



# **WIPO-Standard ST.26** **EINFÜHRUNG**

Webinar-Schulung

# Unsere Themen für heute

- Was ist ein Sequenzprotokoll?
- INSDC-Sequenzdatenbanken
- Warum ein neuer Standard?
- Vorteile von WIPO ST.26
- ST.25 im Vergleich zu ST.26 – wo liegen die Unterschiede?
- Übergang zu ST.26
- XML-Grundlagen
- ST.26 Grundlagen - Teile eines Sequenzprotokolls
- Inhalt von WIPO ST.26
- Einführung in WIPO Sequence

# Was ist ein Sequenzprotokoll?

Ein Sequenzprotokoll...

- enthält Nukleotid- und/oder Aminosäuresequenzen, die in einer Patentanmeldung offenbart sind, und ist Teil der Beschreibung
- enthält beschreibende Informationen über jede Sequenz, die als Annotationen bekannt sind
- entspricht den Anforderungen des maßgeblichen WIPO-Standards (ST.25 oder ST.26)
- ermöglicht die Durchsuchbarkeit der Sequenzdaten einer Erfindung:
  - innerhalb eines Amtes für geistiges Eigentum
  - in öffentlich zugänglichen Datenbanken (INSDC-Datenbanken)

# Hintergrund - INSDC

- INSDC: Internationale Nukleotidsequenz-Datenbank-Zusammenarbeit:
  - DDBJ: DNA-Datenbank von Japan
  - EMBL-EBI: Europäisches Institut für Bioinformatik
  - NCBI: Nationales Informationszentrum für Biotechnologie (GenBank)
  
- Zu den Ämtern für geistiges Eigentum, die Sequenzdaten von veröffentlichten/erteilten Anmeldungen an INSDC-Datenbanken übermitteln, gehören:
  - Europäisches Patentamt
  - Japanisches Patentamt
  - Koreanisches Amt für geistiges Eigentum
  - Patent- und Markenamt der Vereinigten Staaten
  
- INSDC-Datenbanken sind öffentlich durchsuchbar

# Warum ein neuer Standard?

- Derzeit werden Sequenzprotokolle gemäß WIPO ST.25 eingereicht.

Allerdings...

- ist das ST.25-Format nicht mit den Anforderungen der INSDC konform, so dass Daten bei der Eingabe in öffentliche Datenbanken verloren gehen
- sind die ST.25-Regeln nicht eindeutig und Ämter für geistiges Eigentum weltweit interpretieren und setzen die Regeln unterschiedlich um
- werden heute übliche Sequenztypen von den ST.25-Regeln nicht erfasst (Nukleotid-Analoga, D-Aminosäuren, verzweigte Sequenzen) und sind daher in durchsuchbaren Datenbanken nicht vorhanden
- sind die Daten unstrukturiert - das ST.25-Format ist für automatisierte Validierung und Datenaustausch schwer zu verwenden

# Vorteile von WIPO ST.26 (1)

- Akzeptanz eines einzigen Sequenzprotokolls weltweit\*
- Der Standard dient als Anleitung, um eine Übereinstimmung zwischen den Ämtern für geistiges Eigentum bei der Anwendung der Sequenzregeln sicherzustellen
- Klarstellung, welche Sequenzoffenbarungen in einem Sequenzprotokoll enthalten sein müssen oder dürfen und wie diese Sequenzen dargestellt werden müssen
- Verbesserte Einreichungsqualität aufgrund der Struktur von XML-Sequenzprotokollen
- Erhöhte Automatisierung der Datenvalidierung und vereinheitlichte Verarbeitung durch Ämter für geistiges Eigentum

\* mit Ausnahme der erforderlichen Übersetzungen von sprachabhängigen Freitext-Qualifiern in die Sprache der Einreichung für gewisse Ämter für geistiges Eigentum, die Ersatzsequenzprotokolle fordern können

# Vorteile von WIPO ST.26 (2)

- Datenkompatibilität mit den Anforderungen von INSDC-Datenbankanbietern - DDBJ, EBI und NCBI – Sequenzannotationen (Merkmalschlüssel und Qualifier) werden in öffentlich durchsuchbare Datenbanken aufgenommen
- Standardisierung von:
  - Merkmalsannotationen
  - Merkmallagen
  - Qualifiern und Qualifierwerten
  - Darstellung von Sequenzvarianten
- Anforderung der Aufnahme zusätzlicher Sequenztypen (Nukleotid-Analoga, D-Aminosäuren, verzweigte Sequenzen) bedeutet, dass mehr Sequenzdaten durchsuchbar sein werden

# WIPO ST.25 im Vergleich zu ST.26

ST.25	ST.26
ASCII .txt mit numerischen Kennzahlen	XML mit Elementen und Attributen
<u>Muss nicht</u> enthalten: <ul style="list-style-type: none"> <li>- D-Aminosäuren</li> <li>- Lineare Abschnitte von verzweigten Sequenzen</li> <li>- Nukleotid-Analoga</li> </ul>	<u>Muss</u> enthalten: <ul style="list-style-type: none"> <li>- D-Aminosäuren</li> <li>- Lineare Abschnitte von verzweigten Sequenzen</li> <li>- Nukleotid-Analoga</li> </ul>
Annotation zu den Sequenzen: <ul style="list-style-type: none"> <li>- Nur Merkmalschlüssel</li> </ul>	Annotation zu den Sequenzen: <ul style="list-style-type: none"> <li>- Merkmalschlüssel und Qualifier</li> </ul>
<u>Darf</u> Sequenzen enthalten: <ul style="list-style-type: none"> <li>- &lt; 10 genau definierte Nukleotide</li> <li>- &lt; 4 genau definierte Aminosäuren</li> </ul>	<u>Verbotene</u> Sequenzen: <ul style="list-style-type: none"> <li>- &lt; 10 genau definierte Nukleotide</li> <li>- &lt; 4 genau definierte Aminosäuren</li> </ul>

# WIPO ST.25 im Vergleich zu ST.26

## ...für Allgemeine Informationen

ST.25	ST.26
ALLE Informationen zu Prioritätsanmeldungen können enthalten sein	NUR die früheste Prioritätsanmeldung darf enthalten sein
ALLE Anmelder- und Erfindernamen können enthalten sein	Es darf NUR ein Anmelder UND optional EIN Erfinder enthalten sein
Eine Bezeichnung der Erfindung zulässig	Mehrere Bezeichnungen der Erfindung zulässig, jede in einer anderen Sprache
Anmelder-/Erfindernamen und Bezeichnung der Erfindung müssen in lateinischer Grundschrift verfasst sein	Anmelder-/Erfindernamen können unter Verwendung jedes gültigen Unicode-Zeichens zusammen mit einer Übersetzung in lateinische Grundschrift oder Transliteration enthalten sein

# WIPO ST.25 im Vergleich zu ST.26

## ...für Sequenzdaten (1)

ST.25	ST.26
Nur als DNA, RNA oder PRT identifizierte Sequenzen	Als DNA, RNA oder AA identifizierte Sequenzen, zusammen mit einem obligatorischen mol_type-Qualifier zur weiteren Beschreibung des Moleküls
Namen des Organismus: <ul style="list-style-type: none"> <li>- Lateinische Gattung/Art</li> <li>- Name des Virus</li> <li>- „künstliche Sequenz“</li> <li>- „unbekannt“</li> </ul>	Namen des Organismus: <ul style="list-style-type: none"> <li>- Lateinische Gattung/Art</li> <li>- Name des Virus</li> <li>- „synthetisches Konstrukt“</li> <li>- „nicht identifiziert“</li> </ul>
„u“ steht für Uracil in Nukleotidsequenzen	„t“ steht für Uracil in RNA-Sequenzen und Thymin in DNA-Sequenzen
Aminosäuresequenzen, dargestellt durch Abkürzungen mit drei Buchstaben	Aminosäuresequenzen, dargestellt durch Abkürzungen mit einem Buchstaben

# WIPO ST.25 im Vergleich zu ST.26

## ...für Sequenzdaten (2)

ST.25	ST.26
„n“- und „Xaa“-Variablen müssen eine Definition in einem Merkmal haben	Ein Standardwert wird angenommen für „n“- und „X“-Variablen ohne Definition
Format der Merkmallage nicht eindeutig definiert	Streng definierte Formate für Merkmallagen; erlaubt die Verwendung von „<“ und „>“ in allen Sequenztypen sowie „^“, „Verbindung“, „Reihenfolge“ und „Ergänzung“ in Nukleotidsequenzen
„Gemischtmodus“-Sequenzen zulässig - Nukleotidsequenz mit Aminosäuretranslation nachstehend gezeigt	KEIN „Gemischtmodus“; Nukleotidtranslationen sind nur in „Translations“-Qualifiern enthalten

# Übergang zu WIPO ST.26

- CWS/5 hat sich auf ein Übergangsdatum am **1. Januar 2022** geeinigt, das als das ‚**Big-Bang**‘-Datum bezeichnet wird
- **Alle** Ämter für geistiges Eigentum (IPO) werden gleichzeitig auf internationaler (PCT), nationaler und regionaler Ebene umstellen
- Der internationale Anmeldetag (IFD) wird der Stichtag sein, der bestimmt, ob eine Anmeldung unter die Sequenzregel ST.25 oder ST.26 fällt, NICHT das Prioritätsdatum
  - Anmerkung: ST.25 bleibt für Anmeldungen mit einem Anmeldetag vor dem 1. Januar 2022 wirksam

# WIPO ST.26

## ... was muss enthalten sein?

- Nukleotidsequenzen:
  - 10 oder mehr „genau definierte“ und „aufgezählte“\* Reste
  - enthalten Sequenzen mit Nukleotid-Analoga wie Peptid-Nukleinsäuren (PNA) und Glykol-Nukleinsäuren (GNA)
- Aminosäuresequenzen:
  - 4 oder mehr „genau definierte“ und „aufgezählte“ Reste
  - enthalten Sequenzen mit D-Aminosäuren
  - lineare Regionen von verzweigten Sequenzen müssen in ein Sequenzprotokoll aufgenommen werden

# WIPO ST.26

## ... was muss enthalten sein?

- Was ist ein „genau definiertes“ Nukleotid oder eine „genau definierte“ Aminosäure?
  - „genau definiert“ bedeutet jegliches Nukleotid, das nicht durch das Symbol „n“ dargestellt wird, und jegliche Aminosäure, die nicht durch das Symbol „X“ dargestellt wird, aufgeführt in Anlage I. (WIPO-Standard ST.26, pgh. 3(m))
  - nur „genau definierte“ Reste zählen für die Mindestlängenanforderung:
    - 10 oder mehr genau definierte Nukleotide; oder,
    - 4 oder mehr genau definierte Aminosäuren

5'- anctggcaan – 3' nur 8 genau definierte Nukleotide; darf nicht in einem Sequenzprotokoll enthalten sein

5'- agctggcaat – 3' zehn genau definierte Nukleotide; muss in einem Sequenzprotokoll enthalten sein

# WIPO ST.26: XML-Dokument

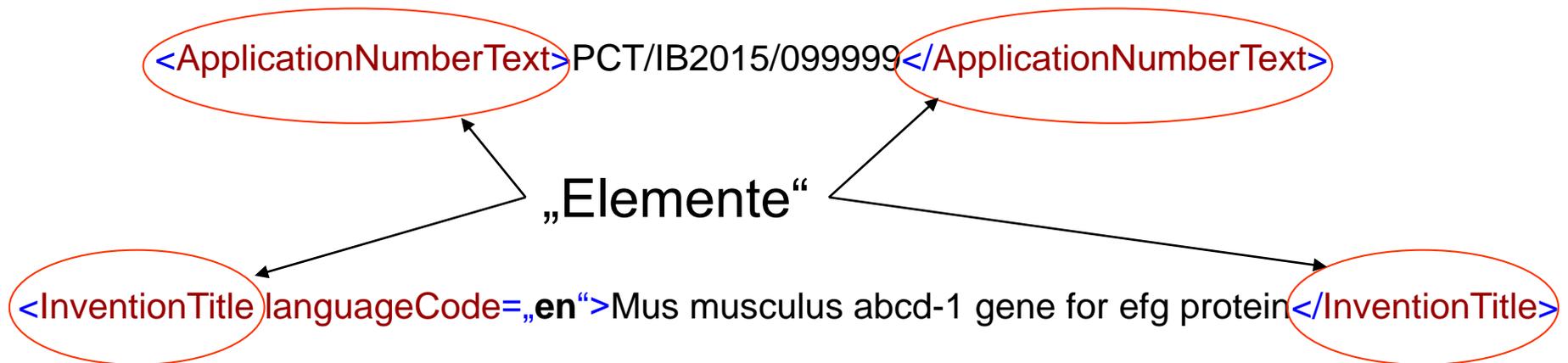
- Muss als eine Datei im XML-1.0-Format bereitgestellt werden
- Muss gegen WIPO ST.26 DTD (Anlage II) und aus dem Inhalt des Standards abgeleitete Geschäftsregeln zu validieren sein
- Muss unter Verwendung von Unicode UTF-8 kodiert sein
- Struktur des Sequenzprotokolls:
  - XML-Deklaration
  - Dokumententyp(DOCTYPE)-Deklaration
  - Wurzelelement
    - Teil mit allgemeinen Informationen
    - Teil mit Sequenzdaten

# Grundlagen von XML

- XML = eXtensible Markup Language
- Informationen werden mit beschreibenden Elementen und Attributen ‚getaggt‘
- Standardisierte Methode zum Datenaustausch, die sowohl von Menschen als auch von Maschinen lesbar ist
- DTD = Document Type Definition - definiert die Struktur und die gesetzmäßigen Elemente und Attribute eines XML-Dokuments

# Grundlagen von XML

## Elemente, Attribute und Werte (1)



# Grundlagen von XML

## Elemente, Attribute und Werte (2)

`<ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>`

„Elementwerte“

`<InventionTitle languageCode=„en“>Mus musculus abcd-1 gene for efg protein</InventionTitle>`

# Grundlagen von XML

## Elemente, Attribute und Werte (3)

```
<ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
```

„Attribut“

```
<InventionTitle languageCode=„en“>Mus musculus abcd-1 gene for efg protein</InventionTitle>
```

„Attributwert“

# Grundlagen von XML

## Reservierte Zeichen

Müssen durch ihre vordefinierten Einheiten in einem Elementwert ersetzt werden.

Reserviertes Zeichen	Vordefinierte Einheiten
<	&lt;
>	&gt;
&	&amp;
“	&quot;
‘	&apos;

Beispiel: Erwünschte Merkmallage ist „<50..62“

`<INSDFeature_location><50..62</INSDFeature_location>`



`<INSDFeature_location>&lt;50..62</INSDFeature_location>`



# WIPO ST.26: Beispiel

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="V1_3" fileName="for WIPO Training" softwareName="WIPO Sequence" softwareVersion="1.1.0-beta4" productionDate="2021-03-24">
  <ApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
    <FilingDate>2015-01-31</FilingDate>
  </ApplicationIdentification>
  <ApplicantFileReference>ABC123</ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
    <FilingDate>2014-01-31</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantName>
  <InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
  <SequenceTotalQuantity>1</SequenceTotalQuantity>
  <SequenceData sequenceIDNumber="1">
    <INSDSeq>
      <INSDSeq_length>52</INSDSeq_length>
      <INSDSeq_moltype>DNA</INSDSeq_moltype>
      <INSDSeq_division>PAT</INSDSeq_division>
      <INSDSeq_feature-table>
        <INSDFeature>
          <INSDFeature_key>source</INSDFeature_key>
          <INSDFeature_location>1..52</INSDFeature_location>
          <INSDFeature_qual>
            <INSDQualifier>
              <INSDQualifier_name>mol_type</INSDQualifier_name>
              <INSDQualifier_value>genomic DNA</INSDQualifier_value>
            </INSDQualifier>
            <INSDQualifier id="q2">
              <INSDQualifier_name>organism</INSDQualifier_name>
              <INSDQualifier_value>Mus musculus</INSDQualifier_value>
            </INSDQualifier>
          </INSDFeature_qual>
        </INSDFeature>
      </INSDSeq_feature-table>
      <INSDSeq_sequence>atgaaattaaaacataaaaarggatgataaaatgagatttgatataaaaagg</INSDSeq_sequence>
    </INSDSeq>
  </SequenceData>
</ST26SequenceListing>

```

# WIPO ST.26: Komponenten (1)

**Zeile 1 - die XML-Deklaration**  
(ST.26 Absatz 39(a))

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="V1_3" fileName="for WIPO Training" softwareName="WIPO Sequence"
softwareVersion="1.1.0-beta4" productionDate="2021-03-24">
  <ApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
    <FilingDate>2015-01-31</FilingDate>
  </ApplicationIdentification>
  <ApplicantFileReference>ABC123</ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
    <FilingDate>2014-01-31</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantName>
  <InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
  <SequenceTotalQuantity>1</SequenceTotalQuantity>
  <SequenceData sequenceIDNumber="1">
```

# WIPO ST.26: Komponenten (2)

**Zeile 2 - die Dokumenttyp(DOCTYPE)-Deklaration**  
(ST.26 Absatz 39(b))

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="V1_3" fileName="for WIPO Training" softwareName="WIPO Sequence"
softwareVersion="1.1.0-beta4" productionDate="2021-03-24">
  <ApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
    <FilingDate>2015-01-31</FilingDate>
  </ApplicationIdentification>
  <ApplicantFileReference>ABC123</ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
    <FilingDate>2014-01-31</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantName>
  <InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
  <SequenceTotalQuantity>1</SequenceTotalQuantity>
  <SequenceData sequenceIDNumber="1">

```

# WIPO ST.26: Komponenten (3)

**Zeile 3 - das Wurzelement**  
(ST.26 Absatz 43)

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="V1_3" fileName="for WIPO Training" softwareName="WIPO Sequence"
softwareVersion="1.1.0-beta4" productionDate="2021-03-24">
```

```
<ApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
  <FilingDate>2015-01-31</FilingDate>
</ApplicationIdentification>
<ApplicantFileReference>ABC123</ApplicantFileReference>
<EarliestPriorityApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
  <FilingDate>2014-01-31</FilingDate>
</EarliestPriorityApplicationIdentification>
<ApplicantName languageCode="en">Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantName>
<InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
<SequenceTotalQuantity>1</SequenceTotalQuantity>
<SequenceData sequenceIDNumber="1">
```

**Allgemeine Informationen:**  
(ST.26 Absätze 38(a), 45-49)

# WIPO ST.26: Komponenten (4)

```

<SequenceData sequenceIDNumber="1">
  <INSDSeq>
    <INSDSeq_length>52</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..52</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>genomic DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q2">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Mus musculus</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaattaaaacataaaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
</ST26SequenceListing>

```

## Sequenzdaten

(ST.26 Absätze 38(b),  
50-100)

# WIPO ST.26: Allgemeine Informationen (1)

- Der Abschnitt „Identifikation der Anmeldung“
  - Anmeldenummer, Anmeldetag und Code des Amtes für geistiges Eigentum sind obligatorisch, wenn bekannt;
  - andernfalls ist nur das Aktenzeichen des Anmelders ausreichend

```
<ApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
  <FilingDate>2015-01-31</FilingDate>
</ApplicationIdentification>
<ApplicantFileReference>ABC123</ApplicantFileReference>
<EarliestPriorityApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
  <FilingDate>2014-01-31</FilingDate>
</EarliestPriorityApplicationIdentification>
```

# WIPO ST.26: Allgemeine Informationen (2)

- Der Abschnitt „Prioritätsanmeldung“
  - es kann nur eine Prioritätsanmeldung in das Sequenzprotokoll aufgenommen werden, und es muss die früheste Prioritätsanmeldung sein;
  - obligatorisch, wenn Priorität beansprucht wird

```
<ApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
  <FilingDate>2015-01-31</FilingDate>
</ApplicationIdentification>
<ApplicantFileReference>ABC123</ApplicantFileReference>
<EarliestPriorityApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
  <FilingDate>2014-01-31</FilingDate>
</EarliestPriorityApplicationIdentification>
```

# WIPO ST.26: Allgemeine Informationen (3)

- Der Abschnitt „Anmelder- und Erfindername“
  - im Sequenzprotokoll dürfen nur ein Anmeldername und ein Erfindername enthalten sein und es muss sich um den „Haupt-“Anmelder und -Erfinder handeln;
  - der Anmeldername ist obligatorisch; der Erfindername ist optional;
  - ein Sprachcode für Anmelder- und Erfindernamen ist obligatorisch;
  - enthält der Anmelder- und/oder Erfindername lateinische Grundschrift, die nicht Unicode ist, muss eine Transliteration oder Übersetzung in lateinische Grundschrift enthalten sein.

```

<ApplicantName languageCode="ja">出願製薬株式会社</ApplicantName>
<ApplicantNameLatin>Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantNameLatin>
<InventorName languageCode="ja">特許 太郎</InventorName>
<InventorNameLatin>Taro Tokkyo</InventorNameLatin>
  
```

# WIPO ST.26: Allgemeine Informationen (4)

- Der Abschnitt „Bezeichnung der Erfindung“
  - mindestens eine Bezeichnung der Erfindung in der Sprache der Einreichung ist obligatorisch;
  - zusätzliche Bezeichnungen in anderen Sprachen können enthalten sein;
  - ein Sprachcode ist für jede Bezeichnung obligatorisch;

```
<InventionTitle languageCode="en">My spectacular invention</InventionTitle>  
<InventionTitle languageCode="de">Meine spektakuläre Erfindung</InventionTitle>  
<InventionTitle languageCode="lv">Manš iespaidīgais izgudrojums</InventionTitle>  
<InventionTitle languageCode="ru">Мое зрелищное изобретение</InventionTitle>
```

# WIPO ST.26: Allgemeine Informationen (5)

- Das Element „Sequenzgesamtmenge“
  - ist obligatorisch;
  - die Gesamtmenge muss übersprungene Sequenzen enthalten

<SequenceTotalQuantity>**6**</SequenceTotalQuantity>

# WIPO ST.26: Sequenzen (1)

- Obligatorische Elemente

```

<SequenceData sequenceIDNumber="1">
  <INSDSeq>
    <INSDSeq_length>52</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..52</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>genomic DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q2">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Mus musculus</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaattaaaacataaaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```

Die „Sequenzkennzahl“  
oder „SEQ ID NO:“

Die Sequenzlänge

Die Art des Moleküls  
(DNA, RNA, oder AA)

Die Einteilung  
(immer „PAT“)

# WIPO ST.26: Sequenzen (2)

- Das obligatorische „Source“- oder „SOURCE“-Merkmal

```

<SequenceData sequenceIDNumber="1">
  <INSDSeq>
    <INSDSeq_length>52</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..52</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>genomic DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q2">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Mus musculus</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaattaaaacataaaaaggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```

Ein und nur ein „Source“-Merkmal ist pro Sequenz erforderlich

Die „Source“-Lage muss sich über die gesamte Sequenz erstrecken

Zwei obligatorische Qualifier: „mol\_type“ und „Organismus“

# WIPO ST.26: Sequenzen (3)

## ST.26 Sequenzdaten:

```

<SequenceData sequenceIDNumber="1">
  <INSDSeq>
    <INSDSeq_length>52</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..52</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>genomic DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q2">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Mus musculus</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaattaaaacataaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```

# WIPO ST.26: Sequenzen (4)

- Die obligatorischen Qualifier „mol\_type“ und „MOL\_TYPE“ -  
Werteauswahlmöglichkeiten:

## DNA

genomische DNA

andere DNA

nicht zugeordnete DNA

## RNA

genomische RNA

mRNA

tRNA

rRNA

andere RNA

transkribierte RNA

virale cRNA

nicht zugeordnete DNA

## AA

Protein

# WIPO ST.26: Sequenzen (5)

- Die obligatorischen Qualifier „Organismus“ und „ORGANISMUS“ -  
Werteauswahlmöglichkeiten:
  - Lateinischer Name von Gattung und Art z. B. „Mus musculus“
  - Gattungsname gefolgt von „sp.“ z. B. „Mus sp.“
  - Virusname z. B. „Torque teno virus 1“
  - „nicht identifiziert“
  - „synthetisches Konstrukt“
- Übliche Namen, wie z. B. „Maus“, dürfen nicht als Organismusname  
verwendet werden. Auf Wunsch können übliche Namen im  
Sequenzprotokoll in einem Anmerkungsqualifier aufgenommen  
werden.

# WIPO ST.26: Sequenzen (6)

## Merkmalschlüssel und Qualifier

Zusätzlich zu dem obligatorischen Merkmal „Source“ oder „SOURCE“ können Anmelder mehrere optionale Merkmale hinzufügen, um die Sequenz weiter zu beschreiben

- unterschiedliche Merkmalschlüssel für Nukleotidsequenzen und Aminosäuresequenzen;
- jedes Merkmal kann einen oder mehrere optionale Qualifier haben und kann einen obligatorischen Qualifier haben

5.12. Feature Key	misc_binding
Definition	site in nucleic acid which covalently or non-covalently binds another moiety that cannot be described by any other binding key (primer_bind or protein_bind)
Mandatory qualifiers	bound_moiety
Optional qualifiers	allele function gene gene_synonym map note
Comment	note that the regulatory feature key and regulatory_class qualifier with the value "ribosome_binding_site" must be used for describing ribosome binding sites

(WIPO ST.26, Anlage I, Abschnitte 5-8)

# WIPO ST.26: Sequenzen (7)

- Nukleotidsequenzen:
  - alle Kleinbuchstabensymbole;
  - keine Leerzeichen, keine Nummerierung;
  - keine „u“-Symbole; „t“ steht für Uracil in der RNA
  - „n“ hat den Standardwert von „jeglichem von ‚a‘, ‚c‘, ‚g‘, oder ‚t/u‘“

Symbol	Nucleotide
a	adenine
c	cytosine
g	guanine
t	thymine in DNA/uracil in RNA (t/u)
m	a or c
r	a or g
w	a or t/u
s	c or g
y	c or t/u
k	g or t/u
v	a or c or g; not t/u
h	a or c or t/u; not g
d	a or g or t/u; not c
b	c or g or t/u; not a
n	a or c or g or t/u; “unknown” or “other”

# WIPO ST.26: Sequenzen (8)

- Aminosäuresequenzen:
  - alle Einzelbuchstaben, Großbuchstabensymbole;
  - keine Leerzeichen, keine Nummerierung;
  - „X“ hat einen Standardwert von „jeglichem von ‚A‘, ‚R‘, ‚N‘, ‚D‘, ‚C‘, ‚Q‘, ‚E‘, ‚G‘, ‚H‘, ‚I‘, ‚L‘, ‚K‘, ‚M‘, ‚F‘, ‚P‘, ‚O‘, ‚S‘, ‚U‘, ‚T‘, ‚W‘, ‚Y‘, oder ‚V‘“

Symbol	Amino acid
A	Alanine
R	Arginine
N	Asparagine
D	Aspartic acid (Aspartate)
C	Cysteine
Q	Glutamine
E	Glutamic acid (Glutamate)
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
O	Pyrrolysine
S	Serine
U	Selenocysteine
T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine
B	Aspartic acid or Asparagine
Z	Glutamine or Glutamic acid
J	Leucine or Isoleucine
X	A or R or N or D or C or Q or E or G or H or I or L or K or M or F or P or O or S or U or T or W or Y or V; "unknown" or "other"

# WIPO ST.26: Sequenzen (8)

- Übersprungene Sequenzen: erlauben einem Anmelder, Sequenzdaten aus einem Sequenzprotokoll zu löschen, ohne dass Teilsequenzsequenzen neu nummeriert werden müssen.
  - INSDSeq\_length, INSDSeq\_moltype, INSDSeq\_division vorhanden, aber ohne Wert;
  - Keine Merkmalstabelle und kein Source-Merkmal;
  - Sequenzelement muss den Wert „000“ haben

```
<SequenceData sequenceIDNumber="7">  
  <INSDSeq>  
    <INSDSeq_length/>  
    <INSDSeq_moltype/>  
    <INSDSeq_division/>  
    <INSDSeq_sequence>000</INSDSeq_sequence>  
  </INSDSeq>  
</SequenceData>
```

# WIPO ST.26: Inhalt

- **Hauptteil** - Anforderungen an die Aufnahme/Darstellung
- **Anlage I** - Festes Vokabular auf der Grundlage von INSDC
- **Anlage II** - ST.26 Dokumententyp DTD
- **Anlage III** - Beispiel ST.26 Sequenzprotokoll XML-Datei
- **Anlage IV** - Zeichenteilmenge des lateinischen Grundcodes für ST.26 XML-Instanz
- **Anlage V** - INDS-Datenaustauschanforderungen (nur IPO)
- **Anlage VI** – Anleitungsdokument mit Beispielen
- **Anhang zu Anlage VI** - XML-Datei mit allen in Anlage VI beispielhaft aufgeführten Sequenzoffenbarungen
- **Anlage VII** - Empfehlung für die Umwandlung eines Sequenzprotokolls von ST.25 nach ST.26

# WIPO ST.26:

## Hauptteil

Absätze	Inhalt
1-9	Einleitung; Definitionen; Anwendungsbereich; Verweise
10-37	Darstellung von Sequenzen
38-49	Struktur des Sequenzprotokolls in XML
50-71	Sequenzdatenteil; Merkmalstabelle; Merkmalschlüssel; Obligatorische Merkmalschlüssel; Lage von Merkmalen
72-84	Merkmalsqualifier; Obligatorische Merkmalsqualifier
85-100	Freier Text; Kodierungssequenzen; Varianten

# WIPO ST.26 Anlage I:

## Festes Vokabular

Abschnitt	Inhalt
1	Liste der Nukleotide (Kleinbuchstaben, Einzelbuchstabensymbole)
2	Liste modifizierter Nukleotide
3	Liste der Aminosäuren (Großbuchstaben, Einzelbuchstabensymbole)
4	Liste modifizierter Aminosäuren
5	Merkmalschlüssel für Nukleotidsequenzen
6	Qualifier für Nukleotidsequenzen
7	Merkmalschlüssel für Aminosäuren (angelehnt an UniProt)
8	Qualifier für Aminosäuresequenzen
9	Tabellen genetischer Codes

# WIPO ST.26 Anlage II:

## ST.26 DTD

- Details zu ST.26 DTD: Derzeitige Version 1.3
  
- Teil Allgemeine Informationen
  - Elemente, die sich auf Informationen zur Patentanmeldung beziehen
  
- Teil Sequenzdaten
  - Teilmenge der INSDC-DTD
  - Eines oder mehrere Sequenzdatenelemente, wobei jedes Element Informationen über eine Sequenz enthält

# WIPO ST.26 Anlage VI:

## Anleitungsdokument

- Enthält 49 reale Beispiele für Sequenzoffenbarungen und eine Erklärung, wie die ST.26-Regeln auf jedes Beispiel anzuwenden sind;
- Jedes Beispiel behandelt:
  1. ob die Sequenz für die Aufnahme in ein Sequenzprotokoll erforderlich, zulässig oder verboten ist;
  2. wenn die Sequenz in ein Sequenzprotokoll aufgenommen werden muss oder darf, wie sie dargestellt werden muss.
- Der Anhang zu Anlage VI ist ein ST.26 XML-Sequenzprotokoll, das alle im Anleitungsdokument dargelegten Beispiele enthält.

# WIPO ST.26 Anlage VII:

## Empfehlung für die Umwandlung eines Sequenzprotokolls von ST.25 zu ST.26

- ... oder wie die Aufnahme von neuem Offenbarungsinhalt vermieden werden kann.
- Die Anforderungen von ST.26 unterscheiden sich von denen von ST.25. ST.26-Regeln erfordern Informationen, die in den ST.25-Regeln nicht erforderlich waren
- Die Umwandlung eines Sequenzprotokolls vom ST.25-Format in das ST.26-Format erfordert immer eine Eingabe vom Anmelder
- Die Umwandlung eines ST.25-konformen Sequenzprotokolls in das ST.26-Format führt nicht zu neuem Offenbarungsinhalt, wenn die Empfehlungen in Anlage VII befolgt werden.
- Zwanzig Umwandlungsszenarien werden mit Empfehlungen und Beispielen vorgestellt

# WIPO Sequence (1)

- Von der WIPO entwickeltes Desktop-Tool zur Unterstützung der Erstellung, Validierung und Erzeugung von ST.26-konformen Sequenzprotokollen
- Die Mitgliedsstaaten haben die WIPO ersucht, dieses gemeinsame Tool für alle Ämter und Anmelder auf internationaler, nationaler und regionaler Ebene zu entwickeln
- Die Verwendung von WIPO Sequence vereinfacht die ST.26-XML-Erstellung mit einer benutzerfreundlichen Oberfläche: Es müssen keine XML-Dateien mehr direkt bearbeitet werden
- Laden Sie die neueste Version kostenlos herunter von:  
<https://www.wipo.int/standards/en/sequence/index.html>

# WIPO Sequence (2)

- Sequenzinformationen können in einem Projekt gespeichert und validiert werden und dann kann ein Sequenzprotokoll im ST.26-Format erzeugt werden
- Daten können importiert werden aus: ST.26-Sequenzprotokollen, ST.26-Projekten, ST.25-Sequenzprotokollen, Dateien im Multi-Sequenz-Format, Dateien im Raw-Format und Dateien im FASTA-Format
- Validierung von Sequenzprotokollen auch im XML-Format
- Maßgebliche Merkmalschlüssel, Qualifier und Organismuskennungen können leicht aus Dropdown-Menüs ausgewählt werden
- Anmelder- und Erfinderinformationen können in einer Datenbank für „Personen und Organisationen“ gespeichert werden
- Unterstützt den Export und Import von XLIFF-Dateien, die von Übersetzern verwendet werden

# WIPO Sequence: Projekte Homepage

WIPO | Sequence    **PROJECTS**    PERSONS & ORGANIZATIONS    ORGANISMS    HELP    PREFERENCES    ENGLISH

[NEW PROJECT](#)   [IMPORT PROJECT](#)   [IMPORT SEQUENCE LISTING](#)   [VALIDATE SEQUENCE LISTING](#)

## PROJECTS

Q

Project name :	Applicant file reference	Applicant name	Invention title	Status	Creation Date :
<a href="#">16079428_no_source_sequence</a>	H0075.70243US00	Ferring B.V.	STABLE LIQUID GONADOTROPIN FORMULATION	invalid	2021-02-01
<a href="#">All features and qualifiers</a>	123abc	Simple Healthkit, Inc.	Compositions and Methods for Treating Cancer	modified	2021-02-01
<a href="#">Ark three letter aa code import from ST.25</a>	abc123	Joe, Smith	Improper 3-letter AA code for import	new	2021-02-26
<a href="#">Beta 4 Test Project</a>	abc123	Stephenson, Adam	Beta 4 Test Project	modified	2021-03-11
<a href="#">DNA RNA with u and t</a>	uspto	Biotech, Inc.	Beta 4 test import ST25	modified	2021-03-15

# WIPO Sequence: Projektdetails

## Abschnitt Allgemeine Informationen

WIPO | Sequence
 FOR WIPO TRAINING
VERIFICATION REPORT
FREE TEXT QUALIFIERS
IMPORT REPORT
DISPLAY THE SEQUENCE LISTING
HELP

 PREFERENCES ENGLISH
 Return to project home

**GENERAL INFORMATION SEQUENCES**

### GENERAL INFORMATION

#### APPLICATION IDENTIFICATION

Application Identified Before the assignment of the application number Application filed Applicant file reference ABC123	IP Office IB - International Bureau of the World Intellectual Property Organization (WIPO) Application number PCT/IB2015/099999 Filing date 2015-01-30	✎
--	--	---

#### PRIORITY IDENTIFICATION

IP Office	Application Number	Filing date	Selected Earliest Priority Application
IB - International Bureau of the World Intellectual Property Organization (WIPO)	<a href="#">PCT/IB2014/111111</a>	2014-01-30	Yes

#### APPLICANT & INVENTOR

Applicant name Shutsugan Pharmaceuticals Kabushiki Kaisha	Primary applicant	✎
---	-------------------	---

#### INVENTION TITLE

Invention title Mus musculus abcd-1 gene for efg protein	Language en - English	✎
--	-----------------------	---

# WIPO Sequence: Projektdetails

## Abschnitt Sequenzen

WIPO | Sequence
 FOR WIPO TRAINING
VERIFICATION REPORT
FREE TEXT QUALIFIERS
IMPORT REPORT
DISPLAY THE SEQUENCE LISTING
HELP

 PREFERENCES ENGLI
 Return to project home

GENERAL INFORMATION SEQUENCES

---

### SEQUENCE 1

Sequence Number (ID) 1	Molecule Type DNA	
Sequence Name test	Organism Mus musculus	
Length 52		

▼ FEATURES

Add feature

Feature Key	Location	Qualifiers
<u>source</u>	1..52	mol_type = genomic DNA organism = Mus musculus

▼ SEQUENCE

```

atgaaattaa  aacataaaaar  ggatgataaa  atgagatttg  atataaaaaa  gg
          
```

52

< 1/1 >

Wie geht es weiter?

[Standards@wipo.int](mailto:Standards@wipo.int)

# Frage-und Antwort-Sitzung

# Glossar: Akronyme

- CWS: Ausschuss für WIPO-Standards
- DDBJ: DNA-Datenbank von Japan
- EMBL-EBI: Europäisches Institut für Bioinformatik
- EPO: Europäisches Patentamt
- INSDC: Internationale Nukleotidsequenz-Datenbank-Zusammenarbeit
- IPO: Amt für geistiges Eigentum
- NCBI: Nationales Informationszentrum für Biotechnologie
- WIPO: Weltorganisation für geistiges Eigentum