

TUTORIAL: USO DE LA BASE DE DATOS RCSB PDB Y VISUALIZACIÓN DE MOLÉCULAS BIOLÓGICAS USANDO EL PROGRAMA AVOGADRO

En este tutorial aprenderemos a buscar estructuras en la base de datos RCSB PDB y a visualizar diferentes moléculas biológicas, viendo cómo el programa Avogadro nos puede ayudar a visualizar su estructura tridimensional y algunas de sus propiedades.

PARTE I: El repositorio de estructuras tridimensionales RCSB PDB

Existe una base de datos, llamada RCSB PDB (RCSB Protein Data Bank). En esta página web se pueden encontrar estructuras de muchísimas proteínas, ADN, ARN y otras estructuras como cápsides virales. Estas estructuras fueron obtenidas por diferentes grupos de investigación, y se encuentran disponibles al público. Cada molécula se identifica por un código de 4 letras/números. La dirección para acceder es:

<http://www.pdb.org/pdb/home/home.do>

Actividad 1: Obtener el archivo de una estructura de ADN conocida del RCSB PDB (código 1BNA).

Acceda a la página y tipee el código 1BNA. Para ello, tipee en la ventana de búsqueda el código correspondiente. Descargue el archivo denominado "PDB file" en formato texto. Este archivo tiene extensión .pdb y es el que vamos a abrir en el programa Avogadro.

The screenshot shows the RCSB PDB website interface. At the top, there's a search bar with the text "Search | All Categories: e.g., PDB ID, molecule name, author". Below the search bar, there are tabs for "Summary", "PDB-101", "Sequence", "Annotations", "Seq. Similarity", "3D Similarity", "Literature", "Biol. & Chem.", "Methods", "Geometry", and "Links". The main content area displays the title "STRUCTURE OF A B-DNA DODECAMER. CONFORMATION AND DYNAMICS" and the PDB ID "1BNA". Below this, there's a section for "Primary Citation" and "PubMed Abstract". On the right side, there's a "Display Files" dropdown menu with options: "FASTA Sequence", "PDB File (Text)", "PDB File (gz)", "mmCIF File", "mmCIF File (gz)", "PDBML/XML File", "PDBML/XML File (gz)", "Structure Factor (Text)", "Structure Factor (gz)", and "Biological Assembly (gz) (A)". A red arrow points to the "PDB File (Text)" option, labeled "3- obtener el archivo PDB file (Text)". Another red arrow points to the "Display Files" dropdown, labeled "2- descarga de archivos". A third red arrow points to the search bar, labeled "1- ventana de búsqueda".

Actividad 2: Obtener el archivo de la estructura de la insulina humana del RCSB PDB sin conocer el código. Si uno no conoce un código, pero desea por ejemplo obtener la estructura de la insulina humana, puede hacer una búsqueda por palabras clave en la ventana de búsqueda. Pruebe hacer esto y obtener una estructura de la insulina humana. Comience tipeando la palabra "insulin" en la ventana de búsqueda. Qué ocurre? Cuántas estructuras o "hits" aparecen?

"hits" de la búsqueda: número de estructuras encontradas

402 Structure Hits 29 Unreleased Structures 212 Citations 154 Ligand Hits 69 Web Page Hits

Query Parameters: Query Details | Save Query to MyPDB

Text Search for: insulin

Query Refinements parámetros para refinar la búsqueda Show

Organism Taxonomy Exp. Method X-Ray Resolution Release Date

Polymer Type Enzyme Classification SCOP Classification

Refine Query Remove Similar: Select Percent Similarity

Cómo podemos acotar la búsqueda para obtener la estructura deseada? Explore las herramientas para refinar la búsqueda: intente seleccionar el organismo "homo sapiens". cuantos "hits" tiene ahora? Ahora restrinja a las estructuras de mejor resolución seleccionando "less than 1.5" dentro de la opción "X-Ray resolution". Así la búsqueda se reduce a unas pocas estructuras (11) de las cuales las primeras tres corresponden a la insulina humana (por ejemplo la primera: código 4EY1). Seleccione esta estructura cliqueando en el título "Human Insulin" de esta estructura y descarguen el archivo "PDB file" (extensión .txt) correspondiente como se explicó en la Actividad 1.

RCSB PDB Query Results

Showing 1 - 11 of 11 Results Results: 25 Page: 1 of 1

Filter: Check All View: Detailed Reports: Select one Sort: Release Date

4EY1 Human Insulin
 Authors: Favero-Retto, M.P., Palmieri, L.C., Lima, L.M.T.R.
 Release: 2013-05-01 Classification: Hormone
 Experiment: X-RAY DIFFRACTION with resolution of 1.47 Å Residue Count: 102
 Compound: 2 Polymers | Display Full Polymer Details | Display for All Results | 2 Ligands | Display Full Ligand Details | Display for All Results |
 Citation: PubMed ID is not available.
 Molecule of the Month: Insulin

4EY9 Human Insulin
 Authors: Favero-Retto, M.P., Palmieri, L.C., Lima, L.M.T.R.
 Release: 2013-05-01 Classification: Hormone
 Experiment: X-RAY DIFFRACTION with resolution of 1.47 Å Residue Count: 102
 Compound: 2 Polymers | Display Full Polymer Details | Display for All Results | 2 Ligands | Display Full Ligand Details | Display for All Results |
 Citation: PubMed ID is not available.

De este modo, podemos utilizar esta base de datos para buscar estructuras de muchas moléculas biológicas, incluyendo proteínas, ADN, ARN, drogas y toxinas, entre otras. Existen actualmente 76.669 estructuras depositadas en la base de datos, y se pueden encontrar muchas cosas interesantes, búsquelas!

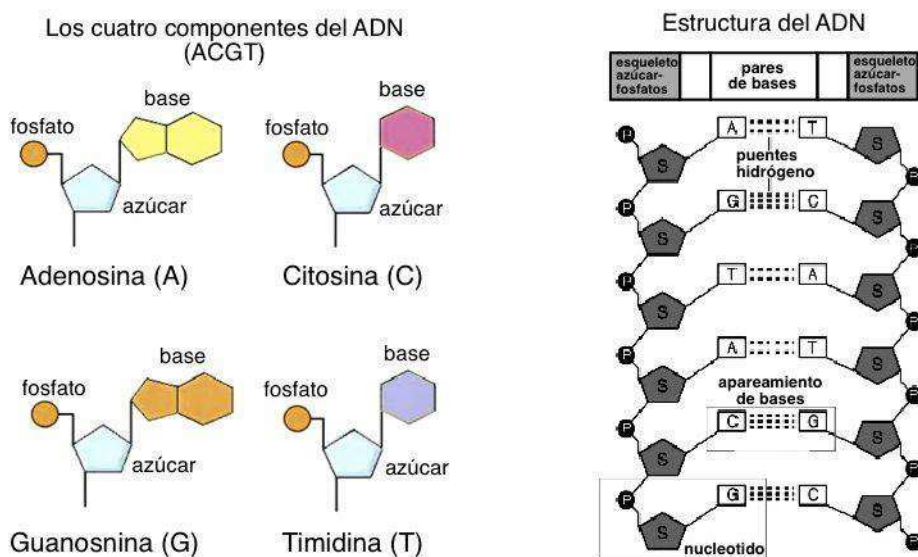
PARTE II: Uso del programa Avogadro para visualizar macromoléculas biológicas

En esta parte del tutorial, vamos a trabajar con archivos .pdb obtenidos del RCSB PDB.

A- MIRANDO ADN

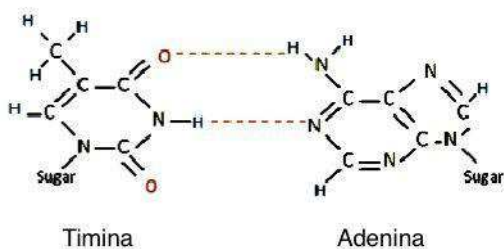
Introducción

El ADN (ácido desoxi-ribonucleico) es una molécula formada por muchas unidades de "ladrillos" llamados nucleósidos-fosfato. Es decir, que el ADN es un polímero de nucleótidos. Existen cuatro tipos de nucleótidos, y en la figura de la izquierda se ve la estructura cada uno. Cada nucleósido-fosfato se compone de una base nitrogenada, un azúcar y un fosfato. Hay cuatro tipos de bases en el ADN (ver figura). En la figura de la derecha se observa la organización de estas unidades en la molécula de ADN.

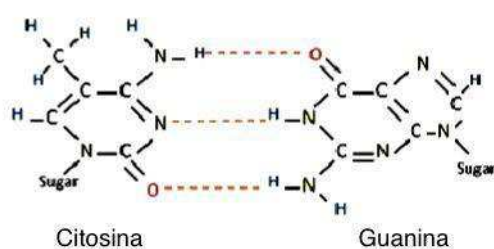


El "apareamiento de bases" es fundamental para la transmisión del código genético desde el ADN al ARN, miremos en más detalle cómo ocurre este apareamiento:

Apareamiento A-T



Apareamiento C-G



En la primera figura, reconocer en la molécula de ADN de la derecha las unidades de nucleósidos-fosfato. De cuántas hebras se compone el ADN? Cómo están unidas estas dos hebras? En la segunda figura, observar los apareamientos de bases. Qué uniones se establecen? En qué se diferencian los dos apareamientos?

Ahora, vamos a intentar relacionar estos esquemas con una estructura real de un fragmento de ADN que visualizaremos en Avogadro.

Actividad 3: Abrir el archivo descargado (1BNA) en el programa Avogadro

Para abrir el archivo descargado en el programa Avogadro, ir a la pestaña Archivo/Abrir, y allí buscar el archivo 1BNA en su computadora. Al seleccionarlo, se abrirá la estructura en la ventana del programa.

También podemos obtener el archivo de ADN 1BNA directamente desde el programa Avogadro.

Una alternativa para importar un archivo del PDB de código conocido es hacerlo directamente desde la ventana del programa Avogadro. Para ello, abra el programa y siga los comandos Archivo/Importar/Buscar del PDB... Allí se abre una ventana para ingresar el código PDB:

Ingresando el código deseado, y tipeando "OK", Avogadro importa directamente el archivo a la ventana del programa.

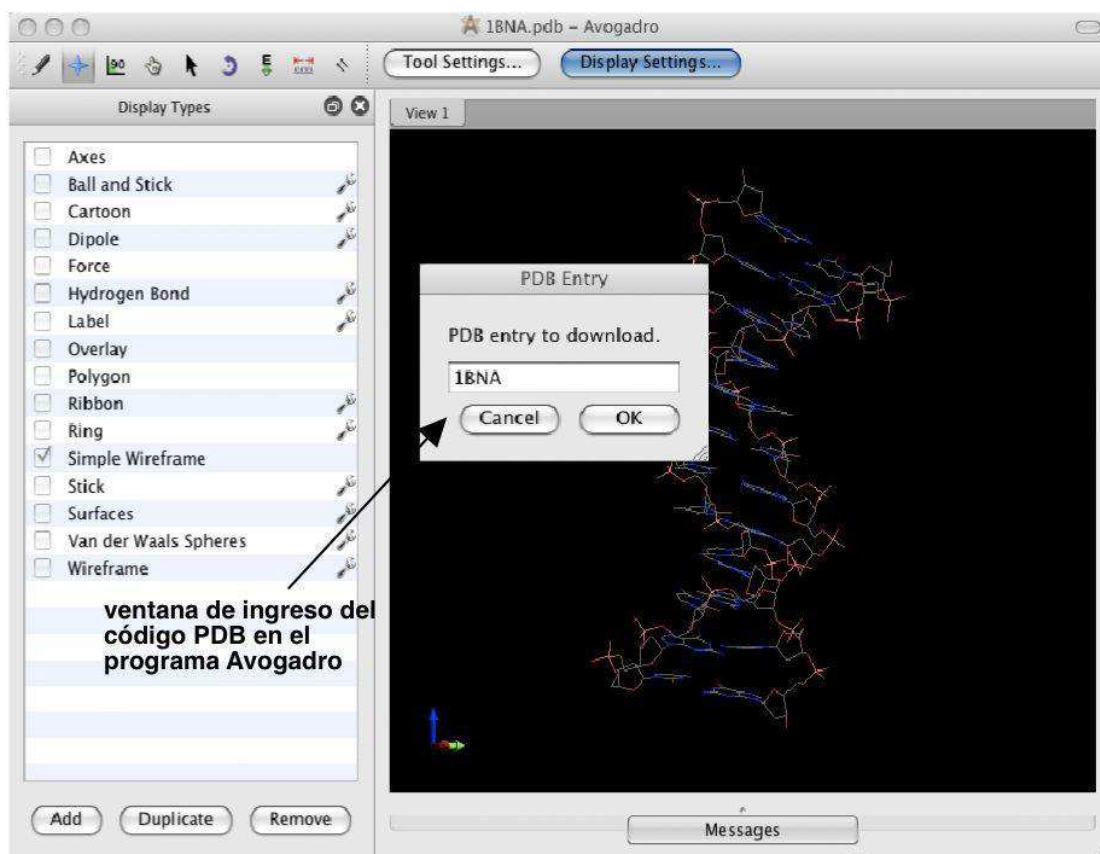
ADN código 1BNA: Información de la molécula en la página de RCSB PDB

Secuencia: 5'-CGCGAATTCGCG-3'

Cadenas: A,B

Preguntas: cuántas unidades componen cada hebra o molécula de ADN?

Cuántas hebras de ADN componen la estructura que vemos?



Actividad 4: Visualizar la molécula de ADN.

Exploren los diferentes modos de representación de la ventana de la izquierda. Cuales funcionan? Alguno de estos modos nos permiten visualizar las bases nitrogenadas? Prueben de usar el "mouse" para rotar y trasladar la molécula. Mirando la estructura, respondan las siguientes preguntas:

1- Elija el modo de representación "Ball and Stick". Qué son las moléculas rojas que "flotan" en torno a la estructura del ADN? Para removerlas, vaya al menú "Seleccionar", elija "seleccionar solvente" y luego en "Edición" seleccione "cortar". Así se remueven las moléculas de agua asociadas a la estructura.

2- Es "plana" la molécula de ADN, como se esquematiza en la figura? Qué forma tiene? En qué región se ubican las bases nitrogenadas?

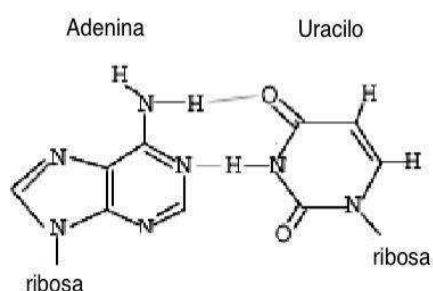
3- Mirando la figura de la introducción, identificar un nucleósido-fosfato en esta estructura, y un apareamiento de bases. Qué parte cambia de una unidad a otra? Puede distinguir los cuatro elementos del código ACTG?

4- Entonces, En qué parte de la molécula de ADN está contenido el "código genético"? Pueden identificar el extremo 5' de cada hebra de ADN y leer su secuencia en la estructura?

B- MIRANDO ARN

Repetimos el mismo procedimiento anterior pero ahora con una molécula de ARN. Usamos el código PDB 4TNA, que corresponde a un ARN de transferencia que carga el aminoácido fenilalanina de la levadura *S. cerevisiae* (levadura de la cerveza). Este ARN tiene 76 nucleótidos de longitud.

En ARN tiene, en vez de Timidina, otra base llamada Uracilo, que se aparea con la Adenina. Su estructura y el apareamiento con adenina se muestran a continuación:



Actividad 5: Responda las siguientes preguntas:

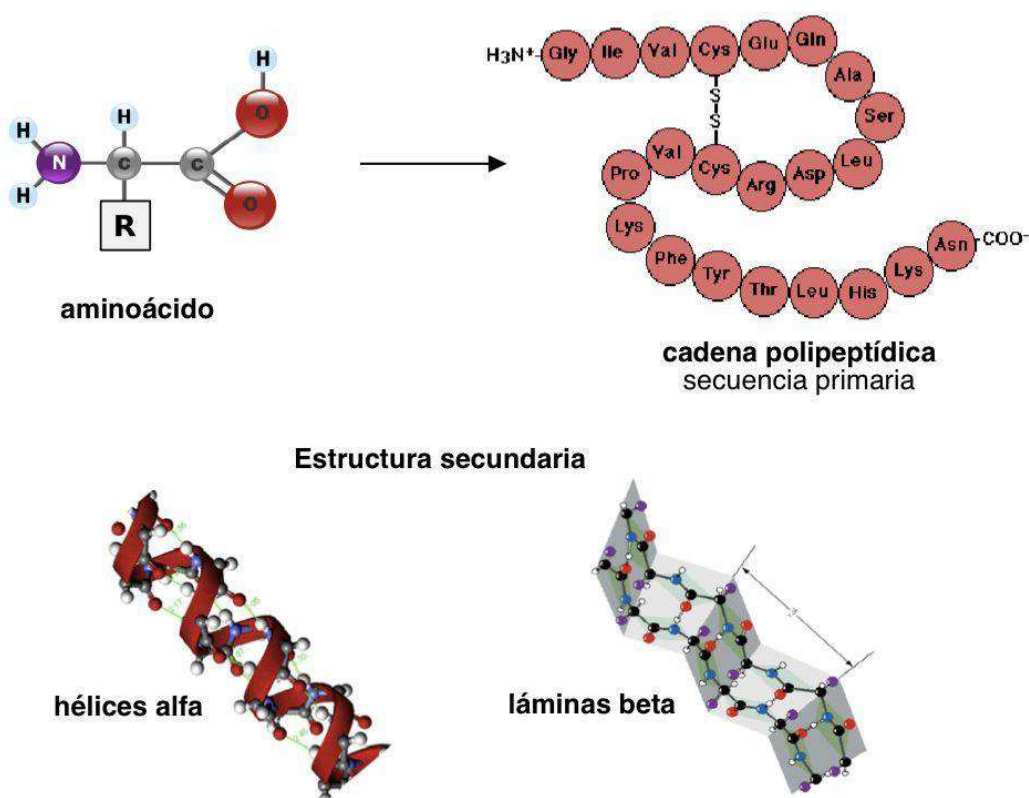
1-Pueden reconocer una base Uracilo dentro de este ARN?

2- Como es la estructura en comparación con la del ADN? Qué función cumplen los ARN de transferencia?

MIRANDO PROTEINAS

Vamos a hacer una primera mirada a las proteínas usando el programa Avogadro. Otros programas gratuitos como el Swiss PDB Viewer también pueden usarse para mirar estructuras de proteínas en una etapa mas avanzada, pero nos introducimos a ellas usando el programa Avogadro.

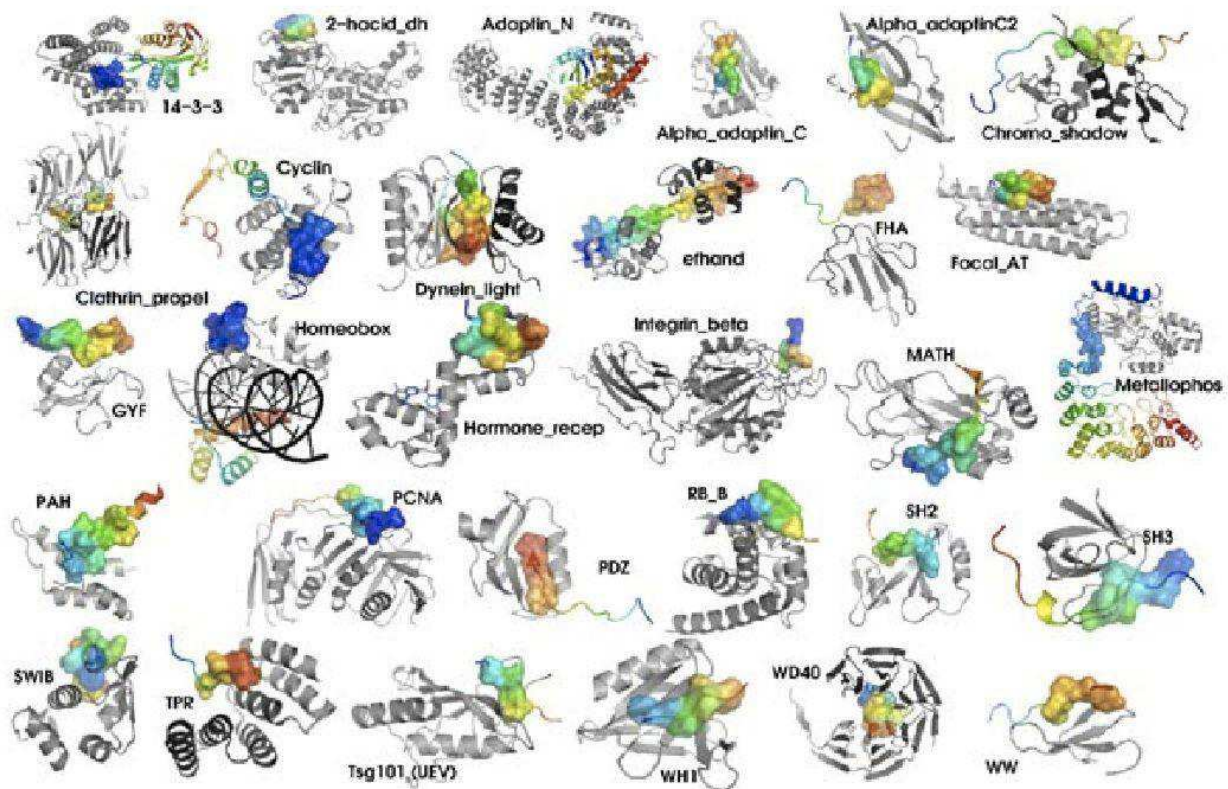
Las proteínas están formadas por aminoácidos, de los cuales existen 20 clases. Los aminoácidos se unen mediante el enlace peptídico en forma de cadenas polipeptídicas, y estas cadenas adquieren a su vez una estructura regular de dos tipos, llamadas hélices alfa y láminas beta.



Actividad 6: Estructura secundaria. Abran las estructuras 3S8W y 1E84. En qué se diferencian estas dos estructuras? Identifique los elementos de estructura secundaria en estas proteínas. Prueben diferentes modos de representación.

Estos elementos de estructura secundaria, a su vez se organizan en el espacio para formar la estructura terciaria "plegada" de la proteína. Finalmente, varias cadenas plegadas pueden asociarse formando estructuras mayores.

Estructura terciaria y cuaternaria: diversidad de estructuras de proteínas



Actividad 7:

Estructura cuaternaria. Abran la estructura de la insulina humana, buscada en la actividad 2 (Archivo 4EY1). Elija la representación "Ribbon" o "Cintas". Cuántas cadenas polipeptídicas componen la insulina humana? Qué elementos de estructura secundaria reconocen en ella?

Puentes de hidrógeno. Utilizando el modo "Wireframe", seleccione mostrar los puentes de hidrógeno en el menú de tipos de representación (menú izquierdo). Con el ratón o "mouse" amplíe la estructura hasta poder visualizar los puentes de hidrógeno (líneas punteadas en amarillo). Ahora, seleccione también la representación en "cintas". Esto marca el esqueleto carbonado o cadena principal. Puede encontrar puentes de hidrógeno entre átomos de la cadena principal que ayudan a mantener la estructura terciaria? Puede encontrar puentes de hidrógeno que involucren a cadenas laterales?

CODIGOS PDB QUE USAMOS EN ESTE TUTORIAL

1BNA: Dodecámero de ADN

4TNA: Molécula de ARN de transferencia

3S8W: Toxina de la bacteria streptomyces, clase todo-beta

1E84: Citocromo C, clase todo-alfa

4EY1: Insulina humana

RECURSOS EDUCATIVOS RECOMENDADOS

Les pasamos la información de algunos recursos de libros educativos, que se encuentran en todas las bibliotecas de colegios secundarios, publicados por el Instituto Nacional de educación técnica (INET) dependiente del Ministerio de Educación de la Nación:

Libro sobre estructura de proteínas:

<http://www.inet.edu.ar/programas/capacitacion/materiales/nuevos/proteinas.html>

Aquí, pueden encontrar material interesante que puede combinarse para enseñar a visualizar y entender diferentes propiedades de proteínas, con muchos ejemplos e ideas de ejercicios que se pueden hacer en clase.

Libro sobre biología molecular:

http://www.inet.edu.ar/programas/capacitacion/materiales/nuevos/bio_logia_molecular.html

Recursos del portal Educ.ar

Videos:

ADN. Capítulo 7 de la serie Explora.Ciencias.

http://www.educ.ar/recursos/ver?rec_id=106154

Infografías

Estructura del ADN

http://aportes.educ.ar/biologia/nucleo-teorico/estado-del-arte/el-libro-de-la-vida-el-adn/estructura_del_adn.php

El dogma central de la biología

http://aportes.educ.ar/biologia/nucleo-teorico/estado-del-arte/como-se-lee-el-libro-de-la-vida/el_dogma_central_de_la_biologi.php

Cuando el dogma no se cumple: el caso de virus y priones

http://aportes.educ.ar/biologia/nucleo-teorico/estado-del-arte/como-se-lee-el-libro-de-la-vida/cuando_el_dogma_no_se_cumple_e.php

El código genético

http://aportes.educ.ar/biologia/nucleo-teorico/estado-del-arte/el-libro-de-la-vida-el-adn/el_codigo_genetico.php

La teoría cinética de los gases:

<http://www.educ.ar/sitios/educar/recursos/embebido?id=20104>

Gases ideales I:

<http://www.educ.ar/sitios/educar/recursos/embebido?id=20077>

Gases ideales II:

<http://www.educ.ar/sitios/educar/recursos/embebido?id=20114>

Portal de Química en aportes Educ.ar

<http://aportes.educ.ar/quimica/>