

# **GUÍA DE ANOTACIÓN Y NORMALIZACIÓN DE COMPUESTOS QUÍMICOS**

**Plan de impulso de las Tecnologías del Lenguaje**

**Obdulia Rabal**

**Ander Intxaurreondo**

**Martin Krallinger**

**Julio 2018**



Este estudio ha sido realizado dentro del ámbito del Plan de Impulso de las Tecnologías del Lenguaje con financiación de la Secretaría de Estado para el Avance Digital, que no comparte necesariamente los contenidos expresados en el mismo. Dichos contenidos son responsabilidad exclusiva de sus autores.

Reservados todos los derechos. Se permite su copia y distribución por cualquier medio siempre que se mantenga el reconocimiento de sus autores, no se haga uso comercial de las obras y no se realice ninguna modificación de las mismas.



## 1 ÍNDICE

1	INTRODUCCIÓN	4	1.1 ENTIDADES QUÍMICAS	52	REGLAS DE ANOTACIÓN	62	1.1 Reglas generales (Reglas-G)	62	1.2 Reglas positivas (Reglas-P): Qué anotar	72	1.3 Reglas Negativas (Reglas-N): Qué NO se debe anotar	162	1.4 Reglas de Ortografía y gramaticales (Reglas-O)	212	1.5 Reglas Multipalabra (Reglas-M): entidades únicas <i>versus</i> entidades múltiples	253	Normalización y clases	314	Referencias	335	Glosario de siglas y acrónimos	34
---	--------------	---	------------------------	----	---------------------	----	---------------------------------	----	---	----	--	-----	--	-----	--	-----	------------------------	-----	-------------	-----	--------------------------------	----



## RESUMEN

Este documento describe las directrices de anotación usadas para la construcción de las anotaciones de menciones químicas, fármacos y biosimilares con relevancia terapéutica en informes clínicos.

## 2 INTRODUCCIÓN

---

El Plan de Impulso de las Tecnologías del Lenguaje (Plan TL) tiene como objetivo fomentar el desarrollo del Procesamiento del Lenguaje Natural (PLN) y la Traducción Automática (TA) en lengua española y lenguas cooficiales. Para ello, el Plan TL define medidas que:

- Aumenten el número, calidad y disponibilidad de las infraestructuras lingüísticas en español y lenguas cooficiales;
- Impulsen la Industria del lenguaje fomentando la transferencia de conocimiento entre el sector investigador y la industria; y
- Incorporen a la Administración como impulsor del sector de PLN.

Uno de los objetivos del proyecto es poner a disposición de la comunidad científica y la industria un corpus biomédico exhaustivo y con licencia abierta que permita ejecutar tareas de PLN sobre *big data* y replicar los experimentos. Este documento describe las directrices de anotación usadas para la construcción de las anotaciones de menciones químicas, fármacos y biosimilares con relevancia terapéutica en informes clínicos.

El propósito final de la tarea es identificar aquellos compuestos químicos / fármacos o biosimilares que aparecen en informes clínicos y que son relevantes porque son utilizados como tratamiento o diagnóstico. **Sin embargo, el contexto (tratamiento o diagnóstico) NO se va a considerar para realizar las anotaciones, ya que el objetivo final es tener la anotación final de la mención.**

Estas guías están basadas en las directrices previamente preparadas para las tareas de CHEMDNER 2013 (Biocreative IV) y CHEMDNER-PATENTS 2015 (Biocreative V y V.5) para la identificación de entidades químicas extraídas de artículos científicos y patentes en inglés, respectivamente.

Los puntos principales que se han modificado son:

- Permitir que se anoten entidades bioquímicas que son aplicadas como tratamiento o como método de diagnóstico: hormonas, anticuerpos, proteínas.
- Eliminar la diferenciación de los tipos de entidad por sub-clases (abreviación, IUPAC...).



- Incorporar la normalización de estructuras químicas a bases de datos.

Este documento proporciona los detalles básicos de la anotación y las convenciones que deberían ser seguidas durante el proceso de construcción del corpus. Las directrices de anotación se refinaron tras sucesivos ciclos iterativos de anotaciones de documentos de muestra, basadas en las sugerencias directas realizadas por los anotadores así como por la observación de inconsistencias detectadas al comparar los resultados de diferentes anotadores.

La tarea de anotación manual básicamente consiste en etiquetar o marcar manualmente las menciones de entidades químicas a través de una interfaz web adaptada (ANNOTATE). El proceso se realizó siguiendo un conjunto de reglas que se detallan específicamente más adelante. El texto para ser etiquetado consistió en informes clínicos en español proporcionados por el Centro Nacional de Investigaciones Oncológicas (CNIO).

## 2.1 ENTIDADES QUÍMICAS

El objetivo principal para definir las entidades químicas y biosimilares anotadas para la construcción del corpus sobre informes clínicos fue primariamente capturar todas las entidades químicas y fármacos que aparecen en un informe clínico, independientemente de su contexto (tratamiento o diagnóstico). Definir qué es una entidad química NO es sencillo, ya que engloba:

- *Términos químicos*: si se representan mejor usando un diagrama químico.
- *Términos bioquímicos*: si normalmente se representan usando una secuencia o un diagrama de bloques.

Por lo tanto, las entidades cubiertas tienen que representar a aquellos tipos de menciones que pueden ser explotados para la extracción de información relevante desde estos puntos de vista de una serie de documentos en castellano.

La anotación realizada para esta tarea fue exhaustiva exclusivamente para los tipos de entidades que se describen en más detalle abajo. Esto implica que otros tipos de menciones biológicas no se anotaron. Conceptos químicos generales y vagos (no estructurales, nombres químicos no específicos), adjetivos, verbos y otros términos (reacciones químicas) están explícitamente descartados de la anotación.

El proceso de anotación en si mismo es altamente dependiente del nivel de conocimiento de los anotadores al etiquetar menciones químicas y biosimilares. Un requisito para realizar la anotación manual fue que los anotadores deberían tener un conocimiento en química, quimioinformática o bioquímica para asegurarse de que las anotaciones sean correctas. Esto facilita también la distribución de un conjunto de reglas de anotación corto y compacto, en lugar



de unas pautas más detalladas para no expertos. En este sentido, se han seguido estrategias similares a las realizadas para las tareas de anotación de genes en otras competiciones BioCreative (Smith et al. 2008). La definición de las menciones de entidades químicas usadas para la tarea CHEMDNER se inspiró en las reglas de anotación de Kolaric et al. (2008) y Corbett et al. (2007).

Las menciones de entidades químicas para esta tarea corresponden a nombres de compuestos químicos específicos, clases específicas de compuestos químicos o fragmentos químicos (estos últimos muy poco frecuentes en informes clínicos). Se anotarán siempre, independientemente de si son para tratamiento terapéutico o en diagnóstico.

Las menciones de entidades biológicas que se refieran a péptidos, proteínas, genes, anticuerpos, enzimas... se anotarán como pertenecientes a la clase PROTEÍNA.

Las menciones de entidades bioquímicas macromoleculares (lípidos y derivados, azúcares, polisacáridos,... se anotarán como cualquier otro compuesto químico.

A continuación se detallan las reglas, incluyendo ejemplos particulares destinados a ayudar en la comprensión de las directrices. En los casos correctos, los diferentes casos están marcados en gris.

Las reglas de anotación están estructuradas en las siguientes clases:

- Reglas-G: generales.
- Reglas-P: positivas
- Reglas-N: negativas
- Reglas-O: ortográficas
- Reglas-M: multipalabra

## 3 REGLAS DE ANOTACIÓN

---

### 3.1 REGLAS GENERALES (REGLAS-G)

#### G1. Uso de recursos de conocimiento externos

En caso de duda del anotador, sobre si corresponde o no a una mención o acerca del tipo de mención química correspondiente, se aconseja consultar a recursos externos en español (ej. Wikipedia o páginas web especializadas en español: Biblioteca Virtual en Salud España<sup>1</sup>, LIS España. Sitios saludables para profesionales<sup>2</sup>, Biblioteca virtual en

---

<sup>1</sup> <http://bvsalud.isciii.es/> (accessed Apr 3, 2017).



salud. Portal Regional de la BVS<sup>3</sup>). Alternativamente, se pueden consultar otros recursos en inglés que faciliten la comprensión del término: Wikipedia, Chemspider, Chemical Suppliers Catalogues (Sigma Aldrich, Tocris,...), Scifinder u otras bases de datos químicas para resolver las dudas.

Un recurso útil para las abreviaturas médicas: <http://www.sedom.es/diccionario/>

## **G2. Menciones que no estén claras**

No anotar menciones para las que no esté claro si corresponde anotarlas (incluso tras consultar algún recurso externo).

## **G3. Revisión de las reglas de anotación**

Si se detectan casos especiales de tipos de menciones que podrían ser de interés o estar relacionadas y las guías no especifican su anotación, se deben reportar estos casos junto con ejemplos para refinar las reglas de anotación.

## **G4. Número de menciones a incluir**

Se anotarán todas las menciones de un texto, independientemente de si están repetidos o no.

### **3.2 REGLAS POSITIVAS (REGLAS-P): QUÉ ANOTAR**

#### **P1. Nombres químicos:**

- Convertibles a un diagrama estructural químico único: átomos individuales, iones, isótopos, elementos químicos puros y moléculas.

Fluor, Hierro, Deuterio, Benceno, Piridina, Plomo(II), acetato de plomo (II)

Y que por lo tanto pueden ser normalizadas frente a una entidad química específica.

- Convertibles a un diagrama Markush o fórmula estructural genérica. Típicamente, funcionalidades químicas generales, fragmentos y clases estructurales.

Amidas, Hidroxipiridinas, ROH, Aminoácidos, grupo metilo, grupo O-H.

Estos nombres químicos pueden aparecer como:

---

<sup>2</sup> <http://bvsalud.isciii.es/lis-profesionales/> (accessed Apr 3, 2017).

<sup>3</sup> <http://pesquisa.bvsalud.org/portal> (accessed Apr 3, 2017).



a. Nombres sistemáticos de menciones químicas (IUPAC y similares a IUPAC):

Incluye nombres sistemáticos multipalabra, nomenclatura siguiendo el estilo CAS (adaptada del inglés) y nombres semi-sistemáticos como las menciones de compuestos químicos siguiendo parcialmente las directrices de la nomenclatura IUPAC. En caso de duda, se pueden consultar las guías IUPAC en español<sup>4-6</sup> y otros recursos en inglés.<sup>7,8</sup>

1,2-dimetil-3-hidroxipiridin-4-ona

Citrato de calcio

1-octanol

Benceno

Cloruro de cloroacetilo

hierro

hierro (III)

hierro (3+)

ácido acetilsalicílico

[(2-pyridyl)-methyl]d-Dap(3)]degarelix

Quinolinas

Éter de diacilglicerol

Aldehído

*incluye términos referidos a grupos químicos generales*

Hydróxido

*incluye términos referidos a grupos químicos generales*

Terpenoides

*términos de familias generales*

b. Identificadores en bases de datos:

Identificadores de bases de datos de compuestos químicos (muchas de ellas en inglés, pero que se han adaptado al uso en español): números de registro CAS, PubChem, identificadores ChEMBL y ChEBI y también nombres de compuestos de compañías. Estos identificadores deber ser etiquetados únicamente si el texto proporciona suficiente información como para deducir que dichos identificadores corresponden a identificadores químicos.

- **Identificadores de compañías:** PD-0332991, FE200486
- **Identificadores de bases de datos:** CID 445154, CHEBI:28262, ChEMBL504, CAS 28718-90-3

c. Fórmulas:

Menciones de fórmulas químicas, notaciones químicas lineales, SMILES, códigos InChI y códigos de 3 letras de nucleótidos, aminoácidos y monosacáridos.





$C_6H_{12}O_6$

EtOAc

Fe, Na, Fe(III), Li<sup>+</sup>, Fe<sup>2+</sup> *Elementos atómicos*

CC(=O)C *Fórmulas en línea*

Glu-Cys-Gly *Códigos 3 letra de aminoácidos*

GlcNAc *Nomenclatura de oligosacáridos*

Asp-Glu-Fmoc *Combinaciones de fórmulas y abreviaciones*

t-BuOK

(Asp<sub>3</sub>Phe<sub>1</sub>)<sub>2</sub>

Éster C-16/C-26

Ésteres de alquilo de C9-C15

MCl<sub>2</sub> *Fórmulas genéricas donde la M es cualquier metal*

ROH *Fórmulas genéricas donde la R es cualquier grupo*

#### d. Nombres triviales:

Nombres triviales y comunes, marcas y nombres comerciales de fármacos:

- **Nombres de fármacos aprobados (International Nonproprietary Name – INN; British Approved Name – BAN- y United States Adopted Name – USAN):** aspirina, Viagra, Degarelix, Dormidina, N-metil sildenafilo, Clindamicina, ...
- **Otros nombres comunes (principalmente para moléculas pequeñas bioquímicas):** adenina, testosterona, mezerina, azalina B, manitol, rosiglitazona, deferiprona, vitamina C, Triacilglicerol, 3H-Testosterona, Bilirrubina,
- **Nombres generales de compuestos asociados al consumo de sustancias:** alcohol, tabaco, gluten, droga y drogas específicas...
- **Aminoácidos:** serina, asparagina
- **Azúcares:** glucosa, lactosa, lactulosa
- **Polímeros:** Polietilenglicol 6000, Brij35, Gelucire 44/14, poloxámero 188

#### e. Abreviaciones:

Menciones de abreviaciones y acrónimos de compuestos químicos y fármacos. Solamente se deben etiquetar



aquellas abreviaciones que se puedan unir claramente a entidades químicas según el conocimiento de fondo del anotador y en las descripciones incluídas en el texto (abreviaciones ad-hoc), sin ningún tipo de duda. No se anotarán si se refieren a pseudoanáforas o similares.

Las abreviaciones/acrónimos y otras definiciones similares que aparezcan justo antes/después de un nombre químico, deben ser etiquetadas separadamente:

DMSO

8-hidroxi-2-(N,N-di-n-propilamino)tetralina (8-OH-DPAT)

2,4-dinitrofenil-sulfenilo (DNPS)

Ácido gamma-aminobutírico (GABA)

[3H]-8-OH-DPAT

5-HT

8-hidroxi-2-(N,N-di-n-propilamino)tetralina

AG = Ácidos grasos

ACs = antocianinas

H-Lys-Trp(NPS)-OMe *incluye acrónimos dentro de nombres químicos*

Hexahidro-1-nitroso-3,5-dinitro-1,3,5-triazina (MNX)

LT = Levotiroxina

CT = Colesterol

AG = Ácido graso

**P2. Nombres generales de clases químicas en las que la definición de la clase incluye información estructural o de composición elemental química de algún tipo**, independientemente de su origen (sintético o productos naturales).

Hidrocarburos, organocloruros, carbohidratos, hidratos de carbono, organometálicos, Ácidos de Lewis, Reactivos de Grignard, policétidos, esteroides, macrólidos, terpenoides, ácidos grasos, nucleótidos, nucleobases, ácidos de Brönsted-Lowry, metales de transición, halógenos, base de Schiff, sal de Wittig, reactivos de Wittig, monosacáridos, azúcares, ácidos grasos saturados, ácidos grasos trans, triglicéridos, ácidos biliares, sales biliares, estatinas, betaínas, ceramida, glucosilceramida, carotenoide, cetocarotenoide, alcaloide, andrógeno, estrógeno, agliconas, cannabinoide, opioide, fosfoinosítido, saponina, sapogenina, parabeno, vitamina, mineralocorticoide, alcano, alqueno, alquino, leucotrieno, hormona esteroide, glucocorticoide, glicoesfingolípido, esfingolípido, aminoglucósido, glucósido, oligosacárido, dipéptido, nonapéptido,



aminoácido, tierras raras, glucosinolato, glicósido, proteínas, corticoides, glicoproteína, lipopolisacárido, ADN, mRNA, gen, lípido, ácido nucleico, ácido ribonucleico, péptido, peptidoglicano, grasas saturadas

### P3. Sacáridos y polisacáridos

Anotar todos, independientemente del número de sacáridos (mono-, di-, tri-, poli-) o si es un polímero.

Glucosa	<i>monosacárido</i>
Sacarosa	<i>disacárido</i>
Estreptomicina	<i>un aminoglicósido trisacárido</i>
Gentamicina	<i>un aminoglicósido trisacárido</i>
Ciclodextrina	<i>oligosacárido cíclico</i>
Almidón	
Almidón acetilado	
Celulosa	
Ácido polisiálico	
Hidroxipropilmetilcelulosa	
Hipromelosa	
Vancomicina	<i>glucopéptido</i>

### P4. Lípidos

Anotar todos los ácidos grasos y sus derivados (incluyendo mono-, di- y triglicéridos), derivados de esteroles, fosfolípidos y estructuras poliméricas.

Glicerol  
Prostaglandina A  
Leucotrieno A4  
Colesterol  
Eicosanoide  
Cefalina

### P5. Nucleótidos, oligonucleótidos y genes

Nucleótidos y oligonucleótidos anotarlos

NADH



NAD+

Nicotinamida adenina dinucleótido

ATP

Adenosina 5'-trifosfato

SAM

S-adenosil metionina

cAMP

Adenosín monofosfato cíclico

Genes y derivados → anotarlos (pertenece a la clase PROTEÍNA)

### P6. Aminoácidos individuales, péptidos y proteínas

Los aminoácidos se anotarán como otro compuesto químico cualquiera.

Salvo excepción (ver abajo), los péptidos y proteínas (independientemente de su tamaño y del contexto en el que aparezcan, tratamiento o diagnóstico), se anotarán siempre y pertenecerán a la clase PROTEÍNA. Incluye enzimas.

Glutación

Hormona liberadora de gonadotropinas (GnRH)

Hormona liberadora de hormona luteinizante (LHRH)

Angiotensina

des-Gly10, [D-Ala6] LHRH Etilamida *también modificaciones químicas*

Queratina

Interferón de fibroblastos humanos

“...coagulopatía (actividad de protrombina 30% y tiempo de cefalina no coagula)...”

“...concentraciones máximas de transaminasas GOT 957 U/L, GPT 237 U/L y GGT 144 U/L, FA 93 U/L, LDH 5031 U/L...”

“...determinación de hierro, ferritina, transferrina...”

Notar que aquí caen también las hormonas peptídicas (que se anotarán siempre) e irán asignadas a la clase PROTEÍNA.

Excepción: los fármacos que son de origen peptídico no irán a la clase PROTEÍNA:

Octreótido



Bacitracina

Sulfato de polimixina B

Ciclosporina A

Degarelix

Gramicidina

“...otros tratamientos, como ciclosporina o gammaglobulina intravenosa” (Ciclosporina a su clase de NORMALIZABLES y gammaglobulina a PROTEÍNA)

PPD tuberculina

A la hora de anotar las proteínas, no se anotarán características especiales (localización en tejidos, etc) que impidan el proceso de normalización en el recurso Snomed:

Actina músculo específica

#### **P7. Polímeros sintéticos y naturales (exceptuando genes y proteínas, sobre los que aplican sus propias reglas).**

Nylon

Poliestireno

Sulfonato sódico de poliestireno

Cloruro de polivinilo (PVC)

Poliamidas

Poliacrilamida (PAM)

Nafion

PoliHis

PoliAsp

Poli(ácido L-láctico)

#### **P8. Minerales, aleaciones y Alótropos**

Calcita

Sílice

Alúmina

Titanita

Diamante

Grafito



Ozono

Bronce

Acero

### **P9. Reactivos de laboratorio, diagnóstico, tintes y nombres de indicadores**

Reactivos comunes de un laboratorio de diagnóstico, de composición química definida

Éter de petróleo

Gel de sílice

Indicador universal

Tamices Moleculares

Tornasol

Aceite de parafina

Salino

Solución fisiológica

Rojo de metileno

Azúl brillante de Coomassie

DAPI

### **P10. Plantas, aceites, esencias, principios de plantas y formulaciones/composiciones generales de varios compuestos**

Pertenecerán a la clase UNCLEAR.

Estragón

Silimarina

Bromelaína

Melanina

Vaselina

Lanolina

Protocolo de quimioterapia INTERFAN 2006

### **P11. Anticuerpos**

Pertenecerán a la clase PROTEÍNA.



Trastuzumab

Alemtuzumab

Brentuximab vedotin

Daclizumab

Denosumab

Herceptin

Kadcyla

MEDI-575

Natalizumab

Ocrelizumab

ASTO = Antiestreptolisina

ANCA = anticuerpos contra los antígenos del citoplasma

Antitiroglobulina (ATG)

Factor reumatoide

IgM e IgG anti-Clamidia e IgM anti-Mycoplasma pneumoniae

## P12. Hormonas

Anotar las hormonas. Aquellas que sean de naturaleza peptídica se enviarán a la clase PÉPTIDOS. Las otras, se enviarán a su correspondiente clase de NORMALIZABLES o NO NORMALIZABLES.

Hormona liberadora de gonadotropinas (GnRH) *clase PÉPTIDO*

Hormona liberadora de hormona luteinizante (LHRH) *clase PÉPTIDO*

Hormona luteinizante (LH) *clase PÉPTIDO*

Insulina *clase PÉPTIDO*

Adrenalina

TSH = Tirotropina

T3 = Triyodotironina

T4 = L-tiroxina

## P13. Vacunas anotadas siempre

Anotar siempre, tanto si están en el contexto de tratamiento como de diagnóstico. Pertencerán a la clase UNCLEAR ya que no es posible determinar que todas ellas tengan naturaleza peptídica.

primera dosis de HBV



vacuna séxtuple [DTPa, Hib, IPV, HB] y meningocócica C

vacuna acelular contra la tos ferina

pero NO si se refiere al virus / bacteria:

“...descartar infección por VEB, CMV, herpes simple de tipos I y II, enterovirus, adenovirus, toxoplasma, rubéola y sífilis...”

“...serologías para VHB, VHC, Bartonella, Mycoplasma y EBV”

### 3.3 REGLAS NEGATIVAS (REGLAS-N): QUÉ NO SE DEBE ANOTAR

**N1. Términos diferentes de nombres químicos:** adjetivos (si están aislados/separados de nombres químicos – ver las reglas M5-M9), pronombres, verbos, otros términos (reacciones), prefijos químicos (si están aislados/separados de nombres químicos), anáforas ó números de compuestos que no corresponden a identificadores.

- **Reacciones químicas:**

Deshidrogenación

Metilación

Hidrólisis

- **Pronombres, anáforas o pseudoanáforas:**

“DAPI es un tinte...**este** compuesto...” *no etiquetar “este”*

“...los agroquímicos triadimefon e imazalil (MIX2) o triadimefon, imazalil, y el fluconazol usado clínicamente (MIX3)...” *no anotar MIX2 or MIX3*

- **Números de compuestos en anáforas:** incluso si los números se combinan con otras palabras (generando anáforas), nunca deberían ser anotados:

...of benceno (2)

*no anotar “2”*

(S)-4-AHCDP (6) and (R)-4-AHCP (7)

*no anotar “6” y “7”*

cis-9, orto-12

*no anotar estas entidades*

cis-isómero 22

*no anotar anáforas o términos generales*

- **Prefijos químicos (fuera de nombres químicos):**

1,4-derivados

*no etiquetar “1,4-”*





**N2. Nombres químicos referentes a roles/actividades químicas o biológicas que no se refieren a conceptos estructurales:**

- **Generalidades:** análogo, sustituyente, inhibidor, hit, agonista, antagonista, activador, efector, antioxidante, sustrato, inactivador, pigmento, agente, estándar, farmacóforo, fármaco, medicamento, promotor, exón, intrón, gen, antifolato, comida, compuesto, derivado,...

gases en sangre

pérdida de grasa

Pero sí que habría que anotar nombres particulares que contengan cualquiera de estas palabras y sí se refieran a una estructura química en particular:

**Compuesto C** *corresponde a una sustancia química (CAS 866405-64-3)*

- **Actividades/roles biológicos generales:** hormona, antibiótico, antígeno, herbicidas, antifúngicos, toxina, metabolito, agentes antineoplásicos, antiestrógenos, placebo, antihistamínicos, antileucotrienos,...

“se comenzó con el tratamiento antibiótico”

“se le administraron analgésicos”

“continuar con un quelante”

“se administraron antiinflamatorios”

Pero sí que habría que anotar casos particulares que puedan ser conectados con una estructura química general o particular:

**Ciguatoxinas** *familia de toxinas*

**Aflatoxinas** *familia de toxinas compuesta de difurano y cumarina*

**Aflatoxina B1** *compuesto químico concreto*

**Hormona tiroidea** *corresponde a 2 compuestos químicos definidos*

- **Actividades / roles Industriales:** Surfactante, Plastificante, Aerosol

“se le administró surfactante”

- **Reactividades:** electrófilo, nucleófilo, aceptor de Michael, dienófilo, quelante, reactivo alquilante, oxidante, catión, anión, lipófilo, zwitterión,...



“...galactosemia (cuerpos reductores en orina...)”

Pero sí que hay que anotar:

- Aquellos casos particulares en los que alguna de estas palabras (catión, anión, zwitterión...) forman parte de un nombre químico de mayor longitud, en los que aportan información estructural (ver regla M4):

Anión carbonato

- Aquellos casos particulares en los que el adjetivo de estas palabras añade información adicional sobre la estructura química (ver regla M5):

PAM aniónica

MF1011 aniónico

siloxano zwitteriónico

- **Actividades / roles de laboratorio / síntesis:** solvente, reactivo, materiales de partida, tampón, catalizador, solución, disolvente

Solución de amoniaco *solamente se anotaría “amoníaco”*

- **Partículas elementales:** neutrón, protón, electrón

### N3. Conceptos estructurales que sean poco específicos:

- **Conceptos estructurales muy generales generales:** átomo, ión, molécula, polímero, estereoisómero, enantiómero, isómero, confórmero, mesómero, conformación, monómero, dímero, trímero, tetrámero, grupos funcionales, transportadoras, terpolímeros,

El estereoisómero 6, pero no 7, activa *no etiquetar*

Terpolímero *no etiquetar*

Pero sí que hay que anotar:

- Clases particulares en las que alguna de estas palabras (ión, dímero, trímero) sean parte de un nombre químico específico mayor (ver regla M4):

Ión cloruro *cloruro como ión*

Catión sodio *sodio como catión, no como elemento químico*



Dímeros de tiol *dímero añade información sobre el tipo de tiol*

Radical hidroxilo

Metales alcalinos

- Casos particulares en los que la forma adjetivo de estas palabras añade información adicional sobre la estructura química (ver regla M5):

Acetónido de triamcinolona polimérico

- **Descriptores topológicos vagos:** macrociclo, catenano, rotaxano,...

Pero sí que hay que anotar:

- Casos particulares en los que alguna de estas palabras (macrociclo) formen parte de un nombre químico específico mayor (ver regla M4):

Macrociclo de tetrapirrol

- Términos topológicos que formen parte de descripciones de fragmentos populares principalmente en patentes químicas en las que se listan motivos de sustitución cubiertos/preferidos de las fórmulas Markush. NO serán nada comunes en informes clínicos.

Compuestos bicíclicos aromáticos y heteroaromáticos

Compuestos de anillo espiro-tetracíclico

Biciclos de 8 hasta 11 miembros

Anillo aromático bicíclico o monociclo de 5 hasta 10 miembros

**N4. Criterio de contexto: hay términos que no son válidos para anotación si ellas no se refieren a términos químicos/biosimilares en el contexto, incluso si coinciden que son la misma palabra (sinónimos y metáforas):**

Fósforo no etiquetar si se refiere a una cerilla

**N8. Palabras especiales que no se van a etiquetar por convención, como los términos generales de consumo del paciente que escapan a la lista concreta (alcohol, tabaco, gluten, drogas) de términos que sí se permiten.**

Orgánico

Aminas orgánicas

Inorgánico



Agua y sus diferentes estados físicos (Vapor, Hielo...) así como sus adjetivos (acuoso)

Protón, (protón tanto referido a la partícula elemental como a  $H^+$ )

Vinagre

Sal de mesa

Alergia al marisco

**N9. NO etiquetar menciones cortas → acrónimos de longitud de 1 letra**

NO etiquetar los acrónimos que son de longitud de 1 letra, excepto si se refieren claramente a elementos químicos (l = yodo); códigos monolettra de aminoácidos (D = Aspartato) o nucleótidos (A, T, R).

Pd/C

N-terminal

Arg-Lys-Phe (RKF)

E(2)

**N10. NO anotar menciones de aminoácidos/nucleótidos que se refieren a mutaciones**

1154C>T (A385V) y 1193T>C (M398T) en los exones

Ser473

Thr-384

**N11. NO etiquetar nombres químicos que forman parte de los nombres de los enzimas**

**Piruvato** quinasa *No anotar piruvato de manera separada*

bomba de **sodio** y **potasio** *No anotar sodio ni potasio*

receptores de **glutamato** *No anotar glutamato*

**glucosa-1-fosfato** uridiltransferasa *No anotar glucosa-1-fosfato*

Etiquetar TODO el nombre del enzima como una única entidad (clase PROTEÍNA).

**N12. NO etiquetar nombres de enfermedades o términos de diagnóstico que se refieran a conceptos químicos (aunque estén asociados al incremento/decremento de un compuesto químico):**

Ferremia *No anotar, aunque indique falta de Fe.*

Acidosis metabólica *No anotar*



GB = glucemia basal

Hipercalciuria

Calcemia iónica

Calcificaciones

déficits vitamínicos

### N13. NO etiquetar células, virus, bacterias...

Plaquetas

Neutrófilos

HIV

## 3.4 REGLAS DE ORTOGRAFÍA Y GRAMATICALES (REGLAS-O)

### O1. Otros idiomas

Los nombres en otros idiomas diferentes del castellano deben ser anotados, independientemente del lenguaje. Especialmente, para el caso de muchos anglicismos / nombres y marcas que están adoptados en el idioma:

Nylon *nombre usado en castellano*

DNA *término que puede ser frecuente (ADN)*

### O2. Errores tipográficos y de conversión

Aquellas menciones de nombres químicos (que sigan las reglas generales de anotación) y que tengan errores tipográficos, deben ser anotados). Esto incluye a las menciones que sufran de errores de conversión automática, generadas con programas de conversión de textos.

C1oro *donde 1 es "uno" no la letra "l"*

1.1 equivalente de Br2en dioxano *donde debería ser "Br<sub>2</sub> en dioxano"*

mometaso-na gotas *donde debería ser mometasona*

### O3. "A B" espacio erróneo

Las palabras separadas por espacios en blanco que deberían ser, de manera correcta, una única mención química → se deben etiquetar como una única entidad.

El grupo acetoxi etilo fue...



#### 04. Nombres químicos tras nombres de personas

Las menciones químicas que incluyen nombres de personas deben etiquetarse si se refieren a estructuras químicas muy claras.

Base de Tröger

Base de Schiff

Reactivos de Grignard

Pero esto aplica únicamente a menciones químicas (no a reacciones químicas):

Síntesis de tiofenos de Gewald

*solamente se etiqueta tiofeno*

#### 05. Límite de sentencias

Las menciones de entidades químicas no pueden atravesar sentencias múltiples.

#### 06. No etiquetar los espacios flanqueantes en blanco

NO etiquetar los espacios en blanco que flanquean una mención química. Los anotadores deberían intentar definir las menciones químicas de la manera más precisa, sin incluir los espacios en blanco u otros espacios que franqueen las menciones (p.ej. puntos al final de una sentencia).

#### 07. No etiquetar comas, puntos de terminación de frase, paréntesis, corchetes...

No incluir como parte de las menciones químicas: comas, puntos de terminación de frases, corchetes, números de referencias, etc, que NO formen parte del nombre químico en sí mismo. NO incluir como parte de la mención química los corchetes que rodean a complejos inorgánicos y líquidos iónicos (únicamente anotar el corchete si éste aparece dentro del nombre):

[Co(CN)<sub>5</sub>]

*no se etiqueta el corchete*

pero:

[Cu(H<sub>2</sub>O)<sub>6</sub>]<sup>2+</sup>

Acetato, bromuro, los nuevos compuestos (aspirina y (carboxialquilo)hidroxipiridina).

Deferiprona (1,2-dimetil-3-hidroxipiridin-4-ona)

[<sup>3</sup>H]

etiquetar corchetes de isótopos



### O8. Incluir los prefijos de la estereoquímica

Incluir dentro de la mención de términos químicos aquellos prefijos que denotan la estereoquímica o regioquímica del compuesto.

cis-metanoglutamato

cis-platino

(S)-Alanina

Ácido (R)-3-(2-(4,4-pentametileno-3-(4-(3-(2-metilfenil)-ureido)-bencil)-2,5-dioxoimidazolidin-1-il)-acetilamino)-3-metil-propiónico

### O9. No etiquetar el símbolo de marca registrada

No etiquetar el símbolo de marca registrada que aparezca como parte de un nombre trivial:

Aspirina®

Mesupron®

A no ser que el símbolo esté anidado dentro del propio nombre:

Tween® 80

Kollidon® VA64

Eudragit® S100

### O10. Etiquetar el nombre de la base de datos si aparece junto al identificador

Para los identificadores de compuestos en bases de datos, anotar el acrónimo de la base de datos si ésta aparece antes o después del identificador:

CID 445154

CHEBI:28262

CAS # 2634-33-5

### O11. No etiquetar signos de puntuación innecesarios

No etiquetar guiones que sirvan para separar conceptos, el apóstrofe en posesivo ó símbolos (e.j. flechas).

R =  $\text{O}$       No anotar "=" cuando significa igual



Pero sí si aportan significado químico:

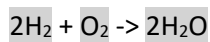
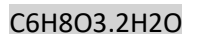
-NH<sub>2</sub>- *notar que ambos guiones deben ser anotados*

**O12. No romper/etiquetar partes de palabras que tienen dentro una mención química: reactividades, enzimas o nombres de proteínas**

Metilante	<i>No etiquetar (reactividad)</i>
Dienófilo	<i>No etiquetar (reactividad)</i>
Acetiltransferasa	<i>No romper</i>
mGluR1alpha, mGluR2	<i>Glu no etiquetar como nombre dentro del receptor</i>

**O13. Números en fórmulas y números como parte del nombre químico**

- Incluir los números de las fórmulas que indican estequiometría:



- Incluir los números que especifican las posiciones de una moléculas únicamente si forman parte del nombre:

Carbono C-2	<i>solo se etiqueta carbono</i>
Posiciones C-2 y C-3	<i>no se etiqueta nada</i>
Posición de sustitución estándar N-1	<i>no se etiqueta nada</i>

Si las posiciones identifican posiciones generales en los compuestos químicos, estas posiciones generales deben anotarse:

Derivado 4-bromo

Sustituyente 5-vinil

Análogos de 5-[2-(1-azirinil)]uracilo

5-viniluracilos

El sustituyente 5-vinil del respectivo 5-viniluracilo.

Análogos 2'-fluoro

Derivado N-metilo

Análogos 5-[2-(1-azirinil)]uracilo

con 5—19 átomos espaciadores entre N6 ó C-8 y el yoduro han sido evaluados





*no etiquetar las posiciones generales N6 y C-8*

- Esta regla de las posiciones generales aplica tanto a las posiciones definidas como números como a las definidas mediante texto (orto, meta, para, o-...).

Análogos de o-nitrofenil

#### **O14. Símbolos de Estado, Carga y Superficie**

Incluir en la mención química los símbolos de estado de oxidación, carga, estado y superficie que se añaden al final de los nombres químicos:

Cu<sup>2+</sup>

Cu(II)

CuSO<sub>4</sub>(ac)

Superficie Au(111)

Isótopo (14)C

#### **3.5 REGLAS MULTIPALABRA (REGLAS-M): ENTIDADES ÚNICAS *VERSUS* ENTIDADES MÚLTIPLES**

**M1. Varios términos consecutivos “Entidad 1 Entidad 2” ó “Entidad1 de Entidad 2” → ¿anotar como una única entidad o como dos entidades diferentes?**

Si aparecen dos o más palabras continuas del tipo: “nombre químico 1 “ y “nombre químico 2” (o separadas por la preposición “de”), cada una de las cuales podrían ser independientemente una entidad por separado:

A) Etiquetar como única entidad si combinadas se refieren a una única entidad específica. Esto es muy común en los siguientes casos:

1) Términos genéricos que siguen la nomenclatura IUPAC:

Acetatos de alquilo

Haluros de isopropilo

Uso de derivados de nucleótidos adenina

Glucósidos limonoides

2) Complejos y compuestos formados por sales / contraiones que definen un único compuesto o aducto:

Citrato de sildenafilo

Cu<sup>2+</sup>· 2H<sub>2</sub>O



Terbinafina-HCl

TBF-HCl

determinación de succinil-acetona en orina

3) Formulaciones combinadas en tratamientos que definen un tratamiento único:

ceftazidima-amikacina

trimetoprima-sulfametoxasol

amoxicilina-clavulánico

amiloride-hidroclorotiazida

• Entidades químicas que forman parte de la otra entidad, típicamente unidas por la preposición de:

Carbono del carbonilo

B) Etiquetar como dos entidades diferentes si se refieren a dos entidades químicas diferentes. Esto es muy común en los siguientes casos:

1) Mezclas de compuestos definidas en los que cada uno de los términos es claramente una entidad separada:

Eluyente se usó hexano-acetato de etilo

Preparación de Pd/C

Reemplazamiento isostérico benceno-tiofeno

Notar que todos los ejemplos de arriba se refieren a casos en los que las menciones aparecen continuas. Si estos términos aparecen separados por otras palabras, se tienen que anotar de manera separada cada entidad.

## M2. Entidades múltiples

Se refiere a menciones de términos químicos que no aparecen descritos en una secuencia continua de caracteres. Es muy común en el caso de menciones de varios términos químicos unidos por cláusulas coordinativas o en enumeraciones de nombres químicos (frecuentemente para eliminar redundancias).

Si se enumeran nombres completos, cada uno de ellos debe ser etiquetado de manera separada:

Ácido cítrico y ácido acético

Carbonato de litio, carbonato de sodio

Hexano – etil acetato, pirano, aspirina / ibuprofeno

Aspirina, azúcar y 4-metoxifenol



Pero si no se describen en una cadena continua de caracteres, se debe etiquetar la cadena total (incluyendo palabras como “y”, “ó” y comas) como una única entidad para evitar la generación de “medias verdades”:

amidas cíclicas de ácidos sulfónicos y de ácidos carboxílicos

ácido acético y cítrico

hidróxido de sodio, potasio y litio

1,2- y 1,3-tiazolilo

Tieno[2,3-d] y tieno[3,2-d] oxazin-4-onas fusionadas

6-Bencil-3-isopropil- or (3-Isopropil)- [1,2,4]triazolo[4,3-a]-piridina

Ácidos piperidinocarboxílicos hidroxilados y amino sustituidos

Vitaminas A-D

Si aparecen símbolos que actúan como separadores (e.j. como comas), cada uno de ellos debe etiquetarse como una única entidad:

dónde R = metilo , etilo , bencilo

Si estos separadores actúan como cualificadores de entidades múltiples, etiquetar la cadena total como una única entidad.

Metil- > Etil- > Propil-amidas

En el caso de anáforas anidadas dentro de menciones MÚLTIPLES, la mención de la anáfora debe mantenerse:

Vitamina A (1) y B (2)

Prostaglandina E1 (PGE1) (1), E2 (PGE2) (2) e I2 (PGI2) (3)

**M3. Una entidad debe reconocerse como la parte más larga que lo define, considerando siempre aquellas palabras que realmente forman parte de la mención (nombre químico total). Las partes no-esenciales de las entidades químicas y modificadores de nombres no se deben etiquetar:**

gas nitrógeno

nanopartículas de oro

grupo metilo

anillo fenilo

sustituyente metilo

análogo de la cafeína



adicción a la **cocaina**  
ciclo del **ácido cítrico**  
Derivados de la **piridina**  
Derivados de **5-(vinil)uracilo**  
Metabolitos de **Bupropión**  
Composición de **Telmisartán**  
Tabletas de **Benzobromanona**  
Compuestos de tipo **piridina**

Esta regla aplica también al anotar menciones múltiples:

Derivados de la **aspirina** y análogos del **ibuprofeno**  
Análogos de la **Viagra** y del **2,3-dihidrobenzofurano**

**M4. Palabras conflictivas: forman parte del nombre de la entidad química o son modificadores?: “Ácido”, “Base”, “Sal”, “Metal”, “Radical”, “Cación”, “Anión”, “Ion”, “Dímero” ( y similares)**

Anotar estos términos si forman parte de un nombre químico:

**Ácido lisérgico**  
**Sal de sodio**  
**Ion sodio**  
**Cación amonio**

O si se refieren a clases explícitas de compuestos:

**Metales de transición**  
**Metales pesados**  
**Óxido de metal**  
**Ácido carboxílico**  
**Sales de hierro**

Pero no etiquetar si aparecen de manera separada, muy generalista:

**Ácido fuerte**  
**Ácido orgánico**  
**Sal orgánica**



En tratamiento con álcali acuosa o ácidos

### M5. Adjetivos con entidades válidas de anotación

Los adjetivos se anotarán únicamente si

- 1) Preceden / continúan una entidad química / biosimilar válida para anotación y
- 2) Añaden información estructural precisa a dicha entidad

En estos casos, el concepto total (adjetivo + nombre químico) se etiquetará como una única entidad. Esto aplica independientemente del nombre raíz del adjetivo (nombres sistemáticos o nombres comunes) o de la terminación del adjetivo (“-ado”, “-ico”, “-ante”, “-olico”).

Bifenilos policlorados

Fenoles acetilados

KOH etanólico

Alcohol alílico

Alcoholes primarios

Hidrocarburos saturados

Óxido nitroso protonado

Hidratos de carbono de cadena corta

Alquilos lineales o ramificados C1-C6

Glucorónido (acetilado)

Grupos hidroxilo secundarios

### M6. No anotar adjetivos con pautas estructurales genéricas o información muy general que no se refiera a una entidad concreta

Éter seco

Plata coloidal

HCl caliente

Nitrógeno molecular

Carbonato sódico acuoso

Clindamicina oral

Amonio sérico



pero:

Hielo seco

*anotar íntegramente porque es  $CO_2$*

Ácido sulfúrico fumante

*anotar íntegramente porque es  $H_2S_2O_7$*

Acero inoxidable

### M8. Adjetivos seguidos de terminos generales: adjetivos sustantivados

En el caso de que los adjetivos acompañen a palabras genéricas químicas que no se anotan (compuesto(s), análogo(s), derivado(s), serie(s)), se anotarán si están sustantivados y añaden información química estructural precisa. Típicamente terminan como “-oide”, “-oico”, “-ado”

Análogos terpenoides

Moléculas orgánicas aromáticas

Espaciador amidico

Compuestos pirazólicos

Análogos etoxilados

Compuestos fluorados

Siempre y cuando la palabra a la que acompañan sea un término químico general – no si está en otro contexto:

Control glicémico

Áreas noradrenérgicas

Dieta hiposódica

**M9. Adjetivos no explícitamente químicos pero que hacen que la entidad química se refiera a un término específico de relevancia en informes clínicos (terapia o diagnóstico) sí se deben anotar.**

T4L = Tiroxina libre      anotar todo el concepto (diferente de tiroxina)

C-LDL = colesterol LDL      anotar todo el concepto (diferente de colesterol)

FiO<sub>2</sub> = fracción inspirada de oxígeno

bilirrubina total: 23,4 mg/dl; bilirrubina indirecta: 22,1 mg/dl

### M10. Adjetivos negativos

Los conceptos negativos que descartan estructuras químicas específicas pero que no definen explícitamente una



estructura química NO se deben anotar.

Compuestos no esteroideos

### M11. Menciones anidadas

Anotar las menciones anidadas como una única entidad

Prostaglandina (PGE1) E1

Ácido acetilénico C(17) bromado

*incluyendo información de la posición*

Proantocianidina (PAC) A2

### M12. Entidades químicas/biosimilares que se solapan con otras entidades no químicas.

Anotar exclusivamente la entidad química / biosimilar de interés.

1. Fórmulas químicas que aparecen en ecuaciones matemáticas (gradiente, concentración...):

$^1\text{H}$  NMR

$d[\text{Na}^+]/dt = x$

[cafeína]=10 mM

Pero no si forma parte de una unidad:

mmHg

2. Compuestos químicos unidos por guiones a conceptos generales:

Oxacilino-resistente

## 4 NORMALIZACIÓN Y CLASES

---

Se han definido cuatro clases de menciones:

1. NORMALIZABLES: aquellas menciones de compuestos químicos y fármacos que se pueden normalizar / estandarizar frente a un identificador único de una base de datos.
2. NO NORMALIZABLES: aquellas menciones de compuestos químicos y fármacos que NO se pueden normalizar
3. PROTEÍNAS: incluye a péptidos, proteínas, genes, hormonas peptídicas y anticuerpos.
4. UNCLEAR: para las plantas, aceites, esencias, principios de plantas y formulaciones/composiciones generales de varios compuestos. Vacunas. Incluye también la lista de términos concretos que sí se van a anotar: Alcohol, Tabaco (cigarrillos), Droga (marihuana, cannabis, opio...), Gluten.



En caso de duda acerca de si la entidad química se refiere a un compuesto único o a una familia de compuestos (por ejemplo “urea”), se tiene que usar el contexto para disambiguar la normalización.

Para normalizar se utilizará de manera jerárquica:

### 1) El recurso Snomed CT: identificador (SCTID)

<http://browser.ihtsdotools.org/>?

(escoger la edición española)

Permite normalizar todo tipo de entidades (incluyendo aquellas que van asignadas a las clases PROTEÍNA y UNCLEAR). Así, aunque una mención, por su naturaleza pertenezca a la clase PROTEÍNA / UNCLEAR, si es normalizable en Snomed, se anotará su identificador (SCTID).

El identificador de Snomed (SCTID) se registrará en el campo comentarios de la herramienta de anotación.

Dentro de las diferentes jerarquías de Snomed (sustancia, prueba, medición...): se escogerán aquellas que se refieran a la sustancia > producto biológico / farmacéutico (con preferencia de sustancia por producto biológico).

Producto biológico / farmacéutico: esta jerarquía representa los medicamentos o productos farmacéuticos en sí, y se diferencia de la de sustancia, con el fin de separar el medicamento de sus componentes. Los conceptos que forman esta jerarquía permiten representar los medicamentos con distintos niveles de detalle.

Sustancia: Contiene conceptos para representar sustancias químicas, componentes de medicamentos, alimentos, agentes causales de alergias, enfermedades, toxicidades

Ejemplos:

- Albendazol como sustancia 387558006
- LHRH como sustancia 49869009
- Para aquellos fármacos de nombre genérico o comercial que no aparezcan como normalizables en Snomed, se hará un esfuerzo en buscar sinónimos posibles antes de ser catalogados como NO NORMALIZABLE.

Adriamicina (NO NORMALIZABLE porque no se recoge en Snomed) = Doxorubicina (NORMALIZABLE, 372817009), luego Adriamicina (372817009)

- Idem si se trata de familias de proteínas:





**Transaminasas** (NO NORMALIZABLE porque no se recoge en Snomed) = **Aminotransferasas** (NORMALIZABLE, 736212005), luego **Transaminasas** (736212005)

- “Anticuerpos Antitransglutaminasa”: Snomed recoge diferentes tipos de anticuerpos de manera separada, pero no el concepto de todos los anticuerpos, por lo que no es normalizable.
- Colesterol como sustancia (colesterol) 84698008
- Triglicéridos como sustancia (triglicérido) 85600001
- Proteína como sustancia (88878007)
- Albúmina como sustancia (52454007)
- Estatina clase que agrupa a varios tipos (96302009)

Respecto a la normalización de vacunas, aquellas vacunas que se encuentren en Snomed sí se normalizarán, pero NO se hará un esfuerzo exhaustivo en normalizarlas (clase UNCLEAR).

## 2) CHEBI para anotar ciertos grupos químicos generales (“hidroxilo”) que no se encuentran en Snomed

- “Hidroxilo” □ “Hydroxyl” □ CHEBI:29191 (anotar todo “CHEBI:29191”).
- “Aminoácido” □ “Amino acid” □ CHEBI:33709 (ya que todas las entradas en Snomed corresponden a procedimientos).
- Fármacos con nombre genérico que no se encuentran recogidos en Snomed (por ejemplo porque tienen grupos químicos generales en su estructura), normalizarlos en ChEBI:

“6-metil-Prednisolona” (no aparece en Snomed, tampoco como “6-metilprednisolona”) → “6-methylprednisolone” → CHEBI:6888

## 5 REFERENCIAS

- 
- [1 Quiñoa Cabana, E.; Riguera Vega, V. Nomenclatura Y Representación de Los Compuestos Orgánicos. Una Guía de Estudio Y Autoevaluación, 2ª edición.; Editorial Schaum, Ed.; 2005.
- [2 Ciriano, M. A.; Román Polo, P. Nomenclatura de Química Inorgánica. Recomendaciones de La



- ] IUPAC de 2005; Zaragoza, E. P. U. de, Ed.; 2007.
- [3 FORMULACIÓN y NOMENCLATURA en QUÍMICA  
] [http://estudiantes.us.es/descargas/Formulacion\\_y\\_Nomenclatura.pdf](http://estudiantes.us.es/descargas/Formulacion_y_Nomenclatura.pdf)
- [4 IUPAC Complete Draft of Chemical Nomenclature  
] [http://www.iupac.org/fileadmin/user\\_upload/publications/recommendations/CompleteDraft.pdf](http://www.iupac.org/fileadmin/user_upload/publications/recommendations/CompleteDraft.pdf)
- [5 IUPAC Nomenclature of Organic Chemistry. ACDLABS  
] <http://www.acdlabs.com/iupac/nomenclature/>

## 6 GLOSARIO DE SIGLAS Y ACRÓNIMOS

---

CNIO	Centro Nacional de Investigaciones Oncológicas
CPI	Compra Pública de Innovación
Plan TL	Plan de Impulso de las Tecnologías del Lenguaje
PLN	Procesamiento del Lenguaje Natural
SEAD	Secretaría de Estado Agenda Digital
TA	Traducción Automática