

COMPONENTES ESTRUCTURALES DE LOS ÁCIDOS NUCLEICOS

BASES NITROGENADAS, NUCLEOSIDOS Y NUCLEOTIDOS

Los ácidos nucleicos: DNA y RNA están constituidos por 3 unidades elementales: ribosa (RNA) o desoxirribosa (DNA), fosfato y bases nitrogenadas. La unión del azúcar y la base se denomina nucleósido y la unión de fosfato, azúcar y base se denomina nucleótido.

NOMENCLATURA

En este punto se repasará la nomenclatura de las unidades estructurales de los ácidos nucleicos: bases nitrogenadas, nucleósidos y nucleótidos.

Las bases nitrogenadas se clasifican en función de su estructura en purinas y pirimidinas (véase Figura 1). Las bases más comunes en los ácidos nucleicos son adenina, timina, citosina y guanina (DNA) y adenina, timina, citosina y uracilo (RNA). No obstante, es común encontrar otras en los ácidos nucleicos, especialmente en el RNA(t). También es posible tener bases alteradas por modificación química de las bases naturales (ej en mutaciones). Por ejemplo, solo de modificaciones oxidativas de las bases de DNA se han descrito más de 100 (véase algún ejemplo en la Figura 2).

Por último es común la presencia de bases no codificantes en nucleósidos y nucleótidos de origen natural que tienen actividades p.ej como antibióticos.

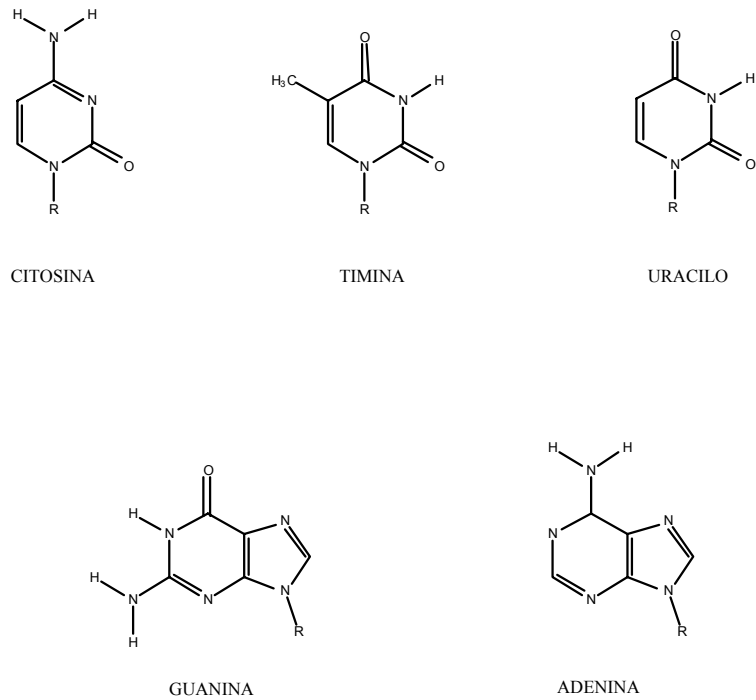


Figura 1: Estructuras de las bases codificantes del DNA (piridiminas (arriba) y purinas (abajo)).

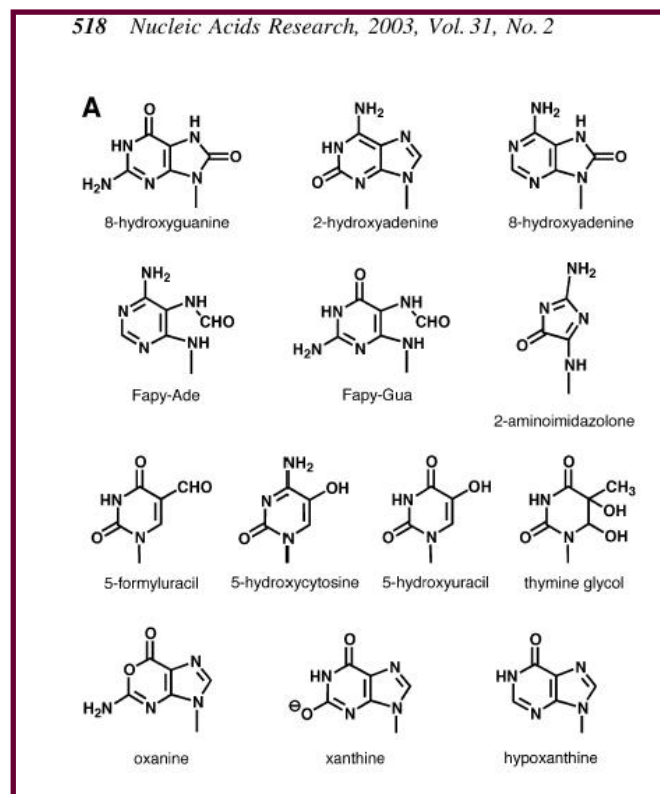
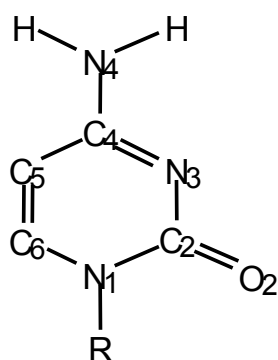
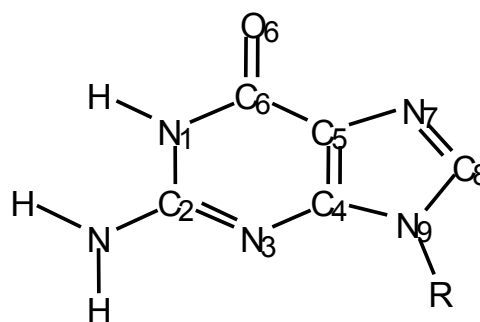


Figura 2. Ejempl,os de oxidaciones comunes de bases de DNA.

La nomenclatura de las bases se basa en la de purinas o pirimidinas. Solo se nombran los átomos del anillo, los sustituyentes sobre él adoptan los mismos números que el átomo al que están unidos (véase Figura 3).



CITOSINA

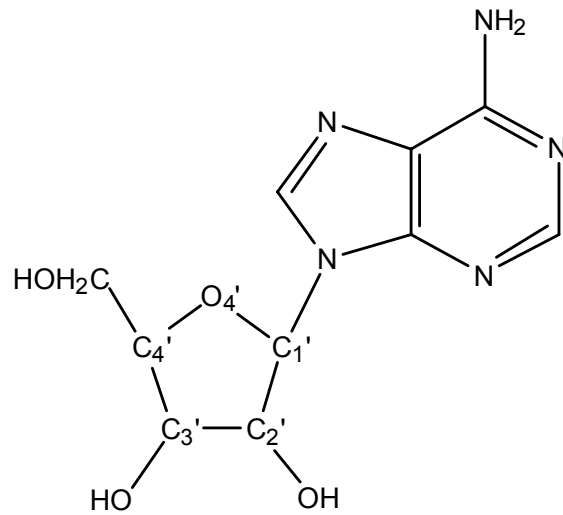


GUANINAA

Figura 3. Esquema de numeración de las bases

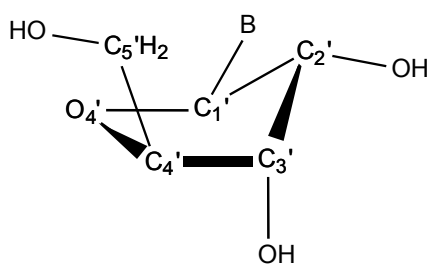
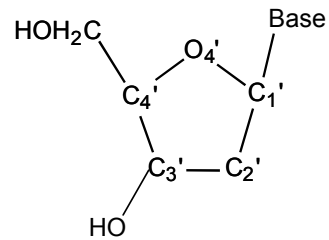
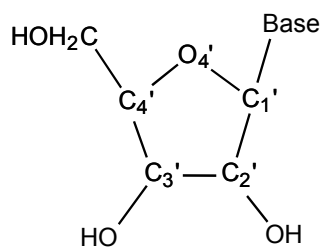
Los nucleósidos más comunes en los ácidos nucleicos surgen como combinación de las bases antes citadas y ribosa (RNA) o 2'-deoxyribosa (DNA). Hay que señalar que es posible, y muy común en fármacos la sustitución de la ribosa o deoxyribosa por otros azúcares.

La nomenclatura de la ribosa comienza por el carbono unido a la base, que se denomina 1' y sigue de tal manera que el carbono unido al grupo CH₂OH es el C4'. Notad que el azúcar siempre se numera con "prima". El C del sustituyente CH₂OH se nombra como 5'. El resto de los sustituyentes sobre el anillo conserva el número del átomo al que está unido. El oxígeno furanósico se numera como O4', pero es común encontrar textos donde se nombre como O1'.

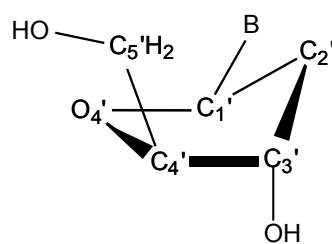


ADENOSINA

Figura 4. Ejemplo de un nucleósido.



RIBOSA



2'deoxy RIBOSA

Figura 5. Azúcares presentes en el RNA (ribosa) y DNA (deoxiribosa)

El grupo(s) fosfato(s) se nombra por la posición del azúcar a la que está unido y por el número de fosfatos esterificados: mono, di o trifosfato.

Los ácidos nucleicos son cadenas de nucleótidos, donde los fosfatos se unen por enlaces esteres a los grupos C3' y C5'. Las cadenas se nomencan a nivel de estructura primaria por la composición de bases. Siempre se nomencla del extremo 5' al 3' (véase 12.4), donde 5' es el extremo que tiene el grupo fosfato libre unido al cabono C5'.

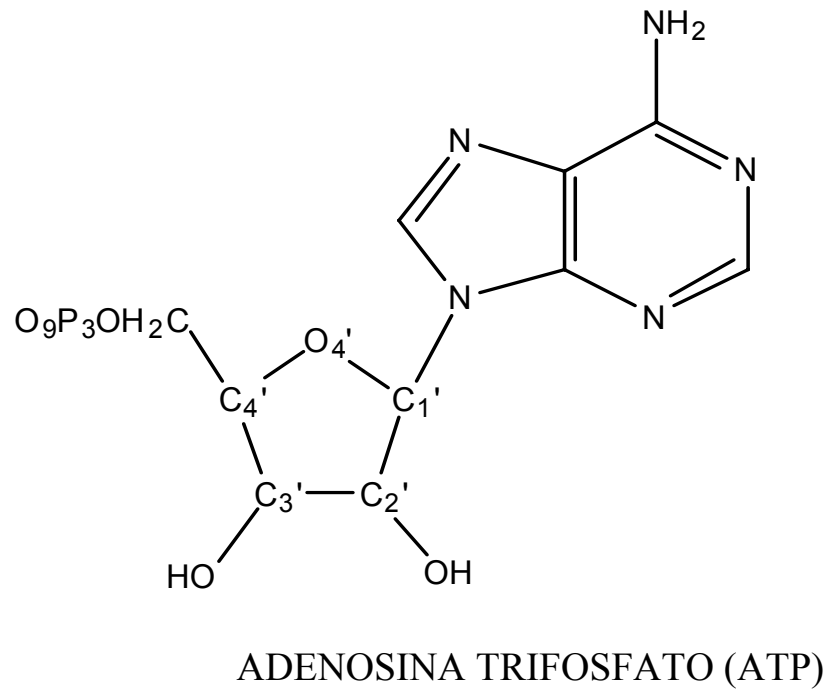


Figura 6. Ejemplo de un nucleótido.

Una pregunta interesante es por que los ácidos nucleicos están formados (en general) por uniones 5'-3' y no por ejemplo 5'-2'?. Realmente es un tema no obvio, que puede tener sus orígenes en el inicio de la vida. En general se sabe que los enlaces 5'-2' son más estables que los 5'-3'. También se ha demostrado que pueden dar lugar a polímeros (5'-2') y que éstos pueden reconocer otros polímeros complementarios (ej. AAAAA reconoce a TTTTT). Algunos autores han sugerido incluso, basándose en datos de calorimetría que estas uniones de cadenas 5'-2' serían incluso más estables que las uniones de cadenas 5'-3'. Cual es entonces la razón por la que in vivo los ácidos nucleicos polimerizan en 5'-3'. Hay varias hipótesis, una es que en un principio existiera otro ácido nucleico que contuviera un azúcar diferente de la ribosa. Otra posibilidad es el primer ácido nucleico fuera un RNA 5'-2', pero después “algo” hiciera

cambiar la preferencia in vivo a 5'-3'. La tercera posibilidad es la de que la unión de polímeros 5'-2' de lugar a estructuras poco eficientes para transmitir la información genética como son triple hélices en lugar de dobles hélices. Datos experimentales de mediados de los noventa apoyan esta tercera hipótesis, aunque dejan aún el campo abierto para discusión

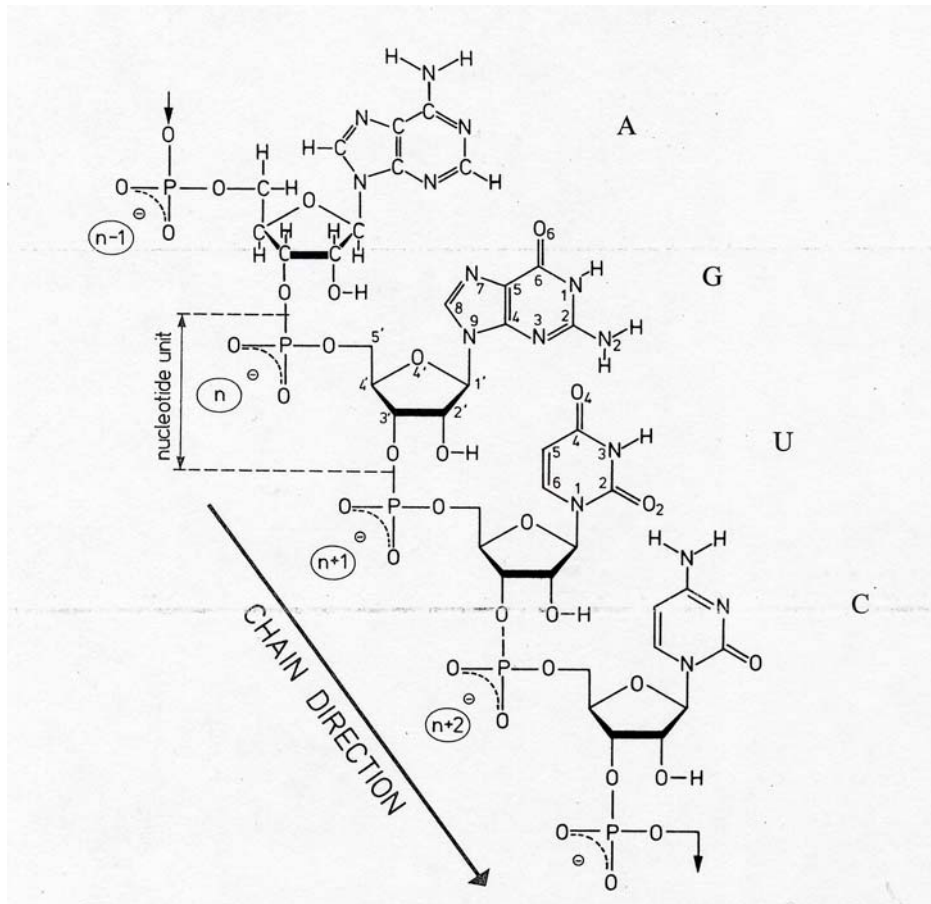


Figura 7. Esquema covalente de una cadena de RNA.

ESTRUCTURA

Las bases son estructuras aromáticas y planas. Los azúcares son anillos furanósicos que adoptan empaquetamientos (puckerings) en los cuales 1 o 2 átomos están fuera del plano formado por el resto. Los enlaces fosfodiéster, los azúcares y los enlaces glicosídico son flexibles mientras que las bases son fundamentalmente rígidas. Una descripción detallada de las geometrías de enlace de estas moléculas se encuentra en la Figura 8.

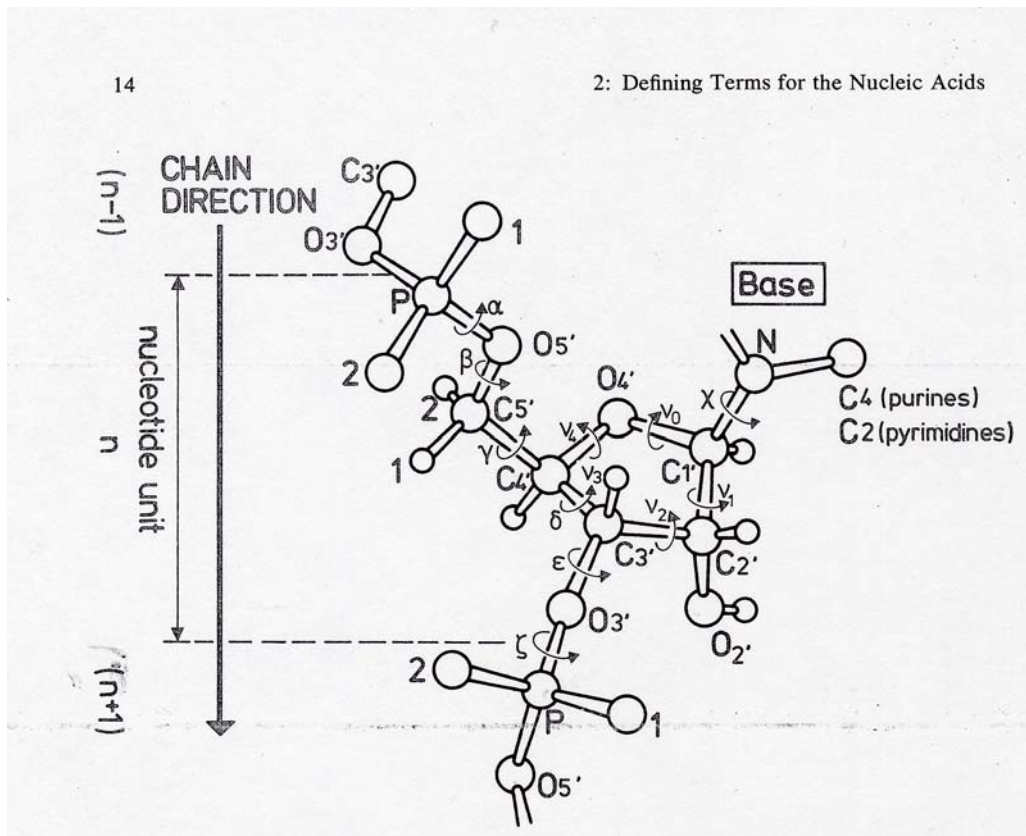


Figura 8. Descripción de la estructura de un nucleótido

PROPIEDADES FISICOQUIMICAS

Las propiedades fisicoquímicas, incluida la reactividad vienen marcadas básicamente por la distribución orbitalica y de carga. En los ácidos nucleicos podemos hablar, en un afán simplificador, de dos comportamientos diferentes. Así, el azúcar y los fosfatos tienen un papel básicamente estructural. Las bases, por el contrario, tienen un papel más funcional, ya que son las que modulan de un modo específico las interacciones que suceden en los ácidos nucleicos. Centraremos pues de un modo especial el estudio de propiedades fisicoquímicas en las bases.

Las bases son estructuras polares, con momentos dipolares que pueden ser elevados y una distribución de carga tal que en ciertos átomos (ej O carbonílicos) se concentra una gran cantidad de carga negativa, mientras que en otros (ej. H unidos a N)

se concentra una gran porción de carga positiva. De hecho, la presencia de una pauta compleja de dadores y aceptores de puente de hidrógeno es lo que confiere a las bases su capacidad de reconocimiento específico de otras bases, de drogas o de residuos aminoácidos. No podemos olvidar que toda la capacidad de interacción no covalente específica del DNA es debida a las bases nitrógenadas.

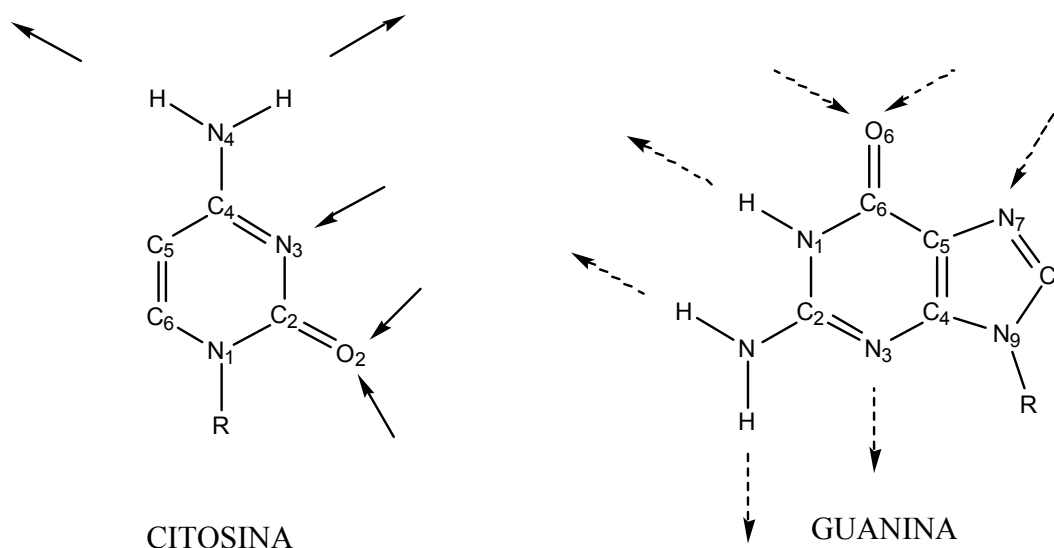


Figura 9. Esquema de dadores/aceptores de puentes de hidrógeno en

Además de esta capacidad de interaccionar no-covalentemente los componentes estructurales del DNA pueden sufrir una gama importante de reacciones químicas.

Protonación:

Hay muchos grupos ionizables en un nucleótido. Así los grupos fosfatos pueden estar en diferentes estados de ionización, las bases también pueden encontrarse en diferentes estados de ionización, y por último en condiciones de alcalinidad extremas hasta los grupos hidroxilo de las ribosas pueden perder un protón (véase Figura 10). Una vez más, las protonaciones/desprotonaciones en las bases son las reacciones de mayor trascendencia funcional (véase Figura 11). La importancia de la reacción está a 2 niveles. Primero a nivel enzimático la protonación de las bases es clave en numerosas

reacciones. Por otro lado, la protonación de ciertas bases es clave para explicar unos tipos de estructura de los ácidos nucleicos. Así la protonación de la citosina parece clave para la estabilización de triple hélice en fragmentos GCC como veremos más adelante. Por otro lado, se ha descrito que la ionización de las bases puede estar detrás de determinados procesos mutagénicos y formas ionizadas de por ejemplo uridina o adenosina se han relacionado con la estabilidad de formas inusuales del DNA.

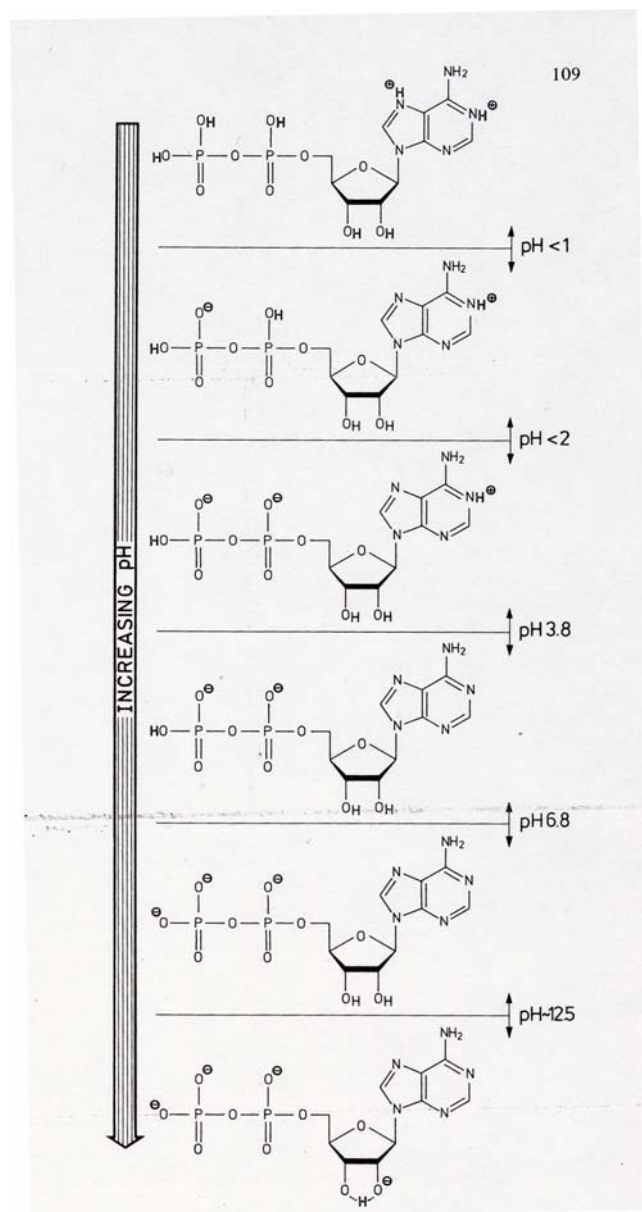


Figura 10. Pautas de protonación de un nucleótido a diferentes pH.

Comp/site	Nucleoside	3'P Nucleot	5'PNucleot
Adenosine/N1	3.52	3.70	3.88
Cytidine/N3	4.17	4.43	4.54
Guanosine/N1	9.42	9.84	10.00
Uridine/N3	9.38	9.96	10.06
Deoxythymidine/ N3	9.93	--	10.47

Figura 11. pKas de las bases, nucleósidos y nucleótidos.

Ataques nucleofílicos

El ataque de grupos nucleofílicos (usualmente con carga negativa) a las bases puede dar a priori lugar a 2 reacciones mayoritarias que compiten: i) abstracción de un protón (normalmente unido a un N), y ii) sustitución nucleofílica. La primera es una reacción en el plano, mientras que la segunda es una reacción donde el nucleófilo ataca en el plano perpendicular a la molécula. Las reacciones nucleofílicas son claves en la vida celular, por ejemplo, la primera etapa del proceso de degradación de nucleósidos no deseados es la sustitución nucleofílica del mismo. Existen un gran número de enzimas dedicados a esta labor, ejemplos son la glutathion S-transferasa y la Adenosine Deaminase. En tanto que esta reacción ocurre catalizada enzimáticamente en un gran número de ocasiones, no existen reglas absolutas de reactividad. No obstante en general los carbonos aromáticos substituidos por oxígenos, o por amino son las dianas preferidas de los nucleófilos. También es común que ataques nucleofílicos sean etapas iniciales en procesos de generación de formas mutadas de las bases.

Hidratación

Las bases nitrogenadas tienden a agregarse en agua por interacciones de stacking como veremos más adelante. De hecho es difícil disolver algunas de ellas porque tienen preferencia a estar en forma sólida donde interacciona con otras nucleobases por redes de stacking y puente de hidrógeno. No obstante, las bases interaccionan muy bien con el agua. De hecho, experimentalmente es muy difícil de medir la energía libre de hidratación de las bases porque es demasiado negativa (hidratación muy favorable). Las únicas aproximaciones a los valores de las energías libres de solvatación se han obtenido por cálculos. A pesar de la dispersión de los datos experimentales, parece claro que las bases tienen energías libres de solvatación muy negativas. La timina es la más hidrofóbica y tiene energías libres de solvatación un poco inferiores a las 10 kcal/mol, mientras que la guanina es la que mejor interacciona con el agua y tiene unas energías libres de solvatación de más de 20 kcal/mol. Esta diferente polaridad tiene mucha importancia en la capacidad de las bases para reconocer a otras bases en solución acuosa.

Tautomerismo

Es un fenómeno que afecta a las bases y que cambia totalmente la capacidad de formar reconocimientos específicos de las bases. De hecho, Watson & Crick no pudieron resolver la estructura del DNA hasta que gracias a un químico teórico (J. Donohue) les sugirió que (en contra de lo que se pensaba en aquel momento) probasen a realizar los apareamientos de las bases con las formas keto-amino. Hoy sabemos que el código genético se mantiene en tanto que los tautómeros keto-amino son más estables que los enol-imino. El cambio tautomérico altera totalmente el pattern de puentes de hidrógeno de las bases, y por lo tanto invierte cualquier tipo de afinidad de las mismas. Como consecuencia, cambios tautoméricos pueden alterar totalmente la reactividad y funcionalidad de las bases (véase la Figura 12).

En principio, podemos hablar de 2 tipos diferentes de tautomerismo: i) el que involucra a átomos del mismo tipo (ej. el N9-H N7-H de la adenina, o el tautomerismo amino-imino de citosina; ii) el que involucra átomos de diferente tipo (p.ej que un H

unido a un N pase a unirse a un O como es el caso de la tautomeria keto-enol de la timina). Fijaos que si se produce un cambio tautomérico en la citosina como el que se muestra arriba de la Figura 12 ésta adopta un esquema de dadores/aceptores en todo igual a la de la citosina, por lo que puede producirse una “transduction” (cambio de purina a purina) del tipo $G \rightarrow A$ en la hebra complementaria (el cambio de una Purina por una pirimidina, o viceversa, se denomina transversion).

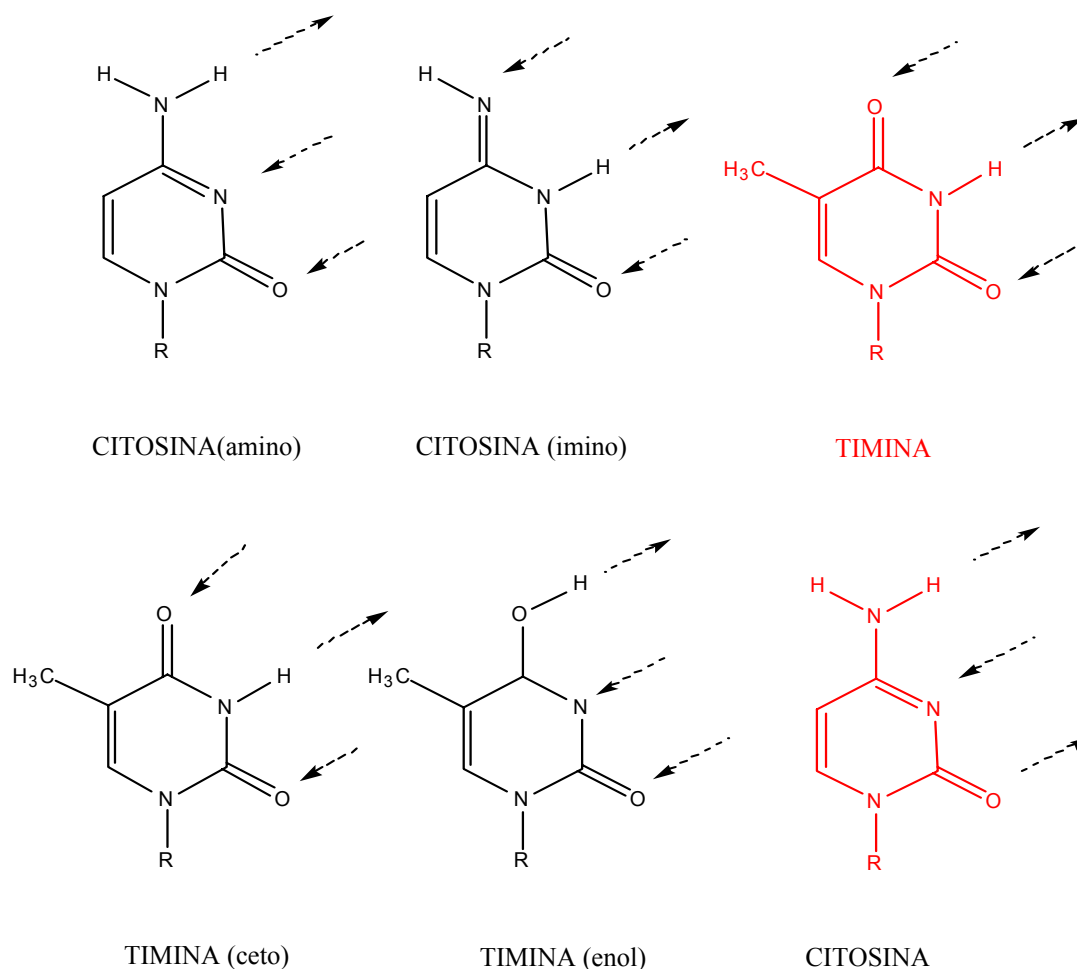


Figura 12. Esquema de cambios tautoméricos y su cambio asociado en el esquema de posibles puentes de hidrógeno. Nótese que debido al cambio tautomérico se pueden confundir Pur y Pyr.

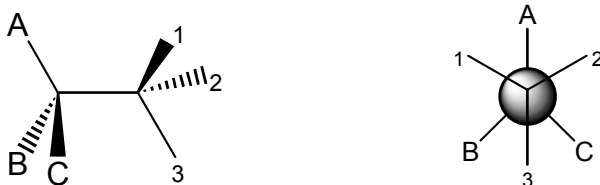
Sabemos hoy tanto por técnicas teóricas como experimentales que las formas keto-amino son las formas más abundantes de las nucleobases codificantes, pero no es siempre el caso de las bases inusuales o no codificantes. En general se ha visto como

para la preferencia de las nucleobases codificantes por los tautómeros keto-amino está marcado por un balance de términos intrínsecos como ligados a la polaridad del medio.

CONFORMACIÓN

Las bases nitrogenadas son estructuras planas y por lo tanto con escasa o nula flexibilidad conformacional por si mismas. La flexibilidad conformacional viene dada por: i) movimientos de rotación respecto al enlace glicosídico (el que une la base y el azúcar), ii) la rotación respecto al enlace C4'-C5', iii) las rotaciones respecto a los enlaces en el grupo fosfato, y iv) los movimientos conformacionales del anillo furanósico.

Los movimientos de rotación respecto a los diferentes enlaces se nombran por el ángulo dihedro, que marca la torsión respecto al enlace en cuestión (véase T12.11) siempre en sentido de las agujas del reloj positivo.



Ejemplo dihedro 1-X-Y-A +60 grados

Las definiciones de los distintos ángulos de torsión se muestran en la Figura 8. Si bien hay una nomenclatura universal para describir las torsiones (la propia de la Química Orgánica), las más importantes para los nucleótidos tienen una nomenclatura propia que es la que discutiremos.

Rotaciones respecto al enlace exocíclico C4'-C5'

Aquí se emplea a menudo la nomenclatura “oficial” pero más aún la nomenclatura tradicional. En ella se mira la posición del O5' respecto al O4'(O1'

dependiendo de la nomenclatura). Si el O5' está entre O4' y C3' en una proyección de Newman es gauche gauche (gg), si está trans al C3' pero al lado del O4' se denomina gauche trans (gt), y si está trans al O4' pero al lado del C3' se denomina trans gauche (tg).

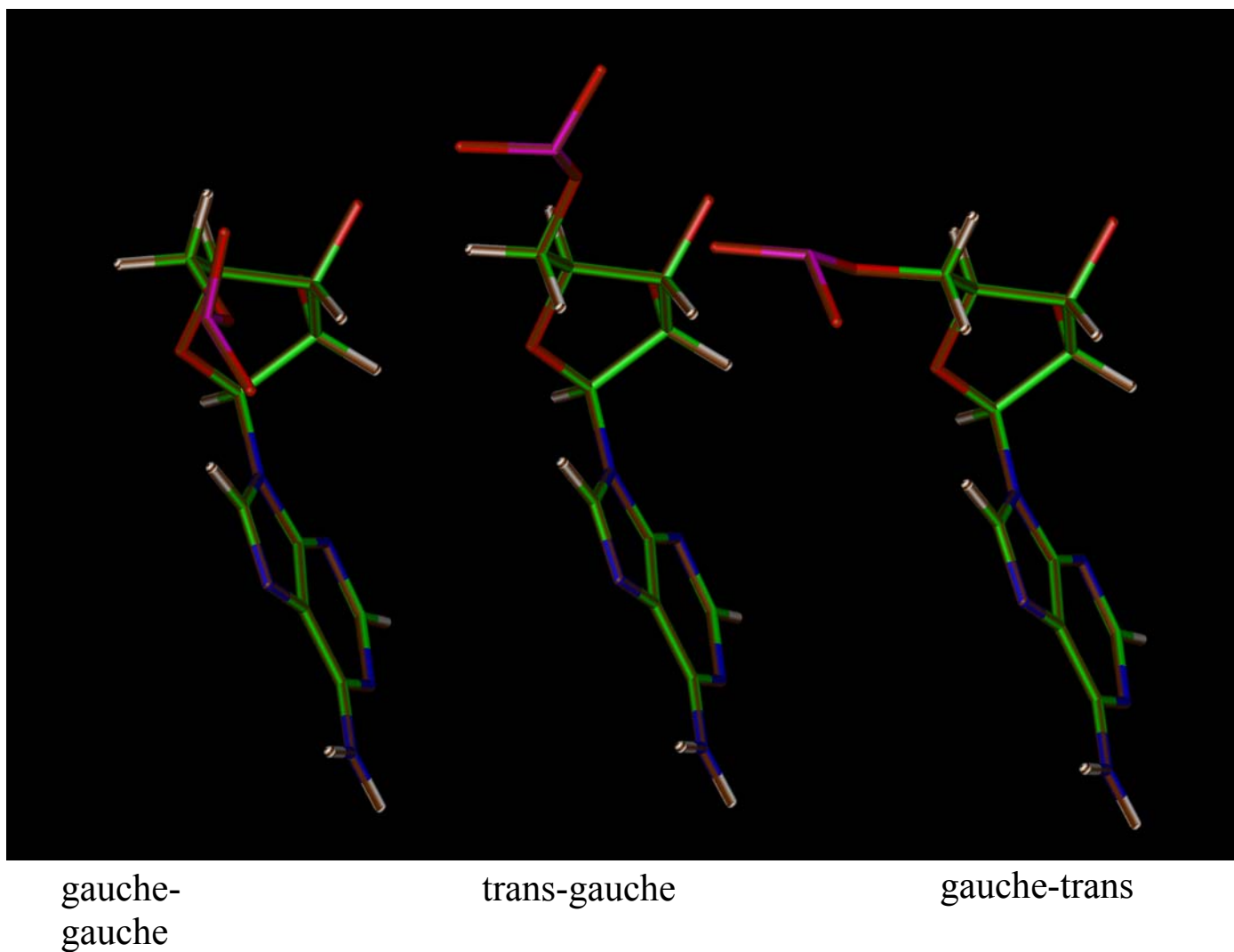


Figura 13. Las tres conformaciones respecto al enlace exocíclico.

La rotación exocíclica es de una moderada importancia para las características conformaciones generales de los nucleósidos libres (si bien puede ser clave en la determinación de reactividad enzimática), pero clave para posicionar correctamente los fosfatos en nucleótidos, y especialmente en ácidos nucleicos.

El enlace C4'-C5' se encuentra mayoritariamente en las 3 conformaciones alternadas: gg, gt y tg. Existe una gran flexibilidad conformacional debido a la muy baja

barrera existente entre estos conformeros. En general la conformación gg es la más abundante tanto en purinas como en pirimidinas y para cualquier tipo de puckering. En pirimidinas las conformaciones gt y tg se dan solo cuando la ribosa es C2'-endo y aparecen con menos abundancia que la gg. En purinas tg se da solo cuando la ribosa es C2'-endo, pero gt es bastante abundante tanto en zona N como S.

Rotación en los enlaces fosfatos C-O y P-O

En cristales se ha visto que en mono- oligo- y poli-nucleótidos el enlace C5'O5' (β) (suele estar en el rango +ap (entorno a 180°). El enlace C3'O3' (dada por ϵ) suele adoptar conformaciones en zona anti, más concretamente -ac (alrededor 220°). Existe una cierta correlación en nucleósidos y nucleótidos entre el puckering (véase más adelante) del azúcar en C3'-endo y el desplazamiento de ϵ hacia ap (alrededor 180°), mientras que si es C2'-endo se da desplazamiento hacia zona -sc. En general, en polinucleótidos solemos encontrar β entorno a $180 \pm 40^\circ$ y ϵ en el rango 180-270°.

Respecto a los enlaces P-O son dos los dihedros que los definen $\chi(\text{PO3'})$ y $\chi(\text{PO5'})$. Para los dos la zona -sc (El primero se suele encontrar en la zona -sc (-gauche) es la más popular, si bien el ángulo χ se presenta también a menudo en conformación -ap.

Rotación respecto al enlace glicosídico

Es la que marca la posición del azúcar respecto a la base. Las conformaciones syn son aquellas en las que la ribosa está eclipsando la base nitrogenada y anti las que el azúcar está en posición opuesta a la base. Posiciones intermedias se conocen como zonas high anti y high syn.

La conformación preferida suele ser la anti frente a syn. La razón de ello estriba en gran medida en la menor repulsión estérica entre la base y la ribosa existente en la conformación anti. La preferencia por anti puede revertirse añadiendo un grupo voluminoso en posición 8. Por ejemplo adenosina es anti fundamentalmente, pero por

contra 8-Br adenosina aparece fundamentalmente en conformación syn. Otros factores que pueden influir en el equilibrio conformacional son los puentes de H, por ejemplo un pte de H entre el grupo O5' y los nitrógenos de las bases puede estabilizar mucho la conformación syn. Por otro lado, la interacción C(8)-H....OH(5') parece estabilizar en cierta medida la conformación anti.

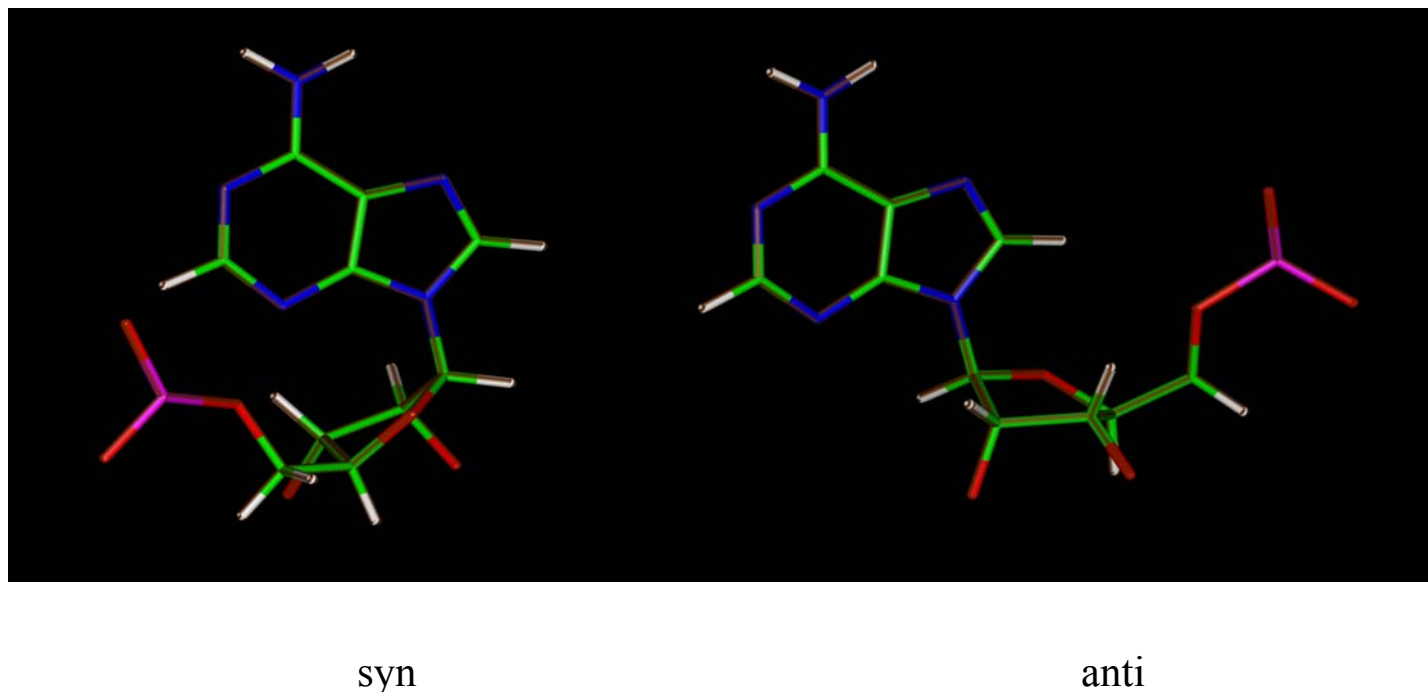


Figura 14. Esquema de adenosina en la conformación anti y syn.

El análisis de estructuras cristalinas depositadas en el Cambridge Data Bank (JACS 112, 8291 (1990)) revela que los nucleósidos se agrupan en 2 familias. Los syn (minoritarios) presentan el ángulo de torsión glicosídico entorno a $49^\circ(\pm 13)$, los anti aparecen entorno a $210^\circ(\pm 23)$. La diferencia energética entre syn y anti es pequeña. Estudios teóricos señalan una diferencia entre syn y anti en el rango 1-2 kcal/mol, los datos de NMR son coherentes con esta diferencia. La barrea anti- $\rightarrow$ syn no es conocida, los datos experimentales sugieren una barrera cercana a las 6 kcal/mol. Cálculos teóricos sugieren una barrea (localizada en zona high syn) de unas 4 kcal/mol. En cualquier caso, la barrera es notable pero no impide el cambio conformacional entre syn y anti. Comentar por último que el análisis de los datos cristalográficos, cálculos teóricos y datos de NMR demuestran que existe una correlación entre la conformación adoptada respecto al enlace glicosídico y el puckering de la ribosa.

Conformación del anillo furanósico

Los anillos furanósicos no se encuentran planos, sino que adoptan conformaciones fuera del plano. Básicamente existen conformaciones tipo envelope (E) y tipo twist (T). En general las formas E son más estables que las formas T.

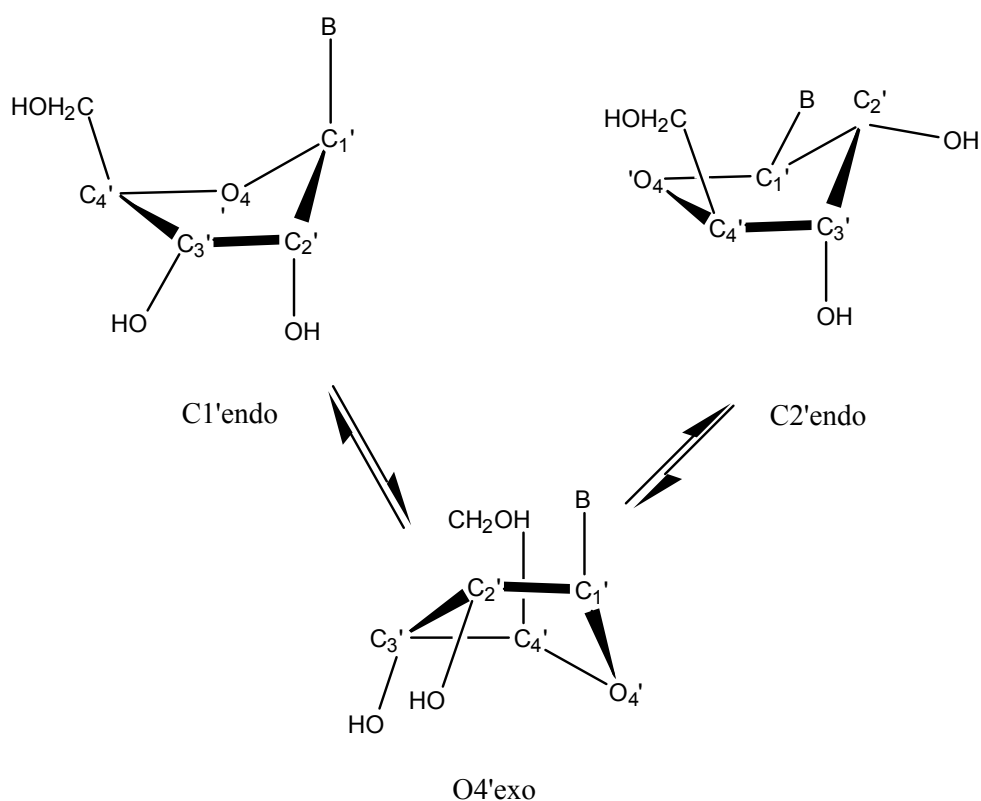


Figura 15. Esquema de 3 tipos de conformaciones E de la ribosa.

La conformación del anillo furanósico es compleja, ya que al existir una restricción de enlace las rotaciones respecto a los distintos enlaces se producen de una forma más o menos concertada, y no de modo quasi-independiente como podría ser el caso de una cadena lineal. La descripción matemática de la conformación de las ribosas fue debida a Altona (entre otros) y se denomina pseudorrotación. Según este modelo podemos definir la conformación de un anillo de cinco miembros a partir del ángulo de fase (P) y de la amplitud de puckering (τ_m). El primero nos da información sobre que átomo o átomos están fuera del plano formado por el resto. El segundo parámetro nos dice que tan fuera del plano están el o los átomos.

Tanto el ángulo de fase como la amplitud de puckering se expresan en función de los ángulos de torsión del anillo furanósico según una expresión analítica compleja. El cambio conformacional se produce normalmente (véase más adelante) a través del denominado “círculo pseudorrotacional” ya que las formas planas son muy inestables energéticamente (véase más abajo). Según el mismo los cambios de conformación ocurren por variación exclusiva del ángulo de fase P , mientras que la amplitud de puckering permanece constante. La amplitud encontrada en cristales y en cálculos teóricos suele estar en el rango $38-40^\circ$. Al ser la amplitud de puckering constante es posible nombrar cualquier conformación de la ribosa según su posición en el ciclo pseudorrotacional.

En el ciclo pseudorrotacional podemos en primer lugar distinguir 4 zonas: norte (N, $P \approx 0$), sur (S, $P \approx 180$), este (E, $P \approx 90$) y oeste (W, $P \approx 270$). Dentro de esta zona es posible encontrar toda una gama de conformaciones. Estas se nombran o por su ángulo de fase, o por dos criterios clásicos. En ambos se mira la posición de los átomos que se encuentran fuera del plano formado por el resto. Se define una dirección “arriba” como aquella en la que está la base y el $C5'$ y la posición “abajo” como la opuesta.

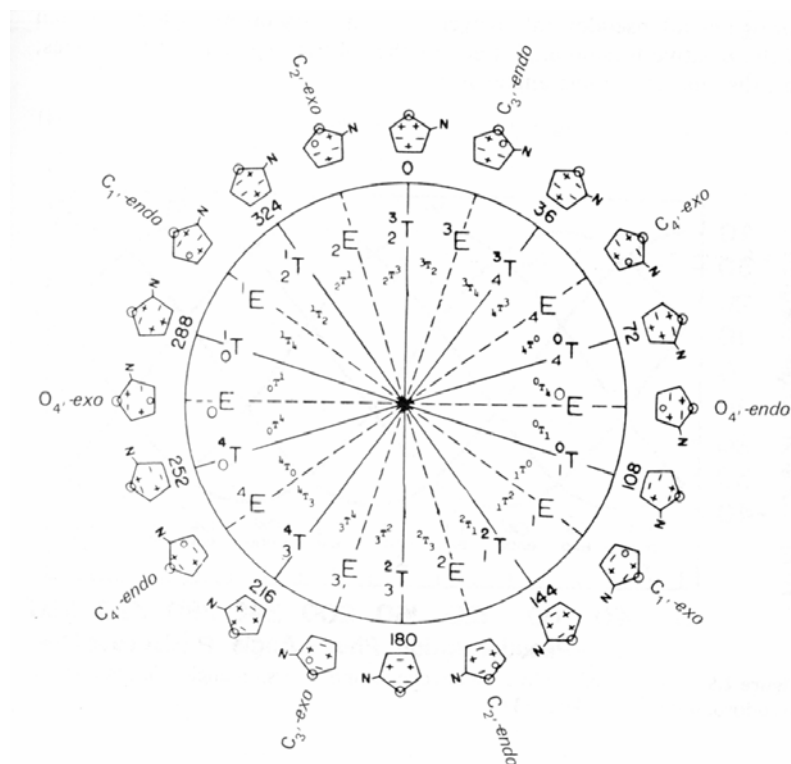


Figura 16. Ejemplo del círculo pseudorrotacional.

- E o T para denominar si es conformación envelope o twisted un superíndice, y un subíndice. El superíndice indica que el átomo que figura en él esta “hacia arriba”. El subíndice quiere decir que este átomo esta “hacia abajo”. Así por ejemplo Ej 2_1T sera una conformación twisted con el átomo 2 hacia arriba y el 1 hacia abajo.

- Endo, exo. Es la más antigua y la más popular. En ella se denomina “endo” si el átomo que está fuera del plano está “hacia arriba”. Se denomina “exo” si dicho átomo se encuentra “hacia abajo”. Por ejemplo, C2’endo (esto es una conformación envelope), o C2’exo-C3’endo (esta es twisted).

Como veremos más adelante, el puckering adoptado por la ribosa es clave para explicar las diferentes conformaciones adoptada por los ácidos nucleicos, y también para explicar la susceptibilidad a la acción enzimática de algunos nucleósidos.

El análisis de datos de NMR, datos de difracción de rayos X, y cálculos teóricos coincide en señalar:

- i) Las conformaciones N y S son las más favorecidas. Especialmente estables son los confórmeros C3’-endo y C2’-endo.
- ii) Las barreras energéticas son pequeñas, lo que explica la interconversión entre uno y otro confórmero a temperatura ambiente. De hecho este cambio conformacional es de los más rápidos del DNA (junto a los que involucran fosfatos).
- iii) Los conformeros tipo “envelope” suelen ser más estables que los “twist”.
- iv) Las transiciones N-S no se dan por un intermediario plano, que corresponde a un máximo energético muy desfavorecido

- v) Las transiciones N-S se dan por zona E siempre (conformación O4'-endo) y nunca por la zona W, que se corresponde también a una conformación muy inestable.

Que la conformación plana sea inestable no sorprende ya que en ella todos los enlaces están eclipsados. Más sorprendente es que el área W sea inestable. Al mirar modelos de ribosa queda claro el porque: las interacciones desfavorables entre el nitrógeno de la base y el carbono C5' de la ribosa.

- El paso entre formas Envelop es traves de formas Twist
- NUNCA se pasa por la conformación plana
- NUNCA se pasa por la conformación W

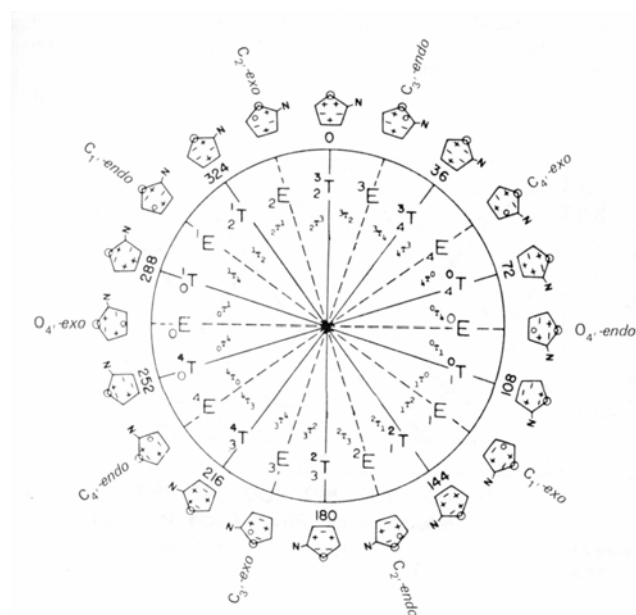


Figura 17. Esquema general y relgas de pseudorrotación.

El análisis detallado de conformaciones en cristal, NMR y cálculos teóricos (JACS 112, 8221 (1990)) revela que la conformación en zona S suele ser C2'-endo con un ángulo de fase sobre 160-164 grados; la conformación en zona N es C3'-endo con ángulo de fase sobre 15-20 grados. Las barreras energéticas dependen un poco del tipo de anillo furanósico, y del nucleósido. Para la adenosina se ha visto que la conformación S es más estable que la N por .5 a 1 kcal/mol, y que la barrera (en O4'-endo) es de 2 a 3 kcal/mol.

Las ribosas y deoxyribosas tienen preferencias conformacionales similares. Añadir grupos diversos a los anillos furanósicos puede introducir modificaciones conformacionales en el anillo furanósico, pero no es fácil de predecir cuáles serán estos cambios. También puede cambiar el puckering si cambia la estereoisomería del mismo. Comentar por último que existe una conexión entre la conformación del anillo furanósico y la conformación respecto al enlace glicosídico. Así, en conformación anti (glicosídico) se pueden encontrar tanto conformación S como N, por el contrario en el conformero syn solo se encuentra el conformero S.

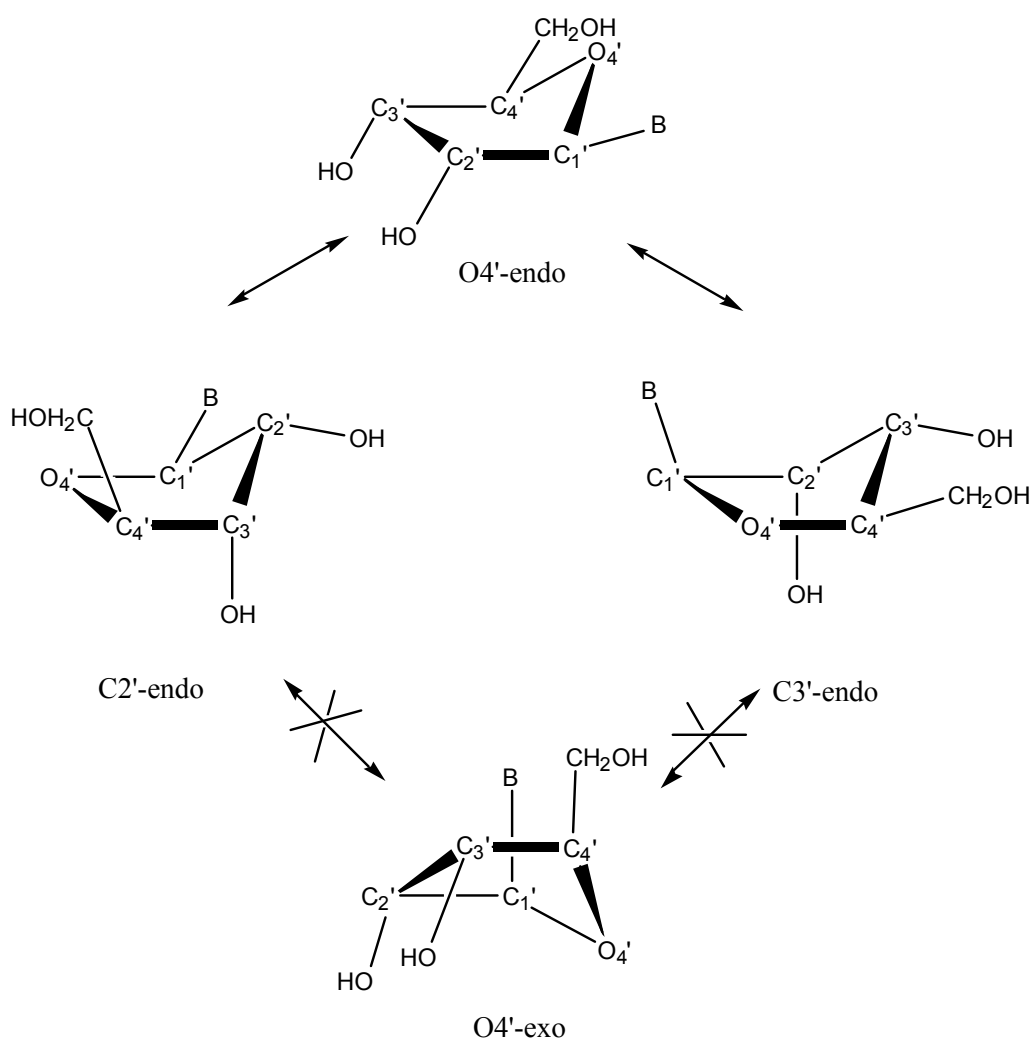


Figura 18. Esquema de una transición N/S a través del círculo pseudorrotacional por la región Este (O4'endo) y Oeste (O4'exo)