste de séquences doit contenir, en sus de la séquence de nucléotides et/ou d'acides aminés proprement dite et juste avant celle-ci, les éléments d'information indiqués ci-après, selon qu'ils s'appliquent en l'espèce et compte tenu des renseignements dont dispose le déposant

1) INFORMATIONS<sup>1</sup> CONCERNANT SEQ ID NO:X:

i) CARACTERISTIQUES DE LA SEQUENCE:

A) LONGUEUR (longueur de la séquence, exprimée en nombre de paires de bases ou de résidus d'acide aminé):

B) TYPE (type de séquence, c'est-à-dire séquence de nucléotides ou séquence d'acides aminés):

C) NOMBRE DE BRINS (s'il s'agit d'un acide nucléique, nombre de brins de la molécule de l'organisme d'origine: à simple brin, à double brin, des deux catégories ou non connu du déposant):

D) CONFIGURATION (molécule de l'organisme d'origine: circulaire, linéaire, des deux catégories ou non connue du déposant):

ii) TYPE DE MOLECULE (type de molécule séquencée dans SEQ ID NO:X):

(l'un au moins des types suivants doit figurer en sous-rubrique sur la liste de séquences):

- ARN génomique
- ADN génomique
- ARNm
- ARNt
- ARNr
- ARNsn
- ARNsc
- pré-ARN
- ADNc pour ARN génomique
- ADNc pour ARNm
- ADNc pour ARNt
- ADNc pour ARNr
- ADNc pour ARNsn
- ADNc pour ARNsc
- Autre acide nucléique (préciser):
- Protéine
- Peptide

iii) HYPOTHETIQUE: (oui/non)

iv) ANTI-SENS: (oui/non)

v) TYPE DE FRAGMENT (pour les protéines et les peptides uniquement, l'un au moins des types de fragment suivants doit figurer sur la liste de séquences):

<sup>1</sup> Das Programm Patentin enthält gegenwärtig einige Begriffe, die von der Terminologie der vorliegenden Dokumente abweichen. So wird beispielsweise statt "DNA" das Kürzel "DNS" verwendet. Eine aktualisierte Version des Programms wird in Kürze zur Verfügung stehen.

<sup>1</sup> Le programme Patentin contient actuellement quelques termes qui varient de la terminologie utilisée dans les présents documents. Par exemple "DNA" y figure pour "ADN". Une version mise à jour sera disponible prochainement.



- N-Terminus
- C-Terminus
- inneres Fragment

(vi) URSPRÜNGLICHE HERKUNFT (ursprüngliche Herkunft des in SEQ ID NO:X sequenzierten Moleküls):

(A) ORGANISMUS (wissenschaftlicher Name des Herkunftsorganismus):

(B) STAMM:

(C) INDIVIDUUM/ISOLAT (Name/Nummer des Individuums/Isolats):

(D) ENTWICKLUNGSSTADIUM (Entwicklungsstadium des Herkunftsorganismus und Angabe, ob das sequenzierte Molekül von einer Keimlinie oder einem umgeordneten Entwicklungsmuster abgeleitet ist):

(E) HAPLOTYP:

(F) GEWEBETYP

(G) ZELLTYP

(H) ZELLINIE

(I) ORGANELLE:

(vii) UNMITTELBARE HERKUNFT (unmittelbare experimentelle Herkunft der Sequenz gemäß SEQ ID NO:X):

(A) BIBLIOTHEK (Art der Bibliothek, Name)

(B) CLON(E)

(viii) POSITION IM GENOM (Position der Sequenz gemäß SEQ ID NO:X im Genom)

(A) CHROMOSOM/SEGMENT (Name/Nummer des Chromosoms/Segments)

(B) KARTENPOSITION

(C) EINHEITEN (Einheiten für die Kartenposition, d h Angabe, ob es sich bei den Einheiten um Genomprozent, Nucleotidnummern oder etwa anderes (bitte angeben) handelt):

(ix) MERKMALE (Beschreibung biologisch signifikanter Stellen in der Sequenz gemäß SEQ ID NO:X) (kann, je nach Zahl der angegebenen Merkmale, mehrmals vorkommen):

Entsprechend dem in den bestehenden Datenbanken für Nichtpatentliteratur eingeführten Merkmalsvokabular können signifikante Merkmale folgendes umfassen

aktiver Bereich  
Allel  
Attenuator  
Bindungsstelle  
CAAT-Signal  
zellulär  
Spaltstelle  
codierende Sequenz  
Konflikt  
Quervernetzung  
D-loop  
Disulfidbindung  
Domäne  
Duplikation  
Enhancer

- N-terminal fragment
- C-terminal fragment
- internal fragment

(vi) ORIGINAL SOURCE (original source of molecule sequenced in SEQ ID NO:X):

(A) ORGANISM (scientific name of source organism):

(B) STRAIN:

(C) INDIVIDUAL/ISOLATE (name/number of individual/isolate):

(D) DEVELOPMENT STAGE (give development stage of source organism and indicate whether derived from germ-line or rearranged developmental pattern):

(E) HAPLOTYPE:

(F) TISSUE TYPE:

(G) CELL TYPE:

(H) CELL LINE:

(I) ORGANELLE:

(vii) IMMEDIATE SOURCE (immediate experimental source of the sequence in SEQ ID NO:X):

(A) LIBRARY (library-type, name):

(B) CLONE(S):

(viii) POSITION IN GENOME (position of sequence in SEQ ID NO:X in genome):

(A) CHROMOSOME/SEGMENT (chromosome/segment-name/number):

(B) MAP POSITION:

(C) UNITS (units for map position, i.e. whether units are genome percent, nucleotide number or other (specify))

(ix) FEATURES (description of points of biological significance in the sequence in SEQ ID NO:X) (may be repeated depending on the number of features indicated):

Following the controlled vocabulary of the features field in the existing non-patent literature databases, significant features might include:

active-site  
allele  
attenuator  
binding-site  
CAAT signal  
cellular  
cleavage-site  
coding sequence  
conflict  
cross-links  
D-loop  
disulfide-bond  
domain  
duplication  
enhancer

- fragment N-terminal
- fragment C-terminal
- fragment interne

(vi) ORIGINE (origine de la molécule séquencée dans SEQ ID NO:X):

(A) ORGANISME (nom scientifique de l'organisme d'origine):

(B) SOUCHE:

(C) INDIVIDUEL OU ISOLE (nom/numéro de l'organisme individuel/isolé):

(D) STADE DE DEVELOPPEMENT (indiquer le stade de développement de l'organisme source et si l'organisme est issu d'une lignée germinale ou d'un schéma de développement réagencé):

(E) HAPLOTYPE:

(F) TYPE DE TISSU:

(G) TYPE DE CELLULE:

(H) LIGNEE CELLULAIRE:

(I) ORGANELLE:

(vii) SOURCE IMMEDIATE (source expérimentale immédiate de la séquence dans SEQ ID NO:X):

(A) BIBLIOTHEQUE (type, nom):

(B) CLONE(S):

(viii) POSITION DANS LE GENOME (position de la séquence SEQ ID NO:X dans le génome):

(A) CHROMOSOME/SEGMENT (nom/numéro du chromosome/segment):

(B) POSITION SUR LA CARTE:

(C) UNITES (unités utilisées pour la position sur la carte: pourcentage du génome, numéro du nucléotide ou autre (préciser)):

(ix) CARACTERISTIQUES (description des points significatifs sur le plan biologique dans la séquence SEQ ID NO:X) (cette rubrique peut être répétée autant de fois qu'il y a de caractéristiques indiquées):

Conformément à la terminologie établie en ce qui concerne les caractéristiques dans les bases de données non-brevet existantes, les caractéristiques importantes peuvent être les suivantes:

site actif  
allèle  
atténuateur  
site de fixation  
signal CAAT  
cellulaire  
site de clivage  
séquence codante  
en opposition  
liaisons croisées  
boucle D  
liaison bisulfure  
domaine  
duplication  
activateur

Exon	exon	exon
GC-Signal	GC signal	signal GC
i-DNA	iDNA	ADNi
Inhibitionsbereich	inhibitory-site	site d'inhibition
Insertionssequenz	insertion sequence	séquence d'insertion
Intron	intron	intron
LTR (long terminal repeat)	LTR (long terminal repeat)	LTR (longue région terminale répétée)
reifes Peptid	mature peptide	peptide mature
modifizierte Base oder Aminosäure	modified base or amino acid	base ou acide aminé modifié(e)
m-RNA	mRNA	ARNm
Mutation	mutation	mutation
Peptid	peptide	peptide
PolyA-Signal	polyA signal	signal poly A
PolyA-Bereich	polyA site	site poly A
Vorläufer-RNA	precursor RNA	ARN précurseur
Primärtranskript	primary transcript	transcript primaire
primerbindend	primer binding	fixation de l'amorce
Promotor	promoter	promoteur
Provirus	provirus	provirus
RBS (Ribosomenbindungsstelle)	RBS (ribosome binding site)	RBS (site de fixation du ribosome)
Repetitionseinheit	repeating unit	unité répétée
repetitiver Abschnitt	repeat region	région répétée
Replikationsursprung	replication origin	origine de réplication
r-RNA	rRNA	ARNr
Satellit	satellite	satellite
sc-RNA	scRNA	ARNsc
Signalpeptid	signal peptide	peptide signal
sn-RNA	snRNA	ARNsn
stem loop	stem loop	structure en épingle à cheveux
TATA-Signal	TATA signal	signal TATA
Terminator	terminator	terminateur
Thiolesterbindung	thiolester-bond	liaison thiolester
Transit-Peptid	transit peptide	peptide de transit
Transposon	transposon	transposon
t-RNA	tRNA	ARNt
unsicher	unsure	indéterminé
Variation	variation	variation
Virion	virion	virion
3'Clip	3'clip	3'clip
3'UTR	3'UTR	3'UTR
5'Clip	5'clip	5'clip
5'UTR	5'UTR	5'UTR
-10-Signal	-10 signal	signal-10
-35-Signal	-35 signal	signal-35
oder sonstige	or others:	ou autres;

Es wird empfohlen, sich bei der Beschreibung von Sequenzmerkmalen an die Merkmalstabelle "The DDBJ/EMBL/GenBank Feature Table: Definition" zu halten

It is recommended to describe features of sequences in accordance with the "The DDBJ/EMBL/GenBank Feature Table: Definition"

Il est recommandé de décrire les caractéristiques des séquences conformément à l'ouvrage intitulé "The DDBJ/EMBL/GenBank Feature Table: Definition".

(A) NAME/SCHLÜSSEL (bitte entsprechende Bezeichnung für das Merkmal angeben):

(A) NAME/KEY (provide appropriate identifier for feature):

A) NOM/CLE (indiquer le numéro d'identification approprié de la caractéristique):

(B) LAGE

(B) LOCATION:

B) EMPLACEMENT:

- von (Nummer der ersten Base/Aminosäure des Merkmals)

- from (number of first base/amino acid in the feature)

- de (numéro de la première base/du premier acide aminé dans la caractéristique)

- bis (Nummer der letzten Base/Aminosäure des Merkmals)

- to (number of last base/amino acid in the feature)

- à (numéro de la dernière base/du dernier acide aminé dans la caractéristique)

- Basenpaare (Ziffern verweisen auf die Positionen der Basenpaare in einer Nucleotidsequenz)

- base pairs (numbers refer to positions of base pairs in a nucleotide sequence)

- paires de bases (les numéros renvoient aux positions des paires de bases dans une séquence de nucléotides)

- Aminosäuren (Ziffern verweisen auf die Positionen der Aminosäurereste in einer Aminosäuresequenz)

- amino acids (numbers refer to positions of amino acid residues in an amino acid sequence)

- acides aminés (les numéros renvoient aux positions des résidus d'acide aminé dans une séquence d'acides aminés)

- Angabe, ob sich das Merkmal auf dem zum Strang des Sequenzprotokolls komplementären Strang befindet

- whether feature is located on the complementary strand to that filed in the Sequence Listing

- déterminer si la caractéristique se situe sur le brin complémentaire de celui figurant dans la liste des séquences

(C) ERMITTLUNGSMETHODE (mit der das Merkmal ermittelt wurde):

- experimentell
- durch Ähnlichkeit mit einer bekannten Sequenz oder einer feststehenden Consensus-Sequenz
- durch Ähnlichkeit mit einem anderen Schema

(D) SONSTIGE ANGABEN:

- assoziierter Phänotyp(en)
- biologische/enzymatische Wirkung
- biologische/enzymatische Wirkung des Produkts
- allgemeine Funktionsklasse des Gens und/oder Genprodukts
- Makromolekülbindung - Makromoleküle, an die sich das Genprodukt binden kann
- subzelluläre Lokalisation - subzelluläre Lokalisation des Genprodukts
- sonstige sachdienliche Angaben

(xi) SEQUENZBESCHREIBUNG:  
SEQ ID NO:X:

BEISPIELE

23 Die Anlagen 1 und 2 veranschaulichen die Wiedergabe der besonders wichtigen Teile eines Sequenzprotokolls. Anlage 1 enthält verschiedene Beispiele für die Darstellung von Sequenzen, die zum Teil der vorliegenden Empfehlung nicht entsprechen. Die Beispiele 4, 6, 8, 9, 13 und 15 zeigen richtig formatierte vollständige Sequenzen. In den Beispielen 5, 10 und 11 werden bestimmte Ausschnitte aus richtig formatierten Sequenzen herausgegriffen, um zu zeigen, wie einzelne Teile der vorliegenden Empfehlung angewendet sind. In Anlage 2 sind Beispiele für die Angabe von Merkmalen in einem Sequenzprotokoll zusammengestellt.

24 Anlage 3 zeigt anhand eines Musters eines Sequenzprotokolls, das die wesentlichen Punkte der vorliegenden Empfehlung veranschaulicht, die vollständige Darstellung der Daten eines Sequenzprotokolls

(C) IDENTIFICATION METHOD (by which the feature was identified):

- experimentally
- by similarity with a known sequence or to an established consensus sequence
- by similarity to some other pattern

(D) OTHER INFORMATION:

- associated phenotype(s)
- biological/enzymatic activity
- biological/enzymatic activity of its product
- general functional class of the gene and/or gene product
- binding macromolecules - macromolecules to which the gene product can bind
- subcellular localisation - subcellular localisation of the gene product
- any other relevant information

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X:

EXAMPLES

23 To illustrate presentation of the especially important parts of a Sequence Listing Appendices 1 and 2 are provided. Examples of various kinds of sequence presentation are given in Appendix 1 and show presentation not complying with this Recommendation as well as presentation that does comply with this Recommendation. Examples 4, 6, 8, 9, 13, 15 represent properly formatted complete sequences. Examples 5, 10, 11 represent particular excerpts of properly formatted sequences to demonstrate application of parts of the Recommendation. Examples of features indication in a Sequence Listing are given in Appendix 2.

24 The complete presentation of data in a Sequence Listing is shown in Appendix 3 which gives a specimen Sequence Listing illustrating the major points of this Recommendation.

C) METHODE D'IDENTIFICATION de la caractéristique:

- expérimentale
- par similitude avec une séquence connue ou avec une séquence consensus connue
- par similitude avec d'autres modèles

D) AUTRES INFORMATIONS IMPORTANTES:

- phénotype(s) associé(s)
- activité biologique/enzymatique
- activité biologique/enzymatique de son produit
- classe de fonction générale du gène et/ou produit du gène
- macromolécules de fixation - macromolécules auxquelles le produit du gène peut se fixer
- localisation subcellulaire du produit du gène
- toute autre information pertinente

xi) DESCRIPTION DE SEQUENCE: SEQ ID NO:X:

EXEMPLES

23 Les appendices I et II illustrent la présentation des parties de la liste de séquences qui revêtent une importance particulière. On trouvera dans l'appendice I divers modes de présentation de séquences, dont certains ne satisfont pas aux dispositions de la présente recommandation. Les exemples 4, 6, 8, 9, 13 et 15 illustrent des séquences complètes convenablement formatées. Les exemples 5, 10 et 11 donnent certains extraits de séquences convenablement formatées pour illustrer l'application de certaines parties de la recommandation. Des exemples d'indication des caractéristiques dans une liste de séquences figurent à l'appendice 2.

24. L'appendice III indique, au moyen d'un modèle illustrant les principaux points de la présente recommandation, la présentation complète des données d'une liste de séquences.

**Anlage 1 zum WIPO-ST. 23****Beispiele für Sequenzangaben****DARSTELLUNG VON AMINOSÄURE-SEQUENZEN**

- 1 Unformatierte Darstellung des Peptids im einbuchstabigen Code (abweichend von der Empfehlung)
- 2 Unformatierte Darstellung desselben Peptids im dreibuchstabigen Code (abweichend von der Empfehlung)
- 3 Unformatierte untergliederte Darstellung desselben Peptids im dreibuchstabigen Code (abweichend von der Empfehlung)
- 4 Formatierte Darstellung desselben Peptids (wie empfohlen)
- 5 Formatiertes Peptid mit System zur Numerierung reifer Proteine (wie empfohlen)
- 6 Formatierte Peptidsequenz mit modifizierten oder seltenen Aminosäuren in den Positionen 4, 14, 37 und 38 (wie empfohlen)

**DARSTELLUNG VON NUCLEOTID-SEQUENZEN**

- 7 Unformatierte Darstellung der Nucleotidsequenz (abweichend von der Empfehlung)
- 8 Formatierte Darstellung derselben Nucleotidsequenz (wie empfohlen)
- 9 Formatierte Nucleotidsequenz mit codierendem Abschnitt und dem durch diese Sequenz codierten Peptid (wie empfohlen)
- 10 Format für den Übergang vom nichtcodierenden zum codierenden Abschnitt (wie empfohlen)
- 11 Format für ein aufgespaltenes Codon (wie empfohlen)

**DARSTELLUNG VON SEQUENZEN, VON DENEN NUR TEILE BEKANNT SIND**

- 12 Unformatiertes Peptid, von dem nur Teile bekannt sind (abweichend von der Empfehlung)
- 13 Formatierte Darstellung desselben Peptids (wie empfohlen)
- 14 Unformatiertes Nucleotid, von dem nur Teile bekannt sind (abweichend von der Empfehlung)
- 15 Formatierte Darstellung desselben Nucleotids (wie empfohlen)

- 1 UNFORMATIERTE DARSTELLUNG DES PEPTIDS IM EINBUCHSTABIGEN CODE (abweichend von der Empfehlung)

**Appendix 1 to WIPO ST. 23****Examples of sequence information****PRESENTATION OF AMINO ACID SEQUENCES**

1. Unformatted Presentation of the Peptide in Single Letter Code (not according to Recommendation)
2. Unformatted Presentation of the Same Peptide in Three Letter Code (not according to Recommendation)
3. Unformatted Separated Presentation of the Same Peptide in Three Letter Code (not according to Recommendation)
4. Formatted Presentation of the Same Peptide (according to Recommendation)
5. Formatted Peptide Showing Mature Protein Numbering Scheme (according to Recommendation)
6. Formatted Peptide Sequence with Modified or Unusual Amino Acids at Positions 4, 14, 37 and 38 (according to Recommendation)

**PRESENTATION OF NUCLEOTIDE SEQUENCES**

7. Unformatted Presentation of the Nucleotide Sequence (not according to Recommendation)
8. Formatted Presentation of the Same Nucleotide Sequence (according to Recommendation)
9. Formatted Nucleotide Sequence with Coding Region and the Peptide Coded by that Sequence (according to Recommendation)
10. Format for Transition from Noncoding to Coding Region (according to Recommendation)
11. Format for a Split Codon (according to Recommendation)

**PRESENTATION OF SEQUENCES OF WHICH ONLY PARTS ARE KNOWN**

12. Unformatted Peptide of Which Only Parts are Known (not according to Recommendation)
13. Formatted Presentation of the Same Peptide (according to Recommendation)
14. Unformatted Nucleotide of Which Only Parts are Known (not according to Recommendation)
15. Formatted Presentation of the Same Nucleotide (according to Recommendation)

- 1 UNFORMATTED PRESENTATION OF THE PEPTIDE IN SINGLE LETTER CODE (not according to Recommendation)

**Appendice 1 à la norme OMPI ST. 23****Exemples d'informations sur les séquences****PRESENTATION DES SEQUENCES D'ACIDES AMINES**

1. Présentation non formatée du peptide selon un code à une lettre (non conforme à cette recommandation)
2. Présentation non formatée du même peptide selon un code à trois lettres (non conforme à cette recommandation)
3. Présentation non formatée du même peptide selon un code à trois lettres avec séparations (non conforme à cette recommandation)
4. Présentation formatée du même peptide (conforme à cette recommandation)
5. Peptide formaté avec système de numérotation des protéines matures (conforme à cette recommandation)
6. Séquence de peptides formatée avec acides aminés modifiés ou peu connus aux positions 4, 14, 37 et 38 (conforme à cette recommandation)

**PRESENTATION DES SEQUENCES DE NUCLEOTIDES**

7. Présentation non formatée de la séquence de nucléotides (non conforme à cette recommandation)
8. Présentation formatée de la même séquence de nucléotides (conforme à cette recommandation)
9. Séquence de nucléotides formatée avec région codante et peptide codé par la séquence (conforme à cette recommandation)
10. Format pour la transition d'une région non codante à une région codante (conforme à cette recommandation)
11. Format pour un codon coupé (conforme à cette recommandation)

**PRESENTATION DES SEQUENCES DONT QUELQUES PARTIES SEULEMENT SONT CONNUES**

12. Peptide non formaté dont quelques parties seulement sont connues (non conforme à cette recommandation)
13. Présentation formatée du même peptide (conforme à cette recommandation)
14. Nucléotide non formaté dont quelques parties seulement sont connues (non conforme à cette recommandation)
15. Présentation formatée du même nucléotide (conforme à cette recommandation).

1. PRESENTATION NON FORMATEE DU PEPTIDE SELON UN CODE A UNE LETTRE (non conforme à cette recommandation)

**AGLLAGSWIPDWTFFVSVPLVLTWYTLTKEPIGEDVYYVDGACNRNSREGKAGYITQQG  
KQRVEKLENTTNQQAELTAIKMALEDGPRVNIIVTDSQYA**

2 UNFORMATIERTE DARSTELLUNG  
DESSELBEN PEPTIDS IM DREIBUCH-  
STABIGEN CODE (abweichend von  
der Empfehlung)

2. UNFORMATTED PRESENTATION  
OF THE SAME PEPTIDE IN THREE  
LETTER CODE (not according to  
Recommendation)

2. PRESENTATION NON FORMATEE  
DU MEME PEPTIDE SELON UN CODE  
A TROIS LETTRES (non conforme à  
cette recommandation)

AlaGlyLeuLeuAlaGlySerTrpIleProAspTrpThrPheValSerValProProLeu  
ValThrLeuTrpTyrThrLeuThrLysGluProIleProGlyGluAspValTyrTyrVal  
AspGlyAlaCysAsnArgAsnSerArgGluGlyLysAlaGlyTyrIleThrGlnGlnGly  
LysGlnArgValGluLysLeuGluAsnThrThrAsnGlnGlnAlaGluLeuThrAlaIle  
LysMetAlaLeuGluAspSerGlyProArgValAsnIleValThrAspSerGlnTyrAla

3 UNFORMATIERTE UNTERGLIEDER-  
TE DARSTELLUNG DESSELBEN PEP-  
TIDS IM DREIBUCHSTABIGEN CODE  
(abweichend von der Empfehlung)

3. UNFORMATTED SEPARATED  
PRESENTATION OF THE SAME PEP-  
TIDE IN THREE LETTER CODE (not  
according to Recommendation)

3 PRESENTATION NON FORMATEE  
DU MEME PEPTIDE SELON UN CODE  
A TROIS LETTRES AVEC SEPARA-  
TIONS (non conforme à cette recom-  
mandation)

Ala Gly Leu Leu Ala Gly Ser Trp Ile Pro Asp Trp Thr Phe Val Ser  
Val Pro Pro Leu Val Thr Leu Trp Tyr Thr Leu Thr Lys Glu Pro Ile  
Pro Gly Glu Asp Val Tyr Tyr Val Asp Gly Ala Cys Asn Arg Asn Ser  
Arg Glu Gly Lys Ala Gly Tyr Ile Thr Gln Gln Lys Gln Arg Val Glu  
Lys Leu Glu Asn Thr Thr Asn Gln Gln Ala Glu Leu Thr Ala Ile Lys  
Met Ala Leu Glu Asp Ser Gly Pro Arg Val Asn Ile Val Thr Asp Ser  
Gln Tyr Ala

4 FORMATIERTE DARSTELLUNG  
DESSELBEN PEPTIDS  
(wie empfohlen)

4 FORMATTED PRESENTATION OF  
THE SAME PEPTIDE (according to  
Recommendation)

4. PRESENTATION FORMATEE DU  
MEME PEPTIDE (conforme à cette  
recommandation)

Ala Gly Leu Leu Ala Gly Ser Trp Ile Pro Asp Trp Thr Phe Val Ser  
1 5 10 15  
Val Pro Pro Leu Val Thr Leu Trp Tyr Thr Leu Thr Lys Glu Pro Ile  
20 25 30  
Pro Gly Glu Asp Val Tyr Tyr Val Asp Gly Ala Cys Asn Arg Asn Ser  
35 40 45  
Arg Glu Gly Lys Ala Gly Tyr Ile Thr Gln Gln Lys Gln Arg Val Glu  
50 55 60  
Lys Leu Glu Asn Thr Thr Asn Gln Gln Ala Glu Leu Thr Ala Ile Lys  
65 70 75 80  
Met Ala Leu Glu Asp Ser Gly Pro Arg Val Asn Ile Val Thr Asp Ser  
85 90 95  
Gln Tyr Ala

5 FORMATIERTES PEPTID MIT SY-  
STEM ZUR NUMERIERUNG REIFER  
PROTEINE (wie empfohlen)

5. FORMATTED PEPTIDE SHOWING  
MATURE PROTEIN NUMBERING  
SCHEME (according to Recommenda-  
tion)

5. PEPTIDE FORMATE AVEC SY-  
STEME DE NUMEROTATION DES PRO-  
TEINES MATURES (conforme à cette  
recommandation)

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu  
-50 -45 -40 -35  
Ala Leu Leu Pro Ala Ala Thr Gln Gly Lys Lys Val Val Leu Gly Lys  
-30 -25 -20

```

Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser
      -15                      -10                      -5
Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn
      1                      5                      10
Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala
15                      20                      25                      30
Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile
      35                      40                      45
Lys Asn Leu Lys Ile Glu .....
      50

```

6 FORMATIERTE PEPTIDSEQUENZ  
MIT MODIFIZIERTEN ODER SELTENE  
AMINOSÄUREN IN DEN POSI-  
TIONEN 4,14, 37 UND 38  
(wie empfohlen)

6. FORMATTED PEPTIDE SEQUENCE  
WITH MODIFIED OR UNUSUAL  
AMINO ACIDS AT POSITIONS 4, 14,  
37 AND 38 (according to Recommen-  
dation)

6. SEQUENCE DE PEPTIDES FORMA-  
TÉE AVEC ACIDES AMINES MODI-  
FIÉS OU PEU CONNUS AUX POSI-  
TIONS 4, 14, 37 ET 38 (conforme à  
cette recommandation)

```

Ala Gly Leu Leu ALa Gly Ser Trp Ile Pro Asp Trp Thr Xaa Val Ser
 1                      5                      10                      15
Val Pro Pro Leu Val Thr Leu Trp Tyr Thr Leu Thr Lys Glu Pro Ile
      20                      25                      30
Pro Gly Glu Asp Val Xaa Tyr Val Asp Gly Ala Cys Asn Arg Asn Ser
      35                      40                      45
Arg Glu Gly Lys Ala Gly Tyr Ile Thr Gln Gln Lys Gln Arg Val Glu
      50                      55                      60
Lys Leu Glu Asn Thr Thr Asn Gln Gln Ala Glu Leu Thr Ala Ile Lys
65                      70                      75                      80
Met Ala Leu Glu Asp Ser Gly Pro Arg Val Asn Ile Val Thr Asp Ser
      85                      90                      95
Gln Tyr Ala

```

Zur näheren Bezeichnung der  
2-Amino adipinsäure in Position 14 und  
der 3-Amino adipinsäure in Position 38  
mittels der unter Nummer 12 aufge-  
führten Abkürzungen wären im Se-  
quenzprotokoll folgende Angaben aus-  
zuweisen:

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: spezielles  
Merkmal

(B) LAGE: 14

(D) SONSTIGE ANGABEN Xaa ist Aad

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL spezielles  
Merkmal

(B) LAGE: 38

(D) SONSTIGE ANGABEN Xaa ist  
bAad

Wie unter Nummer 12 dargelegt, kön-  
nen in einigen Fällen unmodifizierte  
Aminosäuren für modifizierte oder sel-  
tene Aminosäuren in der Sequenz ste-  
hen; zur näheren Bezeichnung von  
Norleucin in Position 4 und Norvalin in  
Position 37 würden im Sequenzproto-  
koll dann beispielsweise folgende An-  
gaben ausgewiesen:

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: spezielles  
Merkmal

To indicate 2-Amino adipic acid at posi-  
tion 14 and 3-Amino adipic acid at  
position 38 using the abbreviations of  
paragraph 12, a Sequence Listing  
would contain the following:

(ix) FEATURE:

(A) NAME/KEY: misc-feature

(B) LOCATION: 14

(D) OTHER INFORMATION: Xaa is Aad

(ix) FEATURE:

(A) NAME/KEY: misc-feature

(B) LOCATION: 38

(D) OTHER INFORMATION: Xaa is  
bAad

As described in paragraph 12, in some  
cases unmodified amino acids may  
represent modified or unusual amino  
acids in the sequence, e.g. to indicate  
Norleucine at position 4 and Norvaline  
at position 37, a Sequence Listing  
would contain the following:

(ix) FEATURE:

(A) NAME/KEY: misc-feature

Pour signaler la présence d'un acide  
2-amino-adipique à la position 14 et  
d'un acide 3-amino-adipique à la posi-  
tion 38 en utilisant les abréviations  
données au paragraphe 12, il faut faire  
figurer les éléments suivants dans la  
liste de séquences:

(ix) CARACTÉRISTIQUE:

A) NOM/CLE: caractéristique particu-  
lière

B) EMPLACEMENT: 14

D) AUTRES INFORMATIONS: Xaa si-  
gnifie Aad

(ix) CARACTÉRISTIQUE:

A) NOM/CLE: caractéristique particu-  
lière

B) EMPLACEMENT: 38

D) AUTRES INFORMATIONS: Xaa si-  
gnifie bAad

Comme indiqué au paragraphe 12, des  
acides aminés non modifiés peuvent  
représenter dans certains cas des  
acides aminés modifiés ou peu  
connus dans la séquence. Ainsi, pour  
signaler la présence de norleucine à la  
position 4 et de norvaline à la position  
37, il faut faire figurer dans la liste de  
séquences les éléments suivants:

(ix) CARACTÉRISTIQUE:

A) NOM/CLE: caractéristique particu-  
lière

(B) LAGE: 4  
(D) SONSTIGE ANGABEN: Leu ist Nle

(B) LOCATION: 4  
(D) OTHER INFORMATION: Leu is Nle

B) EMPLACEMENT: 4  
D) AUTRES INFORMATIONS: Leu signifie Nle

(ix) MERKMAL:  
(A) NAME/SCHLÜSSEL: spezielles Merkmal  
(B) LAGE: 37  
(D) SONSTIGE ANGABEN: Val ist Nva

(ix) FEATURE  
(A) NAME/KEY: misc-feature  
(B) LOCATION: 37  
(D) OTHER INFORMATION: Val is Nva

ix) CARACTERISTIQUE:  
A) NOM/CLE: caractéristique particulière  
B) EMPLACEMENT: 37  
D) AUTRES INFORMATIONS: Val signifie Nva

7. UNFORMATIERTE DARSTELLUNG DER NUCLEOTIDSEQUENZ (abweichend von der Empfehlung)

7. UNFORMATTED PRESENTATION OF THE NUCLEOTIDE SEQUENCE (not according to Recommendation)

7 PRESENTATION NON FORMATEE DE LA SEQUENCE DE NUCLEOTIDES (non conforme à cette recommandation)

CAAGCCCAGAGCCCTGCCATTTCTGTGGGCTCAGGTCCTACTGCTCAGCCCCTTCCTCC  
CTCGGCAAGGCCACAATGAACCGGGAGTCCCTTTTAGGCACCTGCTTCTGGTGCTGCAA  
CTGGCGCTCCTCCCAGCAGCCACTCAGGGAAAGAAAGTGGTGCTGGGCAAAAAGGGGAT  
ACAGTGGAACTGACCTGTACAGCTTCCCAGAAGAAGAGCATAACAATCCACTGGAAAAAC  
TCCAACCAGATAAAGATTCTGGGAAATCAGGGCTCCTTCTTAATAAAGGTCCATCCAAG  
CTGAATGATCGCGTACTCAAGAAGAAGCCTTTGGGACCAAGGAACTTCCCCCTGATC  
ATCAAGAATCTTAAGATAGAAGACTCAGATACTTACATCTGTGAAGTGGAGGACCAGAAG  
GAGGAGGTGCAATTGCTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAG  
GGCAGAGCCTGACCCTGACCTTGGAGAGCCCCCTGGTAGTAGCCCCCTCAGTGCAATGT  
AGGAGTCCAAGGGTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAG  
CTCCAGGATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAAGGAGGAGTTC  
AAAATAGACATCGTGGTGCTAGCTTCCAGAAGGCCCTCCAGCATAGTCTATAAGAAAGAG  
GGGAACAGGTGGAGTTCCTTCCCCTCGCCTTTACAGTTGAAAAGCTGACGGGCAGT  
GGCGAGCTGTGGTGGCAGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTTTGAC  
CTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCAGGACCCTAAGCTCCAGATGGGC  
AAGAAGCTCCCGCTCCACCTCACCTGCCCGAGGCCCTTGCCCTCAGTATGCTGGCTCTGGA  
AACCTCACCTGGCCCTTGAAGCGAAAACAGGAAAGTTGCATCAGGAAGTGAACCTGGTG  
GTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCC  
CCTAAGCTGATGCTGAGCTTGAAGCTGGAGAACAAAGGAGGCAAAGGTCTCGAAGCGGGAG  
AAGGCGGTGTGGGTGCTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGACTCG  
GGACAGGTCTGCTGGAATCCAACATCAAGGTTCTGCCACATGGTCCACCCGGTGCAG  
CCAATGGCCCTGATTGTGCTGGGGGGCGTCGCCGGCCTCCTGCTTTTTCATTGGGCTAGGC  
ATCTTCTTCTGTGTCAGGTGCCGGCACCGAAGGCGCCAAGCAGAGCGGATGTCTCAGATC  
AAGAGACTCCTCAGTGAGAAGAAGACCTGCCAGTGCCCTCACCGGTTTTCAGAAGACATGT  
AGCCCCATTTGAGGCACGAGGCCAGGCAGATCCCCTTGCAGCCTCCCCAGGTGTCTGCC  
CCGCGTTTCTGCTGCGGACCAGATGAATGTAGCAGATCCCACGCTCTGGCCTCCTGTT  
CGTCCCTCCCTACAATTTGCCATTGTTTCTCCTGGGTTAGGCCCGGCTTCACTGGTTGAG  
TGTTGCTCTTAGTTTCCAGAGGCTTAATCACACCGTCCCTCCACGCCATTTCTTTTCT  
TCAAGCCTAGCCCTTCTCTCATTATTTCTCTCTGACCCTTCCCCACTGCTCATTTGGAT  
CC

8 FORMATIERTE DARSTELLUNG DERSELBEN NUCLEOTIDSEQUENZ (wie empfohlen)

8 FORMATTED PRESENTATION OF THE SAME NUCLEOTIDE SEQUENCE (according to Recommendation)

8. PRESENTATION FORMATEE DE LA MEME SEQUENCE DE NUCLEOTIDES (conforme à cette recommandation)

CAAGCCCAGA	GCCCTGCCAT	TTCTGTGGGC	TCAGGTCCTC	ACTGCTCAGC	CCCTTCCTCC	60
CTCGGCAAGG	CCACAATGAA	CCGGGGAGTC	CCTTTTAGGC	ACTTGCTTCT	GGTGCTGCAA	120
CTGGCGCTCC	TCCCAGCAGC	CACTCAGGGA	AAGAAAGTGG	TGCTGGGCAA	AAAAGGGGAT	180
ACAGTGGAAAC	TGACCTGTAC	AGCTTCCCAG	AAGAAGAGCA	TACAATCCA	CTGGAAAAAC	240
TCCAACCAGA	TAAAGATTCT	GGGAAATCAG	GGCTCCTTCT	TAATAAAGG	TCCATCCAAG	300
CTGAATGATC	GCGCTGACTC	AAGAAGAAGC	CCTTGGGACC	AAGGAACTT	CCCCCTGATC	360
ATCAAGAATC	TTAAGATAGA	AGACTCAGAT	ACTTACATCT	GTGAAGTGA	GGACCAGAAG	420
GAGGAGGTGC	AATTGCTAGT	GTTTCGGATTG	ACTGCCAACT	CTGACACCCA	CCTGCTTCAG	480
GGGCAGAGCC	TGACCCTGAC	CCTGGAGAGC	CCCCCTGGTA	GTAGCCCCTC	AGTGCAATGT	540
AGGAGTCCAA	GGGGTAAAAA	CATACAGGGG	GGGAAGACCC	TCTCCGTGTC	TCAGCTGGAG	600
CTCCAGGATA	GTGGCACCTG	GACATGCACT	GTCTTGCAGA	ACCAGAAGAA	GGTGGAGTTC	660





CTG ACC TTG GAG AGC CCC CCT GGT AGT AGC CCC TCA GTG CAA TGT ATT	543
Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg	
95 100 105	
ATG GCC CTG ATT GTG CTG GGG GGC GTC GCC GGC CTC CTG CTT TTC ATT	765
Met Ala Leu Ile Val Leu Gly Gly Val Ala Gly Leu Leu Leu Phe Ile	
350 355 360	
GGG CTA GGC ATC TTC TTC TGT GTC AGG TGC CGG CAC CGA AGG CGC CAA	813
Gly Leu Gly Ile Phe Phe Cys Val Arg Cys Arg Cys His Arg Arg Gln	
365 370 375	
GCA GAG CGG ATG TCT CAG ATC AAG AGA CTC CTC AGT GAG AAG AAG ACC	861
Ala Glu Arg Met Ser Gln Ile Lys Arg Leu Leu Ser Glu Lys Lys Thr	
380 385 390 395	
TGC CAG TGC CCT CAC CGG TTT CAG AAG ACA TGT AGC CCC ATT TGAGGCACGA	913
Cys Gln Cys Pro His Arg Phe Gln Lys Thr Cys Ser Pro Ile	
400 405	
GGCCAGGCAG ATCCCACTTG CAGCCTCCCC AGGTGTCTGC CCCGCGTTTC CTGCCTGCGG	972
ACCAGATGAA TGTAGCAGAT CCCACGCTCT GGCCTCCTGT TCGTCCTCCC TACAATTTGC	1132
CATTGTTTCT CCTGGGTTAG GCCCCGGCTT CACTGGTTGA GTGTTGCTCT CTAGTTTCCA	1192

10 FORMAT FÜR DEN ÜBERGANG VOM NICHTCODIERENDEN ZUM CODIERENDEN ABSCHNITT (wie empfohlen)

10 FORMAT FOR TRANSITION FROM NONCODING TO CODING REGION (according to Recommendation)

10. FORMAT POUR LA TRANSITION D'UNE REGION NON CODANTE A UNE REGION CODANTE (conforme à cette recommandation)

CAAGCCCAGA GCCCTGCCAT TTCTGTGGGC TCAGGTCCCT ACTGCTCAGC CCCTTCCTCC	345
CTGGGCAAGG CCACA ATG AAC CGG GGA GTC CCT TTT AGG CAC TTG CTT CTG	386
Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu	
-45 -40	

11 FORMAT FÜR EIN AUFGESPALTENES CODON (wie empfohlen)

11 FORMAT FOR ASPLIT CODON (according to Recommendation)

11. FORMAT POUR UN CODON COUPE (conforme à cette recommandation)

..... GGT AAA AAC ATA CAG GGG GGG AAG ACC C TCAGACTCCA	368
..... Gly Lys Asn Ile Gln Gly Gly Lys Thr	
115	
TCTGTCTGT TC CAG GAT AGT CGC ACC TGG ACA TGC ACT .....	
Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr .....	
120 125	

CTC codiert für Leu in Position 120 Leu in der Peptidsequenz wird unter die zwei Nucleotidbasen umfassende Gruppe TC gestellt und nicht aufgespalten oder unter dem C in Position 358 der Nucleotidkette angeordnet.

CTC codes for the Leu at 120. The Leu in the peptide sequence is positioned under the group of two nucleotide bases TC rather than being split or positioned under the C at 358 in the nucleotide chain.

Codes CTC pour la Leu à 120 Dans la séquence de peptides, la Leu est positionnée sous le groupe de deux bases de nucléotides TC au lieu d'être coupée ou positionnée sous le C à 358 dans la chaîne de nucléotides.

12. UNFORMATIERTES PEPTID, VON DEM NUR TEILE BEKANNT SIND (abweichend von der Empfehlung)

12. UNFORMATTED PEPTIDE OF WHICH ONLY PARTS ARE KNOWN (not according to Recommendation)

12. PEPTIDE NON FORMATE DONT QUELQUES PARTIES SEULEMENT SONT CONNUES (non conforme à cette recommandation)

Ala Gly Leu Leu Ala Gly Ser Trp Ile Pro Asp Trp Thr Phe Val Ser  
 Val Pro Pro Leu Val Thr Leu Trp ..... Ile  
 Pro Gly Glu Asp Val Tyr Tyr Val Asp Gly Ala Cys Asn Arg Asn Ser  
 Arg Glu Gly ..... Lys Gln Arg Val Glu  
 Lys Leu Glu Asn Thr Thr Asn Gln Gln Ala Glu Leu Thr Ala Ile Lys  
 Met Ala Leu Glu Asp Ser Gly Pro Arg Val Asn Ile Val Thr Asp Ser  
 Gln Tyr Ala

Unbekannte Abschnitte sind durch ..... wiedergegeben.

\* Unknown regions are represented by .....

\*\* Régions non connues représentées par .....

13. FORMATIERTE DARSTELLUNG DESSELBEN PEPTIDS (wie empfohlen)

13 FORMATTED PRESENTATION OF THE SAME PEPTIDE (according to Recommendation)

13. PRESENTATION FORMATEE DU MEME PEPTIDE (conforme à cette recommandation)

SEQ ID NO:1 :

Ala Gly Leu Leu Ala Gly Ser Trp Ile Pro Asp Trp Thr Phe Val Ser  
 1 5 10 15  
 Val Pro Pro Leu Val Thr Leu Trp  
 20

SEQ ID NO:2 :

Ile Pro Gly Glu Asp Val Tyr Tyr Val Asp Gly Ala Cys Asn Arg Asn  
 1 5 10 15  
 Ser Arg Glu Gly  
 20

SEQ ID NO:3 :

Lys Gln Arg Val Glu Lys Leu Glu Asn Thr Thr Asn Gln Gln Ala Glu  
 1 5 10 15  
 Leu Thr Ala Ile Lys Met Ala Leu Glu Asp Ser Gly Pro Arg Val Asn  
 20 25 30  
 Ile Val Thr Asp Ser Gln Tyr Ala  
 35 40

14. UNFORMATIERTES NUCLEOTID, VON DEM NUR TEILE BEKANNT SIND (abweichend von der Empfehlung)

14 UNFORMATTED NUCLEOTIDE OF WHICH ONLY PARTS ARE KNOWN (not according to Recommendation)

14 NUCLEOTIDE NON FORMATE DONT QUELQUES PARTIES SEULEMENT SONT CONNUES (non conforme à cette recommandation)

AAATAGACCTCACCCCTTACCCACTTCCCCTAGCGCTGAAA.....  
 .....AAATTTGCCTAGTCAAAATAAAAGATGCCGAGTCTATAAAAGCGCAAG  
 GACAGTTCAGGAGGTGGCTCGCTCCCTCACCGACCCTCTGGTCACGGAGACTCACCTTGG  
 GGATCCATCCTC.....TCCGTGGGACCGTCTCCCGGCC  
 TCGGCACCTCCTGAACTGCTCCTCCCAAGGTAAGTCTCCTCTAGGGTTCGAGCTCGGCTGC  
 CCCTTAGGTA CTGCTCCCCGAGGGTCTTTAGAGACAC.....  
 CTAGACTCTGCCTTAAACTTCACTTCCGCGTTCTTGTCTCGTTCTTTCCTCTTCGCCGTC  
 ACTGAAAACGAAACCTCAACGCCGCCCTCTTGGC.....  
 TGGAGCGCAGCAAGGGCTAGGGCTTCTGAACCTCTCCGGGAGAGGTCTATTGCTATAGG  
 CAGGCCCGCCCTAGGAGCATTTGTCTTCCCGGGGAAGACAAACAATTGGGGGCTCGTCCGG  
 GATTTGAATTCTCCATTCTCACATTATGGGACAAATCCACGGGCTTCCCCAACTCCAA

Unbekannte Abschnitte sind durch ..... wiedergegeben.

\* Unknown regions are indicated by .....

\*\* Régions non connues représentées par .....

15. FORMATIERTE DARSTELLUNG  
DESSELBEN NUCLEOTIDS  
(wie empfohlen)

15. FORMATTED PRESENTATION OF  
THE SAME NUCLEOTIDE (according  
to Recommendation)

15 PRESENTATION FORMATEE DU  
MEME NUCLEOTIDE (conforme à  
cette recommandation)

SEQ ID NO:1 :

AAATAGACCT CACCCTTACC CACTTCCCCT AGCGCTGAAA 40

SEQ ID NO:2 :

AAATTTGCCT AGTCAAATA AAAGATGCCG AGTCTATAAA AGCGCAAGGA CAGTTCAGGA 60  
GGTGGCTCGC TCCCTCACCG ACCCTCTGGT CACGGAGACT CACCTTGGGG ATCCATCCTC 120

SEQ ID NO:3 :

TCCGTGGGAC CGTCTCCCGG CCTCGGCACC TCCTGAACTG CTCCTCCCAA GGTAAGTCTC 60  
CTCTCAGGTC GAGCTCGGCT GCCCCTTAGG TAGTCGCTCC CCGAGGGTCT TTAGAGACAC 120

SEQ ID NO:4 :

CTAGACTCTG CCTTAACTT CACTTCCGCG TTCTTGTCTC GTTCTTTCCT CTTCGCCGTC 60  
ACTGAAAACG AAACCTCAAC GCCGCCCTCT TGGC 94

SEQ ID NO:5 :

TGGAGCGCAG CAAGGGCTAG GGCTTCCTGA ACCTCTCCGG GAGAGGTCTA TTGCTATAGG 60  
CAGGCCCGCC CTAGGAGCAT TGTCTTCCCG GGAAGACAA ACAATTGGGG GCTCGTCCGG 120  
GATTTGAATT CCTCCATTCT CACATTATGG GACAAATCCA CGGGCTTTC CCAACTCCAA 180

**Anlage 2 zum WIPO-ST. 23****Beispiele für die Angabe von Merkmalen in einem Sequenzprotokoll**

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: Exon  
 (B) LAGE: 194..264  
 (D) SONSTIGE ANGABEN:  
 /Produkt = "Vorläufer der Beta-Untereinheit von Thyrotropin"

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: Signalpeptid  
 (B) LAGE: 4213..4272  
 (D) SONSTIGE ANGABEN:  
 /Produkt = "Signalpeptid der Beta-Untereinheit von Thyrotropin"  
 /Standardname = "TSH-beta"

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: reifes Peptid  
 (B) LAGE: Verbindung (4273..4374, 4752..5003)  
 (D) SONSTIGE ANGABEN:  
 /Produkt = "Beta-Untereinheit von Thyrotropin"  
 /Standardname = "TSH-beta"

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: codierende Sequenz  
 (B) LAGE: Verbindung (4213 4374, 4752 5006)  
 (D) SONSTIGE ANGABEN:  
 /Produkt = "Vorläufer der Beta-Untereinheit von Thyrotropin"  
 /Standardname = "TSH-beta"

**Appendix 2 to WIPO ST. 23****Examples of features indication in a sequence listing**

(ix) FEATURE:

(A) NAME/KEY: exon  
 (B) LOCATION: 194..264  
 (D) OTHER INFORMATION:  
 /product = "Thyrotropin beta subunit precursor"

(ix) FEATURE:

(A) NAME/KEY: signal peptide  
 (B) LOCATION: 4213..4272  
 (D) OTHER INFORMATION:  
 /product = "Thyrotropin beta subunit signal peptide"  
 /standard name = "TSH-beta"

(ix) FEATURE:

(A) NAME/KEY: mature peptide  
 (B) LOCATION: join (4273..4374, 4752..5003)  
 (D) OTHER INFORMATION:  
 /product = "Thyrotropin beta subunit"  
 /standard name = "TSH-beta"

(ix) FEATURE:

(A) NAME/KEY: coding sequence  
 (B) LOCATION: join (4213 4374, 4752..5006)  
 (D) OTHER INFORMATION:  
 /product = "Thyrotropin beta subunit precursor"  
 /standard name = "TSH-beta"

**Appendice 2 à la Norme OMPI ST.23****Exemples d'indication des caractéristiques dans une liste de séquences**

ix) CARACTERISTIQUE:

A) NOM/CLE: exon  
 B) EMBLACEMENT: 194..264  
 D) AUTRES INFORMATIONS:  
 /produit = "Précurseur de sous-unité de la thyrotropine bêta"

ix) CARACTERISTIQUE:

A) NOM/CLE: peptide signal  
 B) EMBLACEMENT: 4213..4272  
 D) AUTRES INFORMATIONS:  
 /produit = "Peptide signal de sous-unité de la thyrotropine bêta"  
 /dénomination courante = "TSH-bêta"

ix) CARACTERISTIQUE:

A) NOM/CLE: peptide mature  
 B) EMBLACEMENT: join (4273..4374, 4752 5003)  
 D) AUTRES INFORMATIONS:  
 /produit = "Sous-unité de la thyrotropine bêta"  
 /dénomination courante = "TSH-bêta"

ix) CARACTERISTIQUE:

A) NOM/CLE: séquence codante  
 B) EMBLACEMENT: join (4213..4374, 4752..5006)  
 D) AUTRES INFORMATIONS:  
 /produit = "Précurseur de sous-unité de la thyrotropine bêta"  
 /dénomination courante = "TSH-bêta"

**Anlage 3 zum WIPO-ST. 23****Appendix 3 to WIPO ST. 23****Appendice 3 à la norme OMPI  
ST. 23****Muster eines Sequenzprotokolls****Specimen sequence listing****Exemple d'une liste de séquences**

ANZAHL DER SEQUENZEN: 1

NUMBER OF SEQUENCES: 1

NOMBRE DE SEQUENCES: 1

(1) ANGABEN ZU SEQ ID NO: 1

(1) INFORMATION FOR SEQ ID NO: 1

1) INFORMATIONS CONCERNANT  
SEQ ID NO:1:

(i) SEQUENZKENNZEICHEN:

(i) SEQUENCE CHARACTERISTICS:

i) CARACTERISTIQUES DE LA SE-  
QUENCE:(A) LÄNGE: 2654 Basenpaare  
(B) ART Nucleinsäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear(A) LENGTH: 2654 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linearA) LONGUEUR: 2654 paires de bases  
B) TYPE: acide nucléique  
C) NOMBRE DE BRINS: simple  
D) CONFIGURATION: linéaire(ii) ART DES MOLEKÜLS: DNA  
(genomisch)

(ii) MOLECULE TYPE: DNA (genomic)

ii) TYPE DE MOLECULE: ADN  
(génomique)

(vi) URSPRÜNGLICHE HERKUNFT:

(vi) ORIGINAL SOURCE:

vi) ORIGINE:

(A) ORGANISMUS: Homo sapiens  
(G) ZELLTYP Leukocyt(A) ORGANISM: Homo sapiens  
(G) CELL TYPE: LeukocyteA) ORGANISME: homo sapiens  
G) TYPE DE CELLULE: leucocyte

(ix) MERKMAL:

(ix) FEATURE:

ix) CARACTERISTIQUE:

(A) NAME/SCHLÜSSEL TATA-Signal  
(B) LAGE: 339..344(A) NAME/KEY: TATA-signal  
(B) LOCATION: 339..344A) NOM/CLE: signal TATA  
B) LOCALISATION: 339..344

(ix) MERKMAL:

(ix) FEATURE:

ix) CARACTERISTIQUE:

(A) NAME/SCHLÜSSEL: Intron  
(B) LAGE: 527..814(A) NAME/KEY: intron  
(B) LOCATION: 527..814A) NOM/CLE: intron  
B) LOCALISATION: 527..814

(ix) MERKMAL:

(ix) FEATURE:

ix) CARACTERISTIQUE:

(A) NAME/SCHLÜSSEL: Intron  
(B) LAGE: 923..1006(A) NAME/KEY: intron  
(B) LOCATION: 923..1006A) NOM/CLE: intron  
B) LOCALISATION: 923 1006

(ix) MERKMAL:

(ix) FEATURE:

ix) CARACTERISTIQUE:

(A) NAME/SCHLÜSSEL: Intron  
(B) LAGE 1113..1359(A) NAME/KEY: intron  
(B) LOCATION: 1113..1359A) NOM/CLE: intron  
B) LOCALISATION: 1113..1359

(ix) MERKMAL

(ix) FEATURE:

ix) CARACTERISTIQUE:

(A) NAME/SCHLÜSSEL: Signalpeptid  
(B) LAGE: 824 1009(A) NAME/KEY: signal peptide  
(B) LOCATION: 824..1009A) NOM/CLE: peptide signal  
B) LOCALISATION: 824..1009

(ix) MERKMAL

(ix) FEATURE:

ix) CARACTERISTIQUE:

(A) NAME/SCHLÜSSEL: reifes Peptid  
(B) LAGE 1010 1772(A) NAME/KEY: mature peptide  
(B) LOCATION: 1010..1772A) NOM/CLE: peptide mature  
B) LOCALISATION: 1010..1772

(ix) MERKMAL

(ix) FEATURE:

ix) CARACTERISTIQUE:

(D) SONSTIGE ANGABEN humanes  
Lymphotoxin-Gen(D) OTHER INFORMATION: human  
lymphotoxin geneD) AUTRES INFORMATIONS: gène  
humain de la lymphotoxine(xi) SEQUENZBESCHREIBUNG  
SEQ ID NO 1(xi) SEQUENCE DESCRIPTION:  
SEQ ID NO:1:xi) DESCRIPTION DE SEQUENCE:  
SÉQ ID NO:1:

AAGGGTGCAG	AGATGTTATA	TATGATTGCT	CTTCAGGGAA	CCGGCCTCCA	GCTCACACCC	60
CAGCTGCTCA	ACCGCCTCCT	CTCTGAATTG	ACTGTCCCTT	CTTTGGAAct	CTAGGCCTGA	120
CCCCACTCCC	TGGCCCTCCC	AGCCACGAT	TCCCCTGACC	CGACTCCCTT	TCCAGGAAct	180
CAGTCGCCTG	AACCCCCAGC	CTGTGGTTCT	CTCCTAGGCC	TCAGCCTTTC	CTGCCTTTGA	240
CTGAAACAGC	AGTATCTTCT	ACACGCTGGG	GCTTCCC GCG	GCCCAGCCCC	GACCTAGAAC	300
CCGCCCGCTG	CCTGCCACGC	TGCCACTGCC	GCTTCCCTCTA	TAAAGGGACC	TGAGCGTCCG	360
CGCGCAGGGG	CTCCACACAG	CAGGTGAGGC	TCTCCTGCCC	CATCTCCTTG	GGCTGCCCGT	420
GCTTCGTGCT	TTGGACTACC	GCCCCGAGTG	TCCTGCCCTC	TGCCTGGGCC	TCGGTCCCTC	480
CTGCACCTGC	TGCCTGGATC	CCCGGCCTGC	CTGGGCCCTGG	GCCTTGGTGG	GTTTGGTTTTT	540
GGTTTCCTTC	TCTGTCTCTG	ACTCTCCATC	TGTCAGTCTC	ATTGTCTCTG	TCACACATTC	600
TCTGTTTCTG	CCATGGTTCC	TCTCTGTTCC	CTTCCCTGTCT	CTCTCTGTCT	CCCTCTGCTC	660
ACCTTGGGGT	TTCTCTGACT	GCATCTTGTC	CCCTTCTCTG	TCCGATCTCT	CTCTCGGGGG	720
TCGGGGGGTG	CTGTCTCCCA	GGGCGGGAGG	TCTGTCTTCC	GCCGCGTGCC	CCGCCCGCT	780
CACTGTCTCT	CTCTCTCTCT	CTCTTTCICT	GCAGGTTCTC	CCC ATG ACA CCA CCT		835

Met Thr Pro Pro

GAA CGT CTC TTC CTC CCA AGG GTG TGT GGC ACC ACC CTA CAC CTC CTC	883
Glu Arg Leu Phe Leu Pro Arg Val Cys Gly Thr Thr Leu His Leu Leu	
-30 -25 -20	
CTT CTG GGG CTG CTG CTG GTT CTG CTG CCT GGG GCC CAG GTGAGGCAGC	932
Leu Leu Gly Leu Leu Leu Val Leu Leu Pro Gly Ala Gln	
-10 -5	
AGGAGAATGG GGGCTGCTGG GGTGGCTCAG CCAAACCTTG AGCCCTAGAG CCCCCCTCAA	992
CTCTGTTCTC CTAG GGG CTC CCT GGT GTT GGC CTC ACA CCT TCA GCT GCC	1042
Gly Leu Pro Gly Val Gly Leu Thr Pro Ser Ala Ala	
1 5 10	
CAG ACT GCC CGT CAG CAC CCC AAG ATG CAT CTT GCC CAC AGC AAC CTC	1090
Gln Thr Ala Arg Gln His Pro Lys Met His Leu Ala His Ser Asn Leu	
15 20 25	
AAA CCT GCT GCT CAC CTC ATT G GTAAACATCC ACCTGACCTC CCAGACATGT	1142
Lys Pro Ala Ala His Leu Ile	
30	
CCCCACCAGC TCTCCTCCTA CCCCTGCCTC AGGAACCCAA GCATCCACCC CTCTCCCCCA	1202
ACTTCCCCCA CGCTAAAAA AACAGAGGGA GCCCACTCCT ATGCCTCCCC CTGCCATCCC	1262
CCAGGAACTC AGTTGTTTCAG TGCCCACTTC CTCAGGGATT GAGACCTCTG ATCCAGACCC	1322
CTGATCTCCC ACCCCCATCC CCTATGGCTC TTCCTAG GA GAC CCC AGC AAG CAG	1376
Gly Asp Pro Ser Lys Gln	
35 40	
AAC TCA CTG CTC TGG AGA GCA AAC ACG GAC CGT GCC TTC CTC CAG GAT	1424
Asn Ser Leu Leu Trp Arg Ala Asn Thr Asp Arg Ala Phe Leu Gln Asp	
45 50 55	
GGT TTC TCC TTG AGC AAC AAT TCT CTC CTG GTC CCC ACC AGT GGC ATC	1472
Gly Phe Ser Leu Ser Asn Asn Ser Leu Leu Val Pro Thr Ser Gly Ile	
60 65 70	
TAC TTC GTC TAC TCC CAG GTG GTC TTC TCT GGG AAA GCC TAC TCT CCC	1520
Tyr Phe Val Tyr Ser Gln Val Val Phe Ser Gly Lys Ala Tyr Ser Pro	
75 80 85	
AAG GCC ACC TCC TCC CCA CTC TAC CTG GCC CAT GAG GCT CAG CTC TTC	1568
Lys Ala Thr Ser Ser Pro Leu Tyr Leu Ala His Glu Val Gln Leu Phe	
90 95 100	
TCC TCC CAG TAC CCC TTC CAT GTG CCT CTC CTC AGC TCC CAG AAG ATG	1616
Ser Ser Gln Tyr Pro Phe His Val Pro Leu Leu Ser Ser Gln Lys Met	
105 105 115 120	
GTG TAT CCA GGG CTG CAG GAA CCC TGG CTG CAC TCG ATG TAC CAC GGG	1664
Val Tyr Pro Gly Leu Gln Glu Pro Trp Leu His Ser Met Tyr His Gly	
125 130 135	
GCT GCG TTC CAG CTC ACC CAG GGA GAC CAG CTA TCC ACC CAC ACA GAT	1712
Ala Ala Phe Gln Leu Thr Gln Gly Asp Gln Leu Ser Thr His Thr Asp	
140 145 150	
GGC ATC CCC CAC CTA GTC CTC AGC CCT AGT ACT GTC TTC TTT GGA GGC	1760
Gly Ile Pro His Leu Val Leu Ser Pro Ser Thr Val Phe Phe Gly Ala	
155 160 165	

TTC GCT CTG TAGAACTTGG AAAAATCCAG AAAGAAAAA TAATTGATTT C	1810
Phe Ala Leu	
170	
AAGACCTTCT CCCCATTTCTG CCTCCATTCT GACCATTTCA GGGGTCGTCA CCACCTCTCC	1870
TTTGGCCATT CCAACAGCTC AAGTCTTCCC TGATCAAGTC ACCGGAGCTT TCAAAGAAGG	1930
AATTCTAGGC ATCCCAGGGG ACCCACACCT CCCTGAACCA TCCCTGATGT CTGTCTGGCT	1990
GAGGATTTCA AGCCTGCCTA GGAATTCCCA GCCCAAAGCT GTTGGTCTTG TCCCACCAGC	2050
TAGGTGGGGC CTAGATCCAC ACACAGAGGA AGAGCAGGCA CATGGAGGAG CTTGGGGGAT	2110
GACTAGAGGC AGGGAGGGGA CTATTTATGA AGGCCAAAA ATTAAATTAT TTATTTATGG	2170
AGGATGGAGA GAGGGGAATA ATAGAAGAAC ATCCAAGGAG AAACAGAGAC AGGCCAAGA	2230
GATGAAGAGT GAGAGGGCAT GCGCACAAAG CTGACCAAGA GAGAAAGAAG TAGGCATGAG	2290
GGATCACAGG GCCCCAGAAG GCAGGGAAAG GCTCTGAAAG CCAGCTGCCG ACCAGAGCCC	2350
CACACGGAGG CATCTGCACC CTCGATGAAG CCAATAAAC CTCTTTTCTC TGAAATGCTG	2410
TCTGCTTGTG TGTGTGTGTC TGGGAGTGAG AACTTCCCAG TCTATCTAAG GAATGGAGGG	2470
AGGGACAGAG GGCTCAAAGG GACGAAGAGC TGTGGGGAGA ACAAAGGAT AAGGGCTCGA	2530
GAGAGCTTCA AGGATATGTG ATGGATCACC AGGTGAGGCC GCCAGACTGC TGCAGGGGAA	2590
GCAAAGGAGA AGCTGAGAAG ATGAAGGAAA AGTCAGGGTC TGGAGGGGGG GGGGTCAGGG	2650
AGCT	2654

## **ANHANG II zum Beschluß des Präsidenten des EPA vom 11. Dezember 1992**

### **Vorschriften über die Einreichung von Sequenzprotokollen auf elektronischem Datenträger**

#### **EINFÜHRUNG**

1 Der WIPO-Standard ST.23 – Empfehlung für die Darstellung von Nucleotid- und Aminosäuresequenzen in Patentanmeldungen und veröffentlichten Patentdokumenten – enthält spezifische, auch das Codierungssystem einschließende Empfehlungen für die Offenbarung von Nucleotid- und Aminosäuresequenzen im Rahmen eines "Sequenzprotokolls". ST 23 enthält jedoch keine Regeln für die Wiedergabe von Sequenzprotokollen auf vom Anmelder selbst erstellten elektronischen Datenträgern, die den Behörden für den gewerblichen Rechtsschutz bei der Einführung eines elektronischen Anmeldesystems und dem Einsatz computergestützter Rechensysteme sehr zustatten kämen

2 Die nachstehenden Vorschriften sollen zur Standardisierung bei der Einreichung von Sequenzprotokollen in computerlesbarer Form beitragen, um die Recherche nach Sequenzdaten, ihre Eingabe in computergestützte Datenbanken und ihren Austausch in elektronischer Form zu erleichtern. Sie entsprechen dem WIPO-Standard ST 24

3 Anlage 1 enthält alle Datenelemente des WIPO-Standards ST.24, deren Verwendung das EPA empfiehlt, und Anlage 2 ein Beispiel

**Anlage 3 enthält die Angaben, deren Verwendung das EPA zwingend vorschreibt.**

#### **DEFINITIONEN**

4 Im Rahmen der nachstehenden Vorschriften gelten folgende Definitionen:

i) Unter einer "Patentanmeldung" ist ein von einem Anmelder oder seinem Anwalt eingereichtes Dokument zu verstehen, mit dem um ein Patent nachgesucht wird. Es enthält in der Regel eine detaillierte Beschreibung der Erfindung, die Ansprüche und gegebenenfalls Zeichnungen, falls diese für das Verständnis der Erfindung erforderlich sind.

ii) Unter einem "Sequenzprotokoll" ist ein letzter, gesonderter Teil der Beschreibung in der maschinenschriftlichen Patentanmeldung zu verstehen, in dem die Nucleotid- und/oder Aminosäuresequenzen und andere verfügbare Informationen im einzelnen offen-

## **ANNEX II to the Decision of the President of the EPO dated 11 December 1992**

### **Rules for the Filing of Sequence Listings on Electronic Data Carriers**

#### **INTRODUCTION**

1. WIPO Standard ST.23 – Recommendation for the Presentation of Nucleotide and Amino Acid Sequences in Patent Applications and in Published Patent Documents provides specific recommendations, including the coding system, for disclosure of nucleotide and amino acid sequences as part of a "Sequence Listing". However, WIPO ST.23 contains no rules for the presentation by applicants of Sequence Listings on electronic data carriers, which would be very useful for industrial property offices wishing to introduce an electronic filing system and to use computerised search systems

2. The Rules set out below are intended to contribute to standardisation in the filing of Sequence Listings in computer-readable form so as to facilitate the searching of sequence data, their input into computerised databases and the exchange of the same in electronic form. They conform to WIPO Standard ST.24.

3 Appendix 1 contains all the data elements in WIPO Standard ST.24. the use of which is recommended by the EPO, and Appendix 2 contains an example.

**Appendix 3 sets out the information prescribed by the EPO.**

#### **DEFINITIONS**

4 For the purposes of these Rules:

(i) the expression "patent application" means a document filed by an applicant or by his/her agent requesting the grant of a patent. It usually contains a detailed description of the invention, the claims, and any drawings needed to understand the invention;

(ii) the expression "Sequence Listing" means a final, separate part of the description in the typed patent application, which gives a detailed disclosure of the nucleotide and/or amino acid sequences and other available information Each sequence disclosed

## **ANNEXE II à la décision du Président de l'OEB en date du 11 décembre 1992**

### **Prescriptions relatives au dépôt des listes de séquences sur des supports électroniques de données**

#### **INTRODUCTION**

1. La norme ST.23 de l'OMPI (Recommandation relative à la présentation des séquences de nucléotides et d'acides aminés dans les demandes de brevet et les documents de brevet publiés) prévoit des recommandations particulières, notamment la présentation normalisée, pour la divulgation des séquences de nucléotides et d'acides aminés dans les demandes de brevet sous la forme d'une "liste de séquences". Elle ne comporte toutefois aucune disposition relative à la présentation, par les déposants, des listes de séquences sur des supports électroniques de données, présentation conseillée aux offices de propriété industrielle qui souhaitent introduire un système de dépôt électronique et utiliser des systèmes de recherche informatisés.

2. Les présentes prescriptions ont été élaborées en vue de normaliser le dépôt des listes de séquences sous une forme déchiffrable par ordinateur et de faciliter ainsi la recherche des données relatives aux dites séquences, l'entrée de ces données dans les bases informatisées et leur échange sous forme électronique Elles correspondent à la norme OMPI ST 24

3 L'appendice 1 établit l'ensemble des renseignements de la norme OMPI ST.24 dont l'application est recommandée par l'OEB, tandis que l'appendice 2 présente un exemple.

**L'appendice 3 indique ceux des renseignements considérés comme obligatoires par l'OEB.**

#### **DEFINITIONS**

4 Aux fins des présentes prescriptions:

i) l'expression "demande de brevet" désigne un document dans lequel le déposant, ou son mandataire, demande la délivrance d'un brevet Ce document contient habituellement une description détaillée de l'invention, les revendications, et les dessins lorsque ceux-ci sont nécessaires à la compréhension de l'invention;

ii) l'expression "liste de séquences" désigne une partie distincte – la partie finale – de la description contenue dans la demande de brevet dactylographiée, qui divulgue de façon détaillée les séquences de nucléotides et/ou d'acides aminés, ainsi que d'autres in-



bart werden. Jede offenbarte Sequenz wird im Sequenzprotokoll gesondert ausgewiesen und mit einer eigenen Kennzahl bezeichnet. Die für Sequenzen und andere verfügbare Informationen zu verwendenden Symbole und das entsprechende Format richten sich nach den Empfehlungen im WIPO-Standard ST. 23

#### SEQUENZPROTOKOLL IN COMPUTERLESBARER FORM

5. Zusätzlich zu dem in der schriftlichen Patentanmeldung enthaltenen Sequenzprotokoll gemäß Nummer 4 ii) ist ein Exemplar des Sequenzprotokolls in computerlesbarer Form auf elektronischem Datenträger einzureichen

Das Sequenzprotokoll sollte die in Anlage 1 dieser Vorschriften genannten Datenelementrubriken umfassen und hat mindestens die in Anlage 3 genannten Datenelemente zu enthalten. Das Beispiel in Anlage 2 veranschaulicht, wie diese Datenelementrubriken zu verwenden sind

6. Die computerlesbare Fassung besteht aus einem auf Diskette aufgezeichneten ausdrückbaren Exemplar des Sequenzprotokolls. Sie ist so zu codieren und zu formatieren, daß anhand der Druckbefehle des Computer-/Betriebssystems der unter Nummer 13 dieser Vorschriften spezifizierten Konfiguration ein gedrucktes Exemplar des Sequenzprotokolls reproduziert werden kann

7. Das gesamte ausdrückbare Exemplar des Sequenzprotokolls sollte nach Möglichkeit in einer einzigen Datei auf einer einzigen Diskette enthalten sein. Die auf der Diskette aufgezeichnete Datei ist mittels IBM-Codetabelle (Code Page) 437 oder einer kompatiblen Codetabelle zu codieren. Als kompatibel gilt eine Codetabelle, die das römische Alphabet und die Ziffern denselben Hexadezimalpositionen zuordnet wie die genannte Codetabelle. Diese Codetabelle ist der bei modernen PCs übliche Standard

8. Die computerlesbare Fassung kann mit beliebigen Mitteln wie Textverarbeitungssystemen, Editoren für Nucleotid-/Aminosäuresequenzen, Computer-Editoren, spezieller Software wie PatentIn oder anderen kundenspezifischen Computerprogrammen erstellt werden, muß jedoch mit einem Computer-/Betriebssystem der unter Nummer 13 genannten Konfiguration lesbar sein und den Anforderungen gemäß den Nummern 6 und 7 dieser Vorschriften entsprechen

9. Die eingereichte Diskette ist mit einem Schreibschutz zu versehen.

10. Die Diskette ist mit einem fest angebrachten, von Hand in großen Druckbuchstaben oder mit Maschine beschrifteten Etikett zu versehen, das

appears separately in the Sequence Listing and is assigned a separate identifier number. Symbols and format used for sequences and other available information follow the recommendations given in WIPO Standard ST. 23.

#### SEQUENCE LISTING IN COMPUTER-READABLE FORM

5. A copy of the Sequence Listing referred to in paragraph 4(ii) above should also be submitted in computer-readable form on an electronic data carrier, in addition to the Sequence Listing contained in the typed patent applications.

The Sequence Listing should contain the Data Element Headings set out in Appendix 1 to these Rules and must at least contain those data elements specified in Appendix 3. The example in Appendix 2 illustrates the use of the Data Element Headings.

6. The computer-readable version comprises a printable copy of the Sequence Listing recorded on diskette. It should be encoded and formatted in such a way that a printed copy of the Sequence Listing may be reproduced using the print commands of the computer/operating-system configuration specified in paragraph 13 of these Rules.

7. The entire printable copy of the Sequence Listing should preferably be contained within one file on a single diskette. The file recorded on the diskette should be encoded using IBM Code Page 437 or a compatible code page. A compatible code page is one that assigns the Roman alphabet and numerals to the same hexadecimal positions as the specified code page. The specified code page is standard for modern personal computers.

8. The computer readable version may be produced by any means such as word-processors, nucleotide/amino acid sequence editors, computer editors, dedicated software such as PatentIn or other custom-made computer programs, however, it must be readable by one of the computer/operating-system configurations specified in paragraph 13 and conform to the specifications given in paragraphs 6 and 7 of these Rules.

9. The diskette submitted should be write-protected

10. The diskette should bear a permanently affixed label on which the following information should be handwritten in block capitals or typed: the

formations disponibles. Chaque séquence divulguée est indiquée séparément dans cette liste et porte un numéro d'identification distinct. Les symboles et le mode de présentation à utiliser pour les séquences et pour d'autres informations disponibles sont conformes aux recommandations données dans la norme ST.23 de l'OMPI.

#### LISTE DE SEQUENCES SOUS UNE FORME DECHIFFRABLE PAR ORDINATEUR

5. Une copie déchiffable par ordinateur d'une liste de séquences visée au paragraphe 4ii) ci-dessus doit être fournie en sus de la liste de séquences figurant dans la demande de brevet dactylographiée.

La liste de séquences devrait contenir les en-têtes d'éléments de données figurant à l'appendice 1 des présentes prescriptions et doit contenir, dans tous les cas, au moins celles énoncées à l'appendice 3. L'exemple donné à l'appendice 2 illustre l'utilisation des en-têtes d'éléments de données

6. Le support déchiffable par ordinateur doit comprendre une copie imprimable de la liste de séquences enregistrée sur disquette. Il doit être codé et formaté de telle sorte qu'une copie imprimée de cette liste de séquences puisse être recréée au moyen des commandes d'impression de l'une des configurations ordinateur/système d'exploitation indiquées au paragraphe 13 ci-dessous

7. La copie imprimable de la liste de séquences doit, de préférence, figurer toute entière dans un seul fichier sur une seule disquette. Le fichier enregistré sur la disquette doit être codé selon la page de code IBM (Code Page) 437 ou une page de code compatible. Une page de code compatible est une page de code qui attribue l'alphabet romain et les chiffres aux mêmes positions hexadécimales que les pages codées indiquées. La page de code indiquée constitue la norme applicable aux ordinateurs personnels modernes

8. La version déchiffable par ordinateur peut être créée par tout moyen, par exemple par des machines de traitement de texte, des éditeurs de séquences de nucléotides ou d'acides aminés, des éditeurs informatiques, du logiciel spécialisé tel que PatentIn ou d'autres programmes informatiques spécialisés; toutefois, il doit pouvoir être déchiffré par l'une des configurations ordinateur/système d'exploitation indiqués au paragraphe 13 et doit être conforme aux spécifications énoncées aux paragraphes 6 et 7 ci-dessus

9. Il y a lieu de protéger en écriture la disquette fournie

10. Sur la disquette doit être apposée une étiquette fixe portant les indications manuscrites (en caractères d'imprimerie) ou dactylographiées sui-

folgende Angaben enthält: das Format der Diskette, den Namen des Anmelders, die Bezeichnung der Erfindung, eine Bezugsnummer, den Zeitpunkt der Aufzeichnung der Daten auf der Diskette sowie Marke und Typenbezeichnung des Computers und des Betriebssystems, mit denen die Datei auf der Diskette erstellt wurde. Bei Verwendung einer anderen Codetabelle als der unter Nummer 7 genannten ist auch die Codetabelle anzugeben.

11 Können aus Platzgründen oder aus einem anderen Grund nicht alle diese Angaben auf dem Etikett der Diskette ausgewiesen werden, so sind auf diesem Etikett nur der Name des Anmelders, die Bezeichnung der Erfindung sowie eine Bezugsnummer anzugeben und die übrigen Angaben auf das Etikett eines ebenfalls mit dem Namen des Anmelders, der Bezeichnung der Erfindung und der Bezugsnummer ausgezeichneten Behältnisses für die Diskette zu übertragen.

12 Wird die Diskette erst nach dem Anmeldetag der betreffenden Anmeldung eingereicht, so sind auf den Etiketten auch der Anmeldetag und die Anmeldenummer anzugeben.

13 Wie unter Nummer 8 dargelegt, können zur Erstellung der computerlesbaren Fassung beliebige Mittel verwendet werden, sofern das Sequenzprotokoll auf der eingereichten Diskette mit einem der Computer-/Betriebssysteme der nachstehend unter a) und b) beschriebenen Konfigurationen gelesen werden kann bzw. dem dort spezifizierten Format entspricht

a) Computer: IBM PC/XT/AT, IBM PS/2 oder kompatible Geräte  
Betriebssystem: PC-DOS oder MS-DOS (Version 2.1 oder höher)

Zeilenende Wagenrücklauf plus Zeilenvorschub  
Seitenumbruch: Formularvorschub oder mehrere Zeilenenden

Dateiende: Ctrl-Z  
Datenträger:  
Diskette – 5,25 Zoll, 360 KB Speicherkapazität  
Diskette – 5,25 Zoll, 1,2 MB Speicherkapazität  
Diskette – 3,50 Zoll, 720 KB Speicherkapazität  
Diskette – 3,50 Zoll, 1,44 MB Speicherkapazität  
Druckbefehl: PRINT Dateiname.Erweiterung

b) Computer: Apple Macintosh  
Betriebssystem: Macintosh  
Macintosh-Dateityp: Text mit Zeilenende-Funktion  
Zeilenende: Vorgegeben durch Textdatei  
Seitenumbruch: Vorgegeben durch Textdatei

format of the diskette, the name of the applicant, the title of the invention, a reference number, the date on which the data was recorded on the diskette and the name and type of computer and operating system which generated the file on the diskette. The code page used should also be indicated if one other than that specified in paragraph 7 is used.

11. If not all this information can be shown on the label affixed to the diskette, for reasons of space or any other reason, the label should indicate only the name of the applicant, the title of the invention and a reference number, and the remaining information should be given on a label affixed to a container for the diskette; the container should likewise bear the name of the applicant, the title of the invention and the reference number

12 If the diskette is submitted after the date of filing of an application, the labels should also include the filing date of the application and the application number.

13. As set out in paragraph 8 above, any means may be used to produce the computer-readable version, as long as the Sequence Listing on the diskette submitted is readable on one of the computer/operating-system configurations or complies with the format specified in sub-paragraphs (a) and (b) below

(a) Computer: IBM PC/XT/AT, IBM PS/2 or compatible;  
Operating System PC-DOS or MS-DOS (Versions 2.1 or above);

Line Terminator: Carriage Return plus Line Feed;  
Pagination: Form Feed or Series of Line Terminators;

End-of-File: Ctrl-Z;  
Data carrier:  
Diskette – 5.25 inch, 360 Kb storage;

Diskette – 5.25 inch, 1.2 Mb storage;

Diskette – 3.50 inch, 720 Kb storage;

Diskette – 3 50 inch, 1 44 Mb storage;

Print command: PRINT filename extension

(b) Apple Macintosh  
Operating system: Macintosh;  
Macintosh File Type: Text with line termination;  
Line Terminator: Pre-defined by text file;  
Pagination: Pre-defined by text file;

vantes: le format de la disquette ainsi que le nom du déposant, le titre de l'invention, un numéro de référence, la date à laquelle les données ont été enregistrées sur la disquette, ainsi que la marque et le type de l'ordinateur et du système d'exploitation au moyen desquels le fichier a été créé sur cette disquette. La page de code utilisée doit aussi être précisée lorsqu'elle est différente de celle indiquée au paragraphe 7 ci-dessus.

11 Si, faute de place ou pour d'autres raisons, ces informations ne peuvent pas toutes être portées sur l'étiquette apposée sur la disquette, le nom du déposant, le titre de l'invention et un numéro de référence doivent y figurer, ces mêmes informations et les indications supplémentaires pouvant être portées sur une étiquette apposée sur l'emballage de la disquette en question.

12. Si la disquette est fournie après la date de dépôt de la demande, cette date et le numéro de la demande doivent figurer aussi sur les étiquettes.

13 Comme indiqué au paragraphe 8 ci-dessus, la liste déchiffrable par ordinateur peut être créée par tout moyen dès lors que la liste de séquences figurant sur la disquette fournie est déchiffrable par l'une des configurations ordinateur/système d'exploitation décrites aux alinéas a) et b) ci-dessous, ou qu'elle est conforme à la présentation indiquée dans ces mêmes alinéas.

a) Ordinateur: IBM PC/XT/AT, IBM PS/2 ou compatible;

Système d'exploitation: PC-DOS ou MS-DOS (version 2.1 ou version supérieure);

Caractère de fin de ligne: retour du chariot et changement de ligne;  
Pagination: caractère de changement de page ou série de caractères de fin de ligne;

Caractère de fin de fichier: Ctrl-Z;  
Supports:

disquette – 5,25 pouces, 360 Ko de mémoire;

disquette – 5,25 pouces, 1,2 Mo de mémoire;

disquette – 3,50 pouces, 720 Ko de mémoire ;

disquette – 3,50 pouces, 1,44 Mo de mémoire ;

Commande d'impression: PRINT nom de fichier extension ;

b) Ordinateur: Apple Macintosh;

Système d'exploitation: Macintosh;  
Type de fichier Macintosh: texte avec fonction de fin de ligne;

Caractère de fin de ligne: prédéfini par le fichier de type de texte;

Pagination: prédéfinie par le fichier de type de texte;

**Datenträger:**

Diskette – 3,50 Zoll, 400 KB Speicher-  
kapazität

Diskette – 3,50 Zoll, 800 KB Speicher-  
kapazität

Diskette – 3,50 Zoll, 1,4 MB Speicher-  
kapazität

Druckbefehl: Verwendung des Befehls  
"PRINT" aus jeder beliebigen Macin-  
tosh-Anwendung, die Textdateien ver-  
arbeitet, wie MacWrite oder TeachText

**Data carrier:**

Diskette – 3.50 inch, 400 Kb storage;

Diskette – 3.50 inch, 800 Kb storage;

Diskette – 3.50 inch, 1.4 Mb storage;

Print command: Use PRINT command  
from any Macintosh application that  
processes text files, such as MacWrite  
or TeachText.

**Supports:**

disquette – 3,50 pouces, 400 Ko de  
mémoire;

disquette – 3,50 pouces, 800 Ko de mé-  
moire;

disquette – 3.50 pouces, 1,4 Mo de  
mémoire;

Commande d'impression: utiliser la  
commande PRINT d'une application  
Macintosh traitant les fichiers de texte,  
par exemple MacWrite ou TeachText

**Anlage 1 zu den Vorschriften über die Einreichung von Sequenzprotokollen auf elektronischem Datenträger**

**Datenelementrubrik WIPO-ST. 23 Nummer**

ALLGEMEINE ANGABEN:<sup>1</sup>  
 ANMELDER:  
 NAME:  
 STRASSE:  
 ORT:  
 BUNDESLAND:  
 LAND:  
 POSTLEITZAHL:  
 TELEFON:  
 TELEFAX:  
 TELEX:  
 BEZEICHNUNG DER ERFINDUNG:  
 ANZAHL DER SEQUENZEN:  
 COMPUTERLESBARE FASSUNG:  
 DATENTRÄGER:  
 COMPUTER:  
 BETRIEBSSYSTEM:  
 SOFTWARE:  
 DATEN DER JETZIGEN ANMELDUNG:  
 ANMELDENUMMER:  
 DATEN DER VORANMELDUNG:  
 ANMELDENUMMER:  
 ANMELDETAG  
 KLASSIFIKATION:  
 ANGABEN ZU SEQ ID NO: 1  
 SEQUENZKENNZEICHEN: 1.i  
 LÄNGE 1.i.A  
 ART 1.i.B  
 STRANGFORM 1.i.C  
 TOPOLOGIE: 1.i.D  
 ART DES MOLEKÜLS: 1.ii  
 HYPOTHETISCH: 1.iii  
 ANTISENSE: 1.iv  
 ART DES FRAGMENTS: 1.v  
 URSPRÜNGLICHE HERKUNFT: 1.vi  
 ORGANISMUS 1.vi.A  
 STAMM 1.vi.B  
 INDIVIDUUM/ISOLAT: 1.vi.C  
 ENTWICKLUNGSSTADIUM: 1.vi.D  
 HAPLOTYP: 1.vi.E  
 GEWEBETYP: 1.vi.F  
 ZELLTYP 1.vi.G  
 ZELLINIE 1.vi.H  
 ORGANELLE: 1.vi.I  
 UNMITTELBARE HERKUNFT 1.vii  
 BIBLIOTHEK: 1.vii.A  
 CLON 1.vii.B  
 POSITION IM GENOM: 1.viii  
 CHROMOSOM/SEGMENT: 1.viii.A  
 KARTENPOSITION: 1.viii.B  
 EINHEITEN 1.viii.C

**Appendix 1 to the Rules for the Filing of Sequence Listings on Electronic Data Carriers**

**Data Element Heading WIPO ST.23 Point**

GENERAL INFORMATION:<sup>1</sup>  
 APPLICANT:  
 NAME:  
 STREET:  
 CITY:  
 STATE OR PROVINCE:  
 COUNTRY:  
 POSTAL CODE:  
 TELEPHONE:  
 FAX:  
 TELEX:  
 TITLE OF INVENTION:  
 NUMBER OF SEQUENCES:  
 COMPUTER-READABLE FORM  
 MEDIUM TYPE:  
 COMPUTER:  
 OPERATING SYSTEM:  
 SOFTWARE:  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 CLASSIFICATION:  
 INFORMATION FOR SEQ. ID. NO: 1  
 SEQUENCE CHARACTERISTICS: 1.i  
 LENGTH: 1.i.A  
 TYPE: 1.i.B  
 STRANDEDNESS: 1.i.C  
 TOPOLOGY: 1.i.D  
 MOLECULE TYPE: 1.ii  
 HYPOTHETICAL: 1.iii  
 ANTI-SENSE: 1.iv  
 FRAGMENT TYPE: 1.v  
 ORIGINAL SOURCE: 1.vi  
 ORGANISM: 1.vi.A  
 STRAIN: 1.vi.B  
 INDIVIDUAL/ISOLATE: 1.vi.C  
 DEVELOPMENT STAGE 1.vi.D  
 HAPLOTYPE: 1.vi.E  
 TISSUE TYPE: 1.vi.F  
 CELL TYPE: 1.vi.G  
 CELL LINE: 1.vi.H  
 ORGANELLE 1.vi.I  
 IMMEDIATE SOURCE: 1.vii  
 LIBRARY: 1.vii.A  
 CLONE: 1.vii.B  
 POSITION IN GENOME: 1.viii  
 CHROMOSOME/SEGMENT: 1.viii.A  
 MAP POSITION: 1.viii.B  
 UNITS: 1.viii.C

**Appendice 1 aux prescriptions relatives au dépôt des listes de séquences sur des supports électroniques de données**

**En-tête d'éléments de données ST.23 OMPi point**

INFORMATIONS GENERALES:<sup>1</sup>  
 DEPOSANT:  
 NOM:  
 RUE:  
 VILLE:  
 ETAT OU PROVINCE:  
 PAYS:  
 CODE POSTAL:  
 TELEPHONE:  
 TELEFAX  
 TELEX:  
 TITRE DE L'INVENTION:  
 NOMBRE DE SEQUENCES:  
 FORME DECHIFFRABLE PAR ORDINATEUR:  
 TYPE DE SUPPORT:  
 ORDINATEUR:  
 SYSTEME D'EXPLOITATION:  
 LOGICIEL:  
 DATE DE LA DEMANDE ACTUELLE  
 NUMERO DE LA DEMANDE:  
 DATE DE LA DEMANDE ANTERIEURE  
 NUMERO DE LA DEMANDE:  
 DATE DE DEPOT:  
 CLASSIFICATION:  
 INFORMATIONS POUR LA SEQ ID NO: 1  
 CARACTERISTIQUES DE LA SEQUENCE 1.i  
 LONGUEUR: 1.i.A  
 TYPE: 1.i.B  
 NOMBRE DE BRINS: 1.i.C  
 CONFIGURATION: 1.i.D  
 TYPE DE MOLECULE: 1.ii  
 HYPOTHETIQUE: 1.iii  
 ANTI-SENS: 1.iv  
 TYPE DE FRAGMENT: 1.v  
 ORIGINE 1.vi  
 ORGANISME: 1.vi.A  
 SOUCHE: 1.vi.B  
 INDIVIDU/ISOLE: 1.vi.C  
 STADE DE DEVELOPPEMENT: 1.vi.D  
 HAPLOTYPE: 1.vi.E  
 TYPE DE TISSU: 1.vi.F  
 TYPE DE CELLULE: 1.vi.G  
 LIGNEE CELLULAIRE: 1.vi.H  
 ORGANELLE: 1.vi.I  
 SOURCE IMMEDIATE 1.vii  
 BIBLIOTHEQUE: 1.vii.A  
 CLONE 1.vii.B  
 POSITION DANS LE GENOME 1.viii  
 CHROMOSOME/SEGMENT: 1.viii.A  
 POSITION SUR LA CARTE: 1.viii.B  
 UNITES: 1.viii.C

<sup>1</sup> Das Programm PatentIn enthält gegenwärtig einige Begriffe die von der Terminologie der vorliegenden Dokumente abweichen. So wird beispielsweise statt "DNA" das Kürzel "DNS" verwendet. Eine aktualisierte Version des Programms wird in Kürze zur Verfügung stehen.

<sup>1</sup> Le programme PatentIn contient actuellement quelques termes qui varient de la terminologie utilisée dans les présents documents. Par exemple "DNA" y figure pour "ADN". Une version mise à jour sera disponible prochainement.

---

MERKMAL:	1.ix	FEATURE:	1.ix	CARACTERISTIQUE	1.ix
NAME/SCHLÜSSEL:	1.ix.A	NAME/KEY:	1.ix.A	NOM/CLE:	1.ix.A
LAGE:	1.ix.B	LOCATION:	1.ix.B	EMPLACEMENT:	1.ix.B
ERMITTLUNGSMETHODE:	1.ix.C	IDENTIFICATION METHOD:	1.ix.C	METHODE D'IDENTIFICATION:	1.ix.C
SONSTIGE ANGABEN:	1.ix.D	OTHER INFORMATION:	1.ix.D	AUTRES RENSEIGNEMENTS:	1.ix.D
SEQUENZBESCHREIBUNG: SEQ ID		SEQUENCE DESCRIPTION: SEQ ID		DESCRIPTION DE LA SEQUENCE:	
NO:....:	1.xi	NO: ...:	1.xi	SEQ ID NO: ...:	1.xi

**Anlage 2 zu den Vorschriften über die Einreichung von Sequenzprotokollen auf elektronischem Datenträger**

**Beispiel  
SEQUENZPROTOKOLL**

**ALLGEMEINE ANGABEN:**

**ANMELDER:**

NAME: Anmelder GmbH  
STRASSE: Erhardtstraße 27  
ORT: München  
LAND: Deutschland  
POSTLEITZAHL: W-8000  
TELEFON: 089-23990  
TELEFAX: 089-23994465  
TELEX: 523656 epmu d

**BEZEICHNUNG DER ERFINDUNG:**

Beispiel eines Sequenzprotokolls

**ANZAHL DER SEQUENZEN: 2**

**COMPUTERLESBARE FASSUNG:**

DATENTRÄGER: Diskette  
COMPUTER: IBM-PC-kompatibel  
BETRIEBSSYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release 1.0, Version 1.25 (EPA)

**ANGABEN ZU SEQ ID NO: 1:**

**SEQUENZKENNZEICHEN:**

LÄNGE: 954 Basenpaare  
ART: Nucleinsäure  
STRANGFORM: Einzelstrang  
TOPOLOGIE: linear

**HYPOTHETISCH: ja**

**ANTISENSE: nein**

URSPRÜNGLICHE HERKUNFT:  
ORGANISMUS: Paramecium sp  
INDIVIDUUM/ISOLAT: XYZ2  
ZELLTYP: einzelliger Organismus

**UNMITTELBARE HERKUNFT:**

BIBLIOTHEK: genomisch  
CLON: Para-XYZ2/36

**MERKMAL**

NAME/SCHLÜSSEL: CDS (codierende Sequenz)\*

LAGE: join (Verbindung) (275..373, 448..498, 679..774)

**MERKMAL:**

NAME/SCHLÜSSEL: mat\_peptide (reifes Peptid)

LAGE: join (Verbindung) (451..498, 679..774)

**MERKMAL:**

NAME/SCHLÜSSEL: sig\_peptide (Signalpeptid)

LAGE: 275..448

**Appendix 2 to the Rules for the Filing of Sequence Listings on Electronic Data Carriers**

**Example  
SEQUENCE LISTING**

**GENERAL INFORMATION:**

**APPLICANT:**

NAME: Applicant Plc  
STREET: Erhardtstrasse 27  
CITY: Munich  
COUNTRY: Germany  
POSTAL CODE: W-8000  
TELEPHONE: 089-23990  
TELEFAX: 089-23994465  
TELEX: 523656 epmu d

TITLE OF INVENTION: Example of a Sequence Listing

**NUMBER OF SEQUENCES: 2**

**COMPUTER READABLE FORM:**

MEDIUM TYPE: Diskette  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release # 1.0, Version 1.25 (EPO)

**INFORMATION FOR SEQ ID NO: 1:**

**SEQUENCE CHARACTERISTICS:**

LENGTH: 954 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

**HYPOTHETICAL: yes**

**ANTI-SENSE: no**

ORIGINAL SOURCE:  
ORGANISM: Paramecium sp  
INDIVIDUAL/ISOLATE: XYZ2  
CELL TYPE: unicellular organism

**IMMEDIATE SOURCE:**

LIBRARY: genomic  
CLONE: Para-XYZ2/36

**FEATURE:**

NAME/KEY: CDS

LOCATION: join (275..373, 448..498, 679..774)

**FEATURE:**

NAME/KEY: mat\_peptide

LOCATION: join(451 498, 679..774)

**FEATURE**

NAME/KEY: sig\_peptide

LOCATION: 275..448

**Appendice 2 aux prescriptions relatives au dépôt des listes de séquences sur des supports électroniques de données**

**Exemple  
LISTE DE SEQUENCES**

**INFORMATIONS GENERALES:**

**DEPOSANT:**

NOM: Déposant SA  
RUE: Erhardtstrasse 27  
VILLE: Munich  
PAYS: Allemagne  
CODE POSTAL: W-8000  
TELEPHONE: 089-23990  
TELEFAX: 089-23994465  
TELEX: 523656 epmu d

TITRE DE L'INVENTION: Exemple d'une Liste de Séquences

**NOMBRE DE SEQUENCES: 2**

**FORME LISIBLE PAR ORDINATEUR:**

TYPE DE SUPPORT: Disquette  
ORDINATEUR: IBM PC compatible  
SYSTEME D'EXPLOITATION: PC-DOS/MS-DOS

LOGICIEL: PatentIn Release 1.0, Version 1.25 (OEB)

**INFORMATIONS POUR LA SEQ ID NO: 1:**

**CARACTERISTIQUES DE LA SEQUENCE:**

LONGUEUR: 954 paires de bases  
TYPE: acide nucléique  
NOMBRE DE BRINS: simple  
CONFIGURATION: linéaire

**HYPOTHETIQUE: oui**

**ANTI-SENS: non**

ORIGINE:  
ORGANISME: Paramecium sp  
INDIVIDU/SOLE: XYZ2  
TYPE DE CELLULE: unicellular organism

**SOURCE IMMEDIATE:**

BIBLIOTHEQUE: genomic  
CLONE: Para-XYZ2/36

**CARACTERISTIQUE ADDITIONNELLE:  
NOM/CLE: CDS (séquence codante)\***

EMPLACEMENT: join(275..373, 448..498, 679..774)

**CARACTERISTIQUE ADDITIONNELLE:  
NOM/CLE: mat\_peptide (peptide mature)**

EMPLACEMENT: join(451 498, 679..774)

**CARACTERISTIQUE ADDITIONNELLE:  
NOM/CLE: sig\_peptide (peptide signal)**

EMPLACEMENT: 275..448

\* Die deutschen Begriffe in Klammern dienen nur zur Erläuterung und sollten im Sequenzprotokoll nicht ausgewiesen werden

\* Les termes en français figurant entre parenthèses sont indiqués pour clarification. Il n'y a pas lieu de les faire apparaître dans la liste de séquences

SEQUENZBESCHREIBUNG: SEQ ID NO 1:	SEQUENCE DESCRIPTION: SEQ ID NO: 1:	DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:	
ATCGGGATAG TACTGGTCAA GACCGGTGGA CACCGGTTAA CCCC GGTTAA GTACCGGTTA			60
TAGGCCATTT CAGGCCAAAT GTGCCCAACT ACGCCAATTG TTTTGCCAAC GGCCAACGTT			120
ACGTTCGTAC GCACGTATGT ACCTAGGTAC TTACGGACGT GACTACGGAC ACTTCCGTAC			180
GTACGTACGT TTACGTACCC ATCCCAACGT AACCACAGTG TGGTCGCAGT GTCCCAGTGT			240
ACACAGACTG CCAGACATTC TTCACAGACA CCCC ATG ACA CCA CCT GAA CGT		Met Thr Pro Pro Glu Arg	292
		-34 -30	
CTC TTC CTC CCA AGG GTG TGT GGC ACC ACC CTA CAC CTC CTC CTT CTG		Leu Phe Leu Pro Arg Val Cys Gly Thr Thr Leu His Leu Leu Leu Leu	340
		-25 -20 -15	
GGG CTG CTG CTG GTT CTG CTG CCT GGG GCC CAT GTGAGGCAGC AGGAGAATGG		Gly Leu Leu Leu Val Leu Leu Pro Gly Ala His	393
		-10 -5	
GGTGGCTCAG CCAAACCTTG AGCCCTAGAG CCCCCCTCAA CTCTGTTCTC CTAG GGG		Gly	450
		-1	
CTC ATG CAT CTT GCC CAC AGC AAC CTC AAA CCT GCT GCT CAC CTC ATT		Leu Met His Leu Ala His Ser Asn Leu Lys Pro Ala Ala His Leu Ile	498
		1 5 10 15	
GTAACATCC ACCTGACCTC CCAGACATGT CCCACCAGC TCTCCTCCTA CCCCTGCCTC			558
AGGAACCCAA GCATCCACCC CTCTCCCCCA ACTTCCCCCA CGCTAAAAAA AACAGAGGGA			618
GCCACTCCT ATGCCTCCCC CTGCCATCCC CCAGGAACTC AGTTGTTTCAG TGCCCACTTC			678
TAC CCC AGC AAG CAG AAC TCA CTG CTC TGG AGA GCA AAC ACG GAC CGT		Tyr Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg Ala Asn Thr Asp Arg	726
		20 25 30	
GCC TTC CTC CAG GAT GGT TTC TCC TTG AGC AAC AAT TCT CTC CTG GTC		Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn Asn Ser Leu Leu Val	774
		35 40 45	
TAGAAAAAAT AATTGATTTT AAGACCTTCT CCCCATTTCTG CCTCCATTCT GACCATTTCA			834
GGGGTCGTCA CCACCTCTCC TTTGGCCATT CCAACAGCTC AAGTCTTCCC TGATCAAGTC			894
ACCGGAGCTT TCAAAGAAGG AATTCTAGGC ATCCCAGGGG ACCCACACCT CCCTGAACCA			954

ANGABEN ZU SEQ ID NO: 2:	INFORMATION FOR SEQ ID NO: 2	INFORMATION POUR LA SEQ ID NO: 2:
SEQUENZKENNZEICHEN:	SEQUENCE CHARACTERISTICS:	CARACTERISTIQUES DE LA SEQUENCE:
LÄNGE: 82 Aminosäuren	LENGTH: 82 amino acids	LONGUEUR: 82 acides aminés
ART: Aminosäure	TYPE: amino acid	TYPE: acide aminé
TOPOLOGIE: linear	TOPOLOGY: linear	CONFIGURATION: linéaire
ART DES MOLEKÜLS: Protein	MOLECULE TYPE: protein	TYPE DE MOLECULE: protéine
SEQUENZBESCHREIBUNG: SEQ ID NO: 2:	SEQUENCE DESCRIPTION: SEQ ID NO: 2:	DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

```

Met Thr Pro Pro Glu Arg Leu Phe Leu Pro Arg Val Cys Gly Thr Thr
-34                -30                -25                -20

Leu His Leu Leu Leu Leu Gly Leu Leu Leu Val Leu Leu Pro Gly Ala
          -15                -10                -5

His Gly Leu Met His Leu Ala His Ser Asn Leu Lys Pro Ala Ala His
          1                5                10

Leu Ile Tyr Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg Ala Asn Thr
15                20                25                30

Asp Arg Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn Asn Ser Leu
          35                40                45

Leu Val

```



### Anlage 3 zu den Vorschriften über die Einreichung von Sequenzprotokollen auf elektronischem Datenträger

#### VOM EPA VORGESCHRIEBENE ANGABEN<sup>1</sup>

Das Sequenzprotokoll hat zusätzlich zu der eigentlichen Nucleotid- und/oder Aminosäuresequenz die nachstehenden Datenelemente zu enthalten, die der Sequenz unmittelbar voranzustellen sind. Die Datenelemente im Sequenzprotokoll sind in der nachstehend dargelegten Reihenfolge und Form wiederzugeben, wobei allerdings die nach den Rubrik titles in Klammern stehenden Erläuterungen wegzulassen sind. Jedes Datenelement hat, wie nachstehend aufgezeigt, in einer neuen Zeile zu beginnen, die wahlweise mit den in Klammern stehenden Nummern/Ziffern/Buchstaben numeriert werden kann, wobei auf den Rubrik title in Großbuchstaben ein Doppelpunkt und danach das Datenelement selbst folgt

#### (1) ALLGEMEINE ANGABEN<sup>2</sup>

##### (i) ANMELDER: (mehrmals)

(A) NAME:

(B) STRASSE

(C) ORT:

(D) BUNDESLAND: (falls zutreffend)

(E) LAND:

(F) POSTLEITZAHL:

##### (ii) BEZEICHNUNG DER ERFINDUNG:

##### (iii) ANZAHL DER SEQUENZEN:

##### (iv) COMPUTERLESBARE FASSUNG<sup>2</sup>:

(A) DATENTRÄGER:

### Appendix 3 to the Rules for the Filing of Sequence Listings on Electronic Data Carriers

#### INFORMATION REQUIRED BY THE EPO<sup>1</sup>

The Sequence Listing must include, in addition to and immediately preceding the actual nucleotide and/or amino acid sequence, the items of information listed below. The information in the Sequence Listing should be presented in the order and form shown below, except that the explanations given in brackets after the Data Element Headings should be omitted. As shown below, each item of information should begin on a new line with the number/numeral/letter in brackets followed by the Data Element Heading in upper case characters, a colon, and then the actual item of information.

#### (1) GENERAL INFORMATION<sup>2</sup>

##### (i) APPLICANT (rep)

(A) NAME:

(B) STREET:

(C) CITY:

(D) STATE OR PROVINCE: (if applicable)

(E) COUNTRY:

(F) POSTAL CODE:

##### (ii) TITLE OF INVENTION

##### (iii) NUMBER OF SEQUENCES

##### (iv) COMPUTER-READABLE FORM<sup>2</sup>:

(A) MEDIUM TYPE:

### Appendice 3 aux prescriptions relatives au dépôt des listes de séquences sur des supports électroniques de données

#### RENSEIGNEMENTS EXIGES PAR L'OEB<sup>1</sup>

La liste de séquences inclut, outre la séquence de nucléotides et/ou d'acides aminés réelle, les renseignements suivants qui devront figurer immédiatement avant ladite séquence. L'ordre et la présentation des informations contenues dans la liste de séquences doit être conforme à la disposition indiquée ci-dessous, à l'exception des explications entre parenthèses suivant les en-têtes d'éléments de données qui doivent être omises. Chaque renseignement doit commencer sur une nouvelle ligne, précédé le cas échéant du numéro/de la lettre entre parenthèses, l'en-tête d'éléments de données en lettres majuscules étant suivi d'un double point et du renseignement fourni, comme indiqué ci-après.

#### 1) INFORMATIONS GENERALES<sup>2</sup>

##### i) DEPOSANT: (rep)

A) NOM:

B) RUE:

C) VILLE:

D) ETAT OU PROVINCE (le cas échéant):

E) PAYS:

F) CODE POSTAL:

##### ii) TITRE DE L'INVENTION:

##### iii) NOMBRE DE SEQUENCES:

##### iv) FORME DECHIFFRABLE PAR ORDINATEUR<sup>2</sup>:

A) TYPE DE SUPPORT:

<sup>1</sup> S Nummer 3 der Vorschriften über die Einreichung von Sequenzprotokollen auf elektronischem Datenträger

<sup>2</sup> Definitionen

ANMELDER NAME

Name des Anmelders (je nach Fall Firma Institution oder Erfinder)

STRASSE

Straße aus der Anschrift des Anmelders wie in Feld 8 des Antrags auf Erteilung eines europäischen Patents angegeben

ORT

Ort aus der Anschrift des Anmelders wie in Feld 8 des Antrags auf Erteilung eines europäischen Patents angegeben

BUNDESLAND

Bundesland oder -staat aus der Anschrift des Anmelders wie in Feld 8 des Antrags auf Erteilung eines europäischen Patents angegeben (nur dann zwingend vorgeschrieben, wenn zutreffend)

LAND

Land aus der Anschrift des Anmelders, wie in Feld 8 des Antrags auf Erteilung eines europäischen Patents angegeben

POSTLEITZAHL

Postleitzahl des Anmelders, wie in Feld 8 des Antrags auf Erteilung eines europäischen Patents angegeben ("keine", wenn nicht zutreffend)

DATENTRÄGER

Art des eingereichten elektronischen Datenträgers (Diskette)

<sup>1</sup> See paragraph 3 of the Rules for the Filing of Sequence Listings on Electronic Data Carriers

<sup>2</sup> Definitions

APPLICANT NAME

Name of applicant (company, institution or inventor as appropriate)

STREET

Street taken from applicant's address as indicated in Section 8 of the Request for Grant of a European Patent

CITY

City taken from applicant's address as indicated in Section 8 of the Request for Grant of a European Patent

STATE OR PROVINCE

State or province taken from applicant's address as indicated in Section 8 of the Request for Grant of a European Patent (only mandatory if applicable)

COUNTRY

Country taken from applicant's address as indicated in Section 8 of the Request for Grant of a European Patent

POSTAL CODE

Applicant's postal code as indicated in Section 8 of the Request for Grant of a European Patent ("none", if not applicable)

MEDIUM TYPE

Type of electronic data carrier submitted (diskette)

<sup>1</sup> Cf. paragraphe 3 des prescriptions relatives au dépôt des listes de séquences sur des supports électroniques de données

<sup>2</sup> Définitions.

NOM DU DEMANDEUR

Nom du demandeur (société, institution ou inventeur, selon le cas)

RUE

Adresse du demandeur (rue), comme indiqué à la rubrique 8 du formulaire de requête en délivrance d'un brevet européen

VILLE

Adresse du demandeur (ville), comme indiqué à la rubrique 8 du formulaire de requête en délivrance d'un brevet européen

ETAT OU PROVINCE

Adresse du demandeur (Etat ou province), comme indiqué à la rubrique 8 du formulaire de requête en délivrance d'un brevet européen (obligatoire, uniquement le cas échéant)

PAYS

Adresse du demandeur (pays), comme indiqué à la rubrique 8 du formulaire de requête en délivrance d'un brevet européen

CODE POSTAL

Code postal du demandeur, comme indiqué à la rubrique 8 du formulaire de requête en délivrance d'un brevet européen (néant, le cas échéant)

TYPE DE SUPPORT

Type de support électronique de données fourni (disquette)

- (B) COMPUTER:  
 (C) BETRIEBSSYSTEM:  
 (D) SOFTWARE:  
 (v) DATEN DER JETZIGEN ANMELDUNG: (falls bekannt)  
 (A) ANMELDENUMMER<sup>2</sup>:  
 (2) ANGABEN ZU SEQ ID NO: X: (mehrmals)  
 (i) SEQUENZKENNZEICHEN:  
 (A) LÄNGE:  
 (B) ART:  
 (C) STRANGFORM:  
 (D) TOPOLOGIE:  
 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: X:

*Anmerkung:* Der Hinweis "mehrmals" bedeutet, daß es zu einer Rubrik mehrere Angaben geben und diese Rubrik daher im Sequenzprotokoll mehrmals vorkommen kann

- (B) COMPUTER:  
 (C) OPERATING SYSTEM:  
 (D) SOFTWARE:  
 (v) CURRENT APPLICATION DATA: (if available)  
 (A) APPLICATION NUMBER<sup>2</sup>:  
 (2) INFORMATION FOR SEQ ID NO: X: (rep)  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH:  
 (B) TYPE:  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY:  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: X:

*Note:* Those items marked "rep" may have multiple responses and may therefore appear several times in the "Sequence Listing".

- B) ORDINATEUR:  
 C) SYSTEME D'EXPLOITATION:  
 D) LOGICIEL:  
 (v) DONNEES RELATIVES A LA DEMANDE EN CAUSE (si disponibles):  
 A) NUMERO DE LA DEMANDE<sup>2</sup>:  
 2) INFORMATION POUR LA SEQUENCE ID NO: X: (rep)  
 i) CARACTERISTIQUES DE LA SEQUENCE:  
 A) LONGUEUR:  
 B) TYPE:  
 C) NOMBRE DE BRINS:  
 D) CONFIGURATION:  
 xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: X:

*Note:* Les points suivis de l'indication "rep" sont susceptibles de recevoir plusieurs réponses et, en tant que tel, le point concerné peut se trouver répété dans la "liste de séquences".

<sup>2</sup> Definitionen (Forts)

COMPUTER  
 Art des Computers mit dem die eingereichte Diskette erstellt wurde  
 BETRIEBSSYSTEM  
 Art des eingesetzten Betriebssystems  
 SOFTWARE  
 Art der Software, mit der die computerlesbare Fassung erstellt wurde  
 ANMELDENUMMER  
 EP-Anmeldenummer einschließlich des Codes EP (Beispiel EP 92000000) oder PCT-Aktenzeichen, bestehend aus den Buchstaben PCT, Schrägstrich, einem zweibuchstabigen Code für das Anmeldeamt, einer zweistelligen Jahreszahl, Schrägstrich und einer fünfstelligen Nummer (Beispiel PCT/EP/92/00000) (nur dann zwingend vorgeschrieben, wenn bekannt)

<sup>2</sup> Definitions (contd.)

COMPUTER  
 Type of computer used to produce diskette submitted  
 OPERATING SYSTEM  
 Type of operating system used  
 SOFTWARE  
 Type of software used to produce the computer-readable version  
 APPLICATION N°  
 EP application number including EP code (example: EP 92000000) or PCT application number comprising the letters PCT, a slash, a two-letter code indicating the receiving Office, a two digit indication of the year, a slash and a five-digit number (example PCT/EP/92/00000) (only mandatory if known)

<sup>2</sup> Définitions (cont.)

ORDINATEUR  
 Type d'ordinateur utilisé avec la disquette fournie  
 SYSTEME D'EXPLOITATION  
 Type de système d'exploitation utilisé  
 LOGICIEL  
 Type de logiciel utilisé pour créer la forme déchiffable par ordinateur  
 N° DE LA DEMANDE  
 Numéro de la demande de brevet européen, y compris le code EP (exemple EP 92000000), ou numéro de la demande PCT, incluant les lettres PCT, un trait transversal, un code à deux lettres indiquant l'office récepteur, deux chiffres indiquant l'année, un trait transversal et un numéro à cinq chiffres (exemple PCT/EP/92/00000) (obligatoire uniquement si disponible)