

**PONTIFICIA UNIVERSIDAD JAVERIANA  
FACULTAD DE CIENCIAS  
PROGRAMA DE MAESTRÍA EN CIENCIAS BIOLÓGICAS**



**ESTUDIO DE LAS PROPENSIONES CONFORMACIONALES DEL AMINO ACIDO  
L-GLUTAMATO EN REGIONES ALTAMENTE ESTRUCTURADAS DE PROTEÍNAS  
RESUELTAS POR CRISTALOGRAFÍA DE RAYOS X.**

**DALIA ANDREA HINCAPIE PARRA**

**TESIS**

**Presentada como requisito parcial para optar al título de MAGISTER EN CIENCIAS  
BIOLÓGICAS**

**JANNETH GONZALEZ SANTOS .PhD.**

**BOGOTA**

**COLOMBIA**

**JUNIO DE 2011**

"La universidad no se hace responsable por los conceptos emitidos por sus alumnos en sus trabajos de tesis. Solo velará por qué no se publique nada contrario al dogma y a la moral católica y por qué las tesis no contengan ataques personales contra persona alguna, antes bien se vea en ellas el anhelo de buscar la verdad y la justicia"

Artículo 23 de la Resolución N° 13 de julio de 1946

**ESTUDIO DE LAS PROPENSIONES CONFORMACIONALES DEL AMINO ACIDO L-  
GLUTAMATO EN REGIONES ALTAMENTE ESTRUCTURADAS DE PROTEINAS  
RESUELTAS POR CRISTALOGRAFIA DE RAYOS X.**

**DALIA ANDREA HINCAPIE PARRA**

**APROBADO**

---

**Janeth González Santos. PhD.**

**Directora de Tesis**

---

**Orlando Emilio Sarmiento. PhD.**

**Codirector de Tesis**

---

**Milton Rueda. PhD.**

**Jurado 1**

---

**Ing. Fabio Avellaneda . MSc.**

**Jurado 2**

---

**Luz Mary Salazar. PhD.**

**Jurado 3**

**ESTUDIO DE LAS PROPENSIONES CONFORMACIONALES DEL AMINO ACIDO L-  
GLUTAMATO EN REGIONES ALTAMENTE ESTRUCTURADAS DE PROTEINAS  
RESUELTAS POR CRISTALOGRAFIA DE RAYOS X.**

**DALIA ANDREA HINCAPIE PARRA**

**APROBADO**

---

**Dra. Ingrid Schuler  
Decana Académica  
Facultad de Ciencias  
Pontificia Universidad Javeriana**

---

**Dr. Manuel Franco  
Director Programa de Posgrado  
Facultad de Ciencias  
Pontificia Universidad Javeriana**

**Bogotá, D.C 20 de Junio de 2011**

Este trabajo lo dedico al Dr. Leonardo Lareo quien soñó con el producto de este trabajo y a todas las personas que aman la investigación y la generación de nuevo conocimiento.

## **AGRADECIMIENTOS**

Quiero expresar mis más sinceros agradecimientos primero que todo a Dios por permitirme tener la oportunidad de hacer parte de este hermoso proyecto, a el Dr. Leonardo Lareo q.e.p.d. quien con su amor y dedicación me abrió las puertas de su mundo de conocimiento y me invito a hacer parte de él, a la Dr. Janeth González que con su dedicación y esfuerzo logró orientar y finalizar el trabajo que me permitirá incurrir en un camino nuevo y quien día a día me brindo frases de apoyo, al Dr. Orlando Acevedo quien con su paciencia y dedicación orientó mis procesos de trabajo matemático y de extracción de datos, al ingeniero Fabio Avellaneda quien me brindo su oportuna asesoría en el trabajo de procesamiento de datos, a mis padres por permitirme estar en este mundo y comprender la maravillas de la investigación también por su comprensión y apoyo, a mi futuro esposo por su amor, compañía y apoyo durante este proceso.

## TABLA DE CONTENIDO

1. PLANTEAMIENTO DEL PROBLEMA .....	14
2. MARCO TEÓRICO.....	16
2.1. Aminoácidos .....	16
2.2. Glutamato.....	18
2.2.1. Abundancia Relativa del Glutamato .....	19
2.3. Técnicas para determinar estructuras atómicas de biomoléculas. ....	20
2.3.1. Cristalografía de Rayos X. ....	21
2.3.2. Resonancia Magnética Nuclear .....	22
2.4. Geometría Molecular.....	23
2.5. Estructura Protéica .....	24
2.5.1. Estructura primaria. ....	25
2.5.2. Estructura secundaria: .....	25
2.5.4. Estructura cuaternaria.....	30
3. OBJETIVOS E HIPOTESIS .....	32
3.1. Hipótesis de la investigación.....	32
3.2. Objetivos de la investigación .....	32
3.2.1. Objetivo General.....	32
3.2.2. Objetivos Específicos.....	32
4. METODOLOGIA.....	34
4.1. Obtención de las estructuras protéicas de humano resueltas por cristalografía de rayos X disponibles en PDB .....	34
4.2. Abstracción de las coordenadas x,y,z de los glutamatos que hacen parte de proteínas de humano resueltas por cristalografía de rayos X .....	35

4.3. Filtrado de datos para abstraer los glutamatos pertenecientes a proteínas de humano resueltas por cristalografía de rayos X, que hacen parte de las regiones altamente estructuradas como hélices alfa y hojas beta.....	35
4.4. Cálculo de las variables internas (longitud de enlace, ángulo de enlace y ángulo dihedro) a partir de las coordenadas de los glutamatos. ....	36
4.5. Clustering y análisis de los datos obtenidos.....	36
4.6. Generación de un análisis estadístico.....	37
5. RESULTADOS Y DISCUSIÓN .....	39
5. CONCLUSIONES.....	54
6. PERSPECTIVAS .....	57
7. BIBLIOGRAFIA .....	58



## ÍNDICE DE FIGURAS

Figura 1. Estructura canónica de un aminoácido.....	16
Figura 2. Diagrama de Venn mostrando la relación de los 20 aminoácidos naturales de acuerdo con una selección de las propiedades físico- químicas.....	17
Figura 3.Estructura del glutamato.....	18
Figura 4.Estructura Secundaria de una proteína Alfa – hélice.....	26
Figura 5.Estructura Secundaria de una proteína: Hoja Plegada en Beta. ....	27
Figura 6.Representación de los ángulos phi y psi descritos por Ramachandran. ....	30
Figura 7. Diagrama abreviado de la metodología .....	38
Figura 8.Grafica comparativa que muestra los valores reportados de cada aminoácido en vertebrados y los valores obtenidos en las estructuras protéicas de humano. ....	40
Figura 9.Diagrama que muestra los vectores de las posiciones atómicas. ....	42
Figura 10.Molécula de Glutamato: nomenclatura PDB. ....	41
Figura 11.Valores de dispersión tanto para alfas como para betas.....	53

## ÍNDICE DE TABLAS

Tabla 1. Abundancia relativa del los aminoácidos en las proteínas. ....	20
Tabla 2. Vectores de Posición atómica para estructuras protéicas resueltas por cristalografía de rayos x. ....	42
Tabla 3. Inferencia estadística de distancia de enlace para alfa y beta.....	45
Tabla 4. Inferencia estadística de ángulos de enlace para alfa y beta.....	47
Tabla 5. Inferencia estadística de ángulos diedros para alfa y beta .....	48
Tabla 6. Comparación de los ángulos $\phi^*$ (N-CA-C-N) y $\psi^*$ (C-N-CA-C) propuesto por Ramachandran y propuesto por comparación de hipótesis de medias.....	53

## ÍNDICE DE ANEXOS

Anexo A. ALGORITMO AGM (Aminoacid Geometry Made Easy) .....	67
Anexo B. DATOS BÁSICOS DE LAS PROTEÍNAS DE HUMANO RESUELTAS POR CRISTALOGRAFÍA DE RAYOS X. ....	68
Anexo C. COORDENADAS x,y,z. ....	69
Anexo D. SCRIPTS EMPLEADOS PARA LA MINERIA DE DATOS.....	70
Anexo E. TABLA DE DISTANCIAS, ANGULOS Y DIEDROS DE LOS MONOMEROS ENCONTRADOS.....	71
Anexo F. SELECCIÓN DE HOJAS ALFA Y BETA.....	72
Anexo G. COMPOSICION DEL GRUPO DE DATOS .....	73
Anexo H. GRÁFICOS DE MEDIAS DE DISTANCIAS .....	79
Anexo I. GRÁFICOS DE MEDIAS DE ÁNGULOS.....	81
Anexo J. GRÁFICOS DE MEDIAS DE DIEDROS .....	83
Anexo K. ESTADÍSTICA POR INDICADOR .....	85

## INTRODUCCIÓN

Aunque el estudio de la conformación en biomoléculas es fundamental para el entendimiento de sus funciones biológicas, la información concerniente a estudios estructurales y conformacionales de las estructuras de macromoléculas a partir de los datos reportados experimentalmente resonancia magnética nuclear (RMN) y cristalografía de rayos X (DRX) y de los estados estructurales preferenciales particularmente para proteínas y los residuos de aminoácidos que las constituyen son limitados y en muchos casos inexistentes.

Desde este punto de vista, el conocimiento estructural y de parámetros básicos de las biomoléculas, nos permite acercarnos al entendimiento de diferentes procesos biológicos que implican la interacción entre las mismas. Tal es el caso de las proteínas las cuales al igual que otras macromoléculas poseen distribuciones espaciales de carga, en donde las fuerzas que mantienen su estructura, estabilidad y función biológica dependen de una compleja interrelación entre los potenciales que enlazan y pliegan su configuración espacial [78]. Considerando los aminoácidos que constituyen a las proteínas, tiene especial interés, el L-glutamato; debido a que además de ser considerado como el principal neurotransmisor excitador del Sistema Nervioso Central (SNC); desempeña un papel fundamental en los procesos que involucran la diferenciación neuronal y el desarrollo del SNC, tiene un papel único en la incorporación del nitrógeno en los esqueletos carbonados de los aminoácidos no esenciales y es fundamental en los procesos de desaminación y de detoxificación por remoción de los iones amonio ( $\text{NH}_4^+$ ), por ser uno de los aminoácidos más abundantes presentes en las proteínas [52].

El estudio de aminoácidos como el L-glutamato o como el de otros compuestos requiere del conocimiento de parámetros básicos como las propiedades específicas de la molécula, las cuales constituyen la fuente de información primaria. Sin embargo, la información reportada sobre los parámetros moleculares y estructurales básicos que permitan derivar información biológica para proponer modelos de correlación estadística es escasa o en muchos casos inexistente; por lo cual se hace necesario desarrollar estudios que permitan establecer relaciones entre las variaciones de las variables internas (distancias de enlace, ángulos de enlace y ángulos diedros) de las estructuras macromoleculares y las regiones altamente estructuradas que juegan un papel importante en la función de esta. De acuerdo con lo anterior el presente trabajo consistió en el análisis estructural de los glutamatos pertenecientes a las proteínas resueltas por cristalografía de rayos X y reportadas en el banco de datos de proteínas PDB hasta Septiembre de 2009 [80] (PDB). Este estudio aplicó algoritmos matemáticos para inferir estadísticamente la relación entre las variables internas derivadas de las coordenadas atómicas reportadas experimentalmente para los glutamatos y su respectiva posición en regiones altamente estructuradas de las proteínas.

## 1. PLANTEAMIENTO DEL PROBLEMA

En la naturaleza se han identificado más de 300 aminoácidos, no todos ellos codificados genéticamente, ni constituyentes de secuencias protéicas. En la mayoría de los organismos se reconocen 20 aminoácidos con codificación genética, aunque en microorganismos se han identificado hasta 22 aminoácidos genéticamente codificados.

De todo este grupo de moléculas se puede identificar un aminoácido fundamental para los procesos vitales, tiene un papel único en la incorporación del nitrógeno en los esqueletos carbonados de los aminoácidos no esenciales, y es fundamental en los procesos de deaminación y de detoxificación por remoción de los iones amonio ( $\text{NH}_4$ ). Este aminoácido particular, además una de las más abundantes moléculas de la naturaleza, es el glutamato (Glu, E) [14][59]

Dentro de las funciones metabólicas del L-glutamato (Glu) se ha encontrado que es el principal neurotransmisor excitador del Sistema Nervioso Central (SNC) y ejerce su función por medio de receptores (GluRs), donde desempeña un papel fundamental en los procesos que involucran la diferenciación neuronal y el desarrollo del SNC[57][58].

El estudio de ésta como de otras moléculas biológicas en cuanto a su estructura requiere del conocimiento de parámetros básicos como las propiedades específicas como las fuerzas de Van der Waals y los puentes de hidrógeno de la molécula, las cuales, constituyen la fuente de información primaria. Sin embargo, la información existente sobre los parámetros moleculares básicos de las conformaciones moleculares de las estructuras

del aminoácido L-glutamato presentes en los datos reportados experimentalmente (DRX) y de los estados estructurales preferenciales cuando el glutamato hace parte la cadena amino acídica de una estructura protéica son limitados y en muchos casos inexistentes.

Desde este punto de vista, el conocimiento estructural y de parámetros básicos de las biomoléculas, además de brindar información hasta ahora escasa en la literatura sobre las variaciones de las variables internas de aminoácidos es de particular interés cuando se encuentran haciendo parte de una cadena amino acídica en una región determinada como por ejemplo un dominio funcional de interés particular, también nos permite acercarnos al entendimiento de diferentes procesos biológicos que implican la interacción entre biomoléculas como por ejemplo el caso de la acción de ciertas drogas que actúan en moléculas endógenas y alimentar con esta información a las bases de datos mundiales de estructuras de proteínas más consultadas por la comunidad científica en general como lo es la base de datos PDB (Protein Data Bank) [80]. De acuerdo con lo anterior, el presente trabajo se centró en el estudio de las propensiones conformacionales del L-glutamato en regiones altamente estructuradas de proteínas resueltas por cristalografía de rayos X.

## 2. MARCO TEÓRICO

### 2.1. Aminoácidos

Las proteínas se encuentran constituidas por residuos de aminoácidos los cuales están unidos por un tipo específico de enlace peptídico, el cual fúe descrito en la década de los treinta, Linus Pauling descubrieron en ensayos acerca de los cambios conformacionales de las proteínas que el enlace peptídico tiene unas características que permiten que el enlace peptídico sea polar, por la mayor densidad de carga electrónica alrededor del oxígeno, y además tiene características de doble enlace, con lo que es plano y puede presentar isomería cis-trans [27]. Todos los aminoácidos cuentan en su estructura con un grupo carboxilo y un grupo amino unido a un carbono alfa. Estos grupos funcionales se encuentran deprotonado y protonado respectivamente a pH fisiológico (Fig. 1). De acuerdo con su cadena lateral los aminoácidos varían en estructura, tamaño y carga eléctrica influyendo así en su comportamiento físico y químico [1,2,3,4].

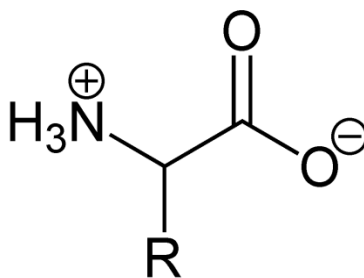


Figura 1. Estructura canónica de un aminoácido.



Los aminoácidos pueden clasificarse dentro de diferentes grupos de acuerdo a sus propiedades físicas y químicas [2], de tal manera que por sus propiedades pueden estar compartiendo características entre sí y también simultáneamente pueden hacer parte de varios grupos como se demuestra a continuación en la siguiente figura 2

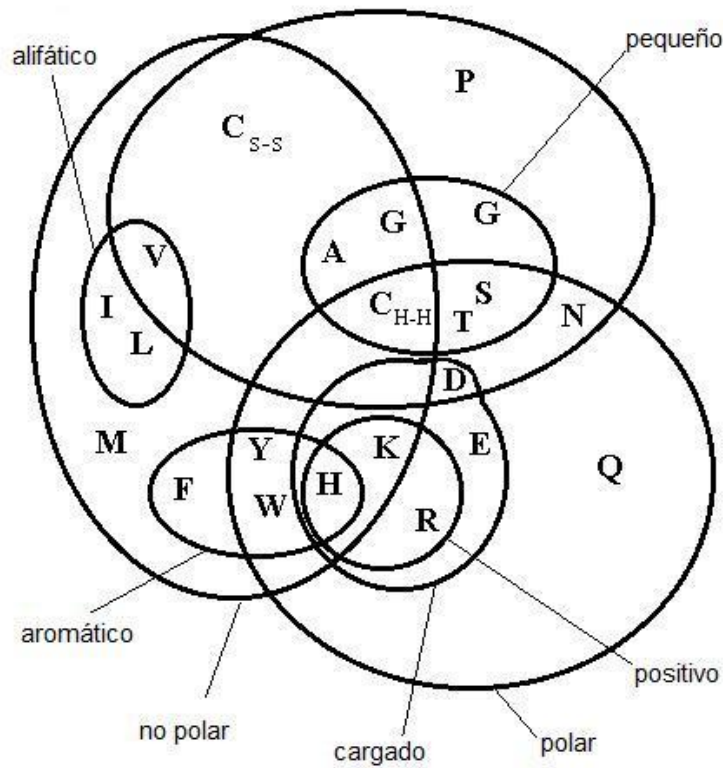


Figura 2. Diagrama de Venn mostrando la relación de los 20 aminoácidos naturales de acuerdo con una selección de las propiedades físico- químicas, las cuales se creen son importantes en la determinación de la estructura protéica. Tomada de <http://prowl.rockefeller.edu/aainfo/pchem.htm>

## 2.2. Glutamato

El glutamato es una macromolécula que participa en muchos de los procesos asociados al sistema nervioso central y periférico. [5,6,56]. Su estructura se explica en la figura 3.

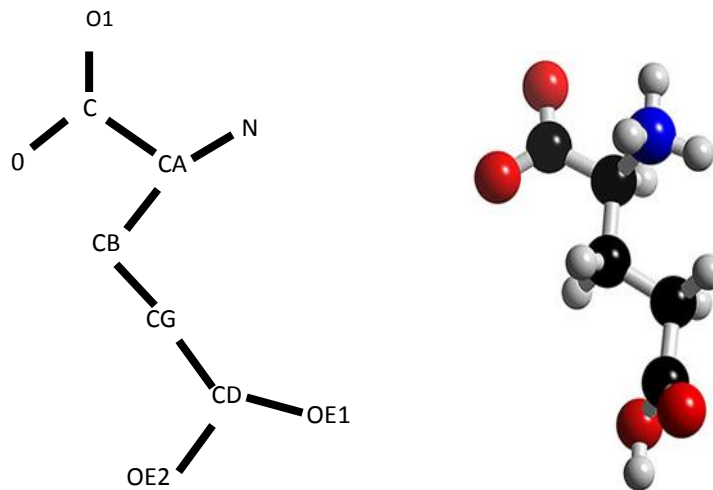


Figura 3. Estructura del glutamato. Presenta un carbono alfa (CA) que se une al grupo carboxilo y al grupo amino, un carbono beta (CB), un carbono gama (CG) y un carbono delta (CD), además presenta ángulos diedricos phi y psi Desarrollada con el software Avogadro.

El glutamato tiene una masa molecular relativa de 129.2 Da, un volumen molecular de 138.4 Å<sup>3</sup> y es uno de los aminoácidos con superficie accesible relativamente alta de 190.0 Å<sup>2</sup>. el grupo carboxilo de su cadena lateral se ioniza a un pH superior de 5.0 y es por ello un aminoácido extremadamente polar. su grupo carboxilo es muy reactivo, su valor energético de hidrofobicidad es alto ( 12.58 Kcal/mol), [7,50,51,57].

El glutamato está presente en la mayoría de los tejidos animales [9,10,11] pero su mayor concentración esta en el cerebro con niveles de hasta 10nM [8,12,13,52,55]., aunque en

una pequeña cantidad también se encuentra en el líquido extracelular [15]. Es quien actúa en todo el sistema nervioso central permitiendo el 60% de sinápsis las cuales todas son excitatorias heterogéneamente dando lugar a lo que se le conoce como reservorios de glutamato [22], los cuales se dividen en reservorio glial y reservorio neural; estos, a su vez se pueden dividir en cuatro reservorios funcionales diferentes: 1) el reservorio de glutamato para neurotransmisión, 2) el reservorio del glutamato glial, 3) el reservorio precursor del GABA, 4) el reservorio metabólico [14,23]. Es importante anotar que estudios recientes han presentado al glutamato como una de las moléculas responsables de la evolución del cerebro [10,23,42].

El L-glutamato, por ser considerado un aminoácido dicarboxílico, cargado a pH fisiológico (7.4), puede establecer puentes de hidrógeno con otras biomoléculas (agonistas o antagonistas) de forma directa o indirecta por medio de una molécula de agua. De acuerdo con lo anterior, este aminoácido generalmente se encuentra presente en los sitios de interacción de proteínas como enzimas y receptores y su localización en la topología estructural de la proteína generalmente es a nivel de regiones expuestas al solvente. [24,25]

### **2.2.1. Abundancia Relativa del Glutamato**

Cada uno de los aminoácidos presenta valores que hacen referencia a su abundancia dentro de las proteínas de los vertebrados [16,17]. De acuerdo con estudios preliminares se han reportado algunos rangos significativos (Tabla 1) entre el uso de codones y la frecuencia de los aminoácidos para vertebrados (gráfica 2) [17,18].

<b>AMINO ACIDOS</b>	<b>CODONES</b>	<b>FRECUENCIA OBSERVADA EN VERTEBRADOS</b>
ALANINA	GCU, GCA, GCC, GCG	7,34%
ARGININA	CGU, CGA, CGC, CGG, AGA, AGG	5,20%
ASPARAGINA	AAU, AAC	5,12%
ACIDO ASPARTICO	GAU, GAC	5,9%
CISTEINA	UGU, UGC	1,76%
<b>ACIDO GLUTAMICO</b>	<b>GAA, GAG</b>	<b>6,22%</b>
GLUTAMINA	CAA, CAG	3,7%
GLICINA	GGU, GGA, GGC, GGG	3,96%
HISTIDINA	CAU, CAC	2,26%
ISOLEUCINA	AUU, AUA, AUC	5,76%
LEUCINA	CUU, CUA, CUC, CUG, UUA, UUG	9,36%
LISINA	AAA, AAG	5,81%
METIONINA	AUG	2,32%
FENILALANINA	UUU, UUC	4,12%
PROLINA	CCU, CCA, CCC, CCB	5,0%
SERINA	UCU, UCA, UCC, UCG, AGU, AGC	7,38%
TREONINA	ACU, ACA, ACC, ACG	6,2%
TRIPTOFANO	UGG	1,34%
TIROSINA	UAU, UAC	3,3%
VALINA	GUU, GUA, GUC, GUG	6,48%

Tabla 1. Abundancia relativa de los aminoácidos en las proteínas. En la primera columna se muestra el nombre aminoácido, en la segunda columna se muestra los codones que codifican para estos y en la tercera columna se muestra el valor en porcentaje de la abundancia de cada aminoácido dentro de una estructura protéica. Tomada de <http://www.russelllab.org/aas/>.

### 2.3. Técnicas para determinar estructuras atómicas de biomoléculas.

La cristalografía de rayos X y la Resonancia Magnética Nuclear en solución son las dos técnicas experimentales mayormente usadas para determinar las estructuras atómicas de biomoléculas.

### 2.3.1. Cristalografía de Rayos X.

La cristalografía de rayos X es una técnica que consiste en hacer pasar un haz de rayos X a través de un cristal en estudio (sólido limitado por una serie de caras planas las cuales, forman parejas paralelas, en lados opuestos del cristal) [19,79]. El haz se escinde en varias direcciones debido a la simetría de la agrupación de átomos y, por difracción, da lugar a un patrón de intensidades que puede interpretarse según la ubicación de los átomos en el cristal [57,58]. Cuando el haz de rayos X incide sobre un cristal, provoca que los átomos que conforman a este dispersen a la onda incidente tal que cada uno de ellos produce un fenómeno de interferencia que para determinadas direcciones de incidencia será destructivo y para otras constructivo surgiendo así el fenómeno de difracción [20,49]. La información que proporciona la difracción de rayos X, se puede referenciar de dos formas, por un lado, la geometría de las direcciones de difracción (condicionadas por el tamaño y forma de la celdilla elemental del cristal) nos ofrecen información sobre el sistema cristalino [21] y por otro lado, los resultados se relacionan con la naturaleza de los átomos y las posiciones que ocupan en la red, tal que su medida constituye la información tridimensional necesaria para conocer la estructura interna del cristal [21]. Dentro de las técnicas más empleadas se encuentra el método de Laue en el que se utiliza un Policromático de Rayos X que incide sobre un cristal fijo y perpendicularmente a este se sitúa una placa fotográfica plana encerrada en un sobre a prueba de luz [58]. En la actualidad, este método se utiliza para determinar la simetría: si un cristal se orienta de tal manera que el haz incidente sea paralelo a un elemento de simetría, la disposición de las manchas en la fotografía revela su simetría [20,21].

Otro método es el de rotación o cristal giratorio en donde el cristal se orienta de tal manera que puede hacerse girar según uno de los ejes cristalográficos principales [22]. Finalmente el método powder en el que la muestra se pulveriza tan finamente como sea posible y se asocia con un material amorfo, en forma de eje acicular de 0.2 a 0.3 mm de diámetro, la muestra generalmente se hace girar en el haz de rayos X durante la exposición [20,22].

### **2.3.2. Resonancia Magnética Nuclear**

La información acerca de la estructura y dinámica de moléculas en disolución por RMN se obtiene a partir de las señales de absorción que se detectan cuando núcleos atómicos que poseen número cuántico de espín diferente de cero y que se encuentran situados dentro de un campo magnético se irradian con una radiación electromagnética de frecuencia equivalente a la diferencia de energía entre sus niveles de espín nuclear [23,59]. Por otro lado, los espines nucleares pueden interactuar entre sí, bien a través de los electrones de los enlaces que los separan (acoplamiento escalar) o bien directamente a través del espacio (acoplamiento dipolar) provocando, en un caso, desdoblamientos de las señales de resonancia y, en el otro, contribuyendo al fenómeno de relajación [23,61]. El espectro de RMN de una molécula compleja, estará formado por un gran número de líneas de absorción con frecuencias, intensidades y en algunas ocasiones anchos distintos [24,62].

Dentro de esta técnica los desplazamientos químicos son relevantes en la descripción de molecular; por ello, es importante resaltar que dicho aspecto depende de su entorno químico y es por tanto, muy sensible a la conformación de la molécula. [26,27].

## 2.4. Geometría Molecular

La geometría molecular o estructura molecular se refiere a la disposición tri-dimensional de los átomos que constituyen una molécula, dicha geometría está dada por la coordenadas  $x,y,z$  de cada uno de los átomos y las conectividades existentes entre ellos (distancias, ángulos y diedros). Se ha demostrado que la geometría molecular determina diferentes propiedades de las moléculas, como son la reactividad, polaridad, actividad biológica, entre otras [78].

Uno de los aspectos de mayor importancia en el estudio de la geometría molecular corresponde a los valores de las variables internas de la molécula como son: a) la longitud de enlace, b) los ángulos de enlace y c) los ángulos diedros.

La longitud de enlace es la magnitud que mide la relación espacial existente entre dos átomos que se unen por medio de enlaces covalentes. Los ángulos de enlace se constituyen por la relación existente entre tres átomos y los ángulos diedricos se constituyen por la relación existente entre cuatro átomos [26,28].

En 1963 el bioquímico Indú G. N Ramachandran y colaboradores [35], usaron un modelo de esferas rígidas para explorar el espacio conformacional permitido para cada átomo en una estructura polipeptídica. A partir de este estudio fue posible visualizar dichas representaciones como mapas bidimensionales donde los dos ejes corresponden a los ángulos diedros  $\phi$  y  $\psi$  [35,36], los cuales se describen en la sección 2.1. El mapa muestra las regiones permitidas y no permitidas estéricamente para todos los aminoácidos, excepto para Glicina (único aminoácido asimétrico) que no cuenta con cadena lateral y por lo tanto puede adoptar ángulos  $\phi$  y  $\psi$  en los cuatro cuadrantes del gráfico de Ramachandran [89].

## 2.5. Estructura Protéica

Las proteínas son biomoléculas formadas por cadenas lineales de aminoácidos: estas desempeñan diversas funciones en el organismo: 1) estructural, 2) inmunológica, 3) enzimática, 4) contráctil, 5) homeostática, 6) protectora o defensiva y 7) transducción de señales [26] entre otras; la función de una proteína está directamente relacionada con la estructura la cual depende de: 1) el orden de los aminoácidos, los cuales pueden tener varias conformaciones gracias al movimiento de los electrones dentro de sus átomos; lo cual, puede traducirse en un número astronómico de posibilidades de estructuras para la misma proteína, 2) el correcto plegamiento de los diferentes niveles estructurales; ya que, la cadena se pliega hasta alcanzar la formación nativa, en la que la proteína es funcional, si ocurre un error la proteína puede ser inactivada, 3) condiciones fisiológicas como la temperatura, y el pH , 4) fuerzas no covalentes, que a su vez incluyen interacciones electrostáticas, fuerzas de van der Wals y puentes de hidrógeno [30].

La estructura de las proteínas puede jerarquizarse en una serie de niveles, interdependientes. Estos niveles corresponden a:

1. Estructura primaria
2. Estructura secundaria
3. Estructura terciaria
4. Estructura cuaternaria



### **2.5.1. Estructura primaria.**

Es la forma de organización más simple de las proteínas y corresponde a la secuencia lineal de aminoácidos; es decir, el número de aminoácidos presentes y el orden como se enlazan en dicha estructura.[28]. Los aminoácidos secuencialmente forman la proteína por medio de enlaces peptídicos [53,54].

Conocer la estructura primaria no solamente es importante para entender su función si no también en el estudio de enfermedades genéticas; ya que es posible que el origen de una enfermedad radique en una secuencia anormal [29].

### **2.5.2. Estructura secundaria:**

Esta definida como la conformación presente en regiones locales de la proteína, que está estabilizada a través de enlaces de hidrógeno entre los elementos del enlace peptídico. [29,47]. Hay varios elementos de estructura secundaria que adquieren una forma real de hélice, de los que el más abundante es la hélice alfa, hipotetizada por Pauling en 1951 y posteriormente demostrada su existencia mediante Cristalografía [39,40,47]. En ésta como en todas las hélices se establecen puentes de hidrógeno entre los átomos del esqueleto polipeptídico dentro de ella misma. En este caso se establecen entre el carbonilo del residuo  $i$  y el amino del residuo  $i + 4$ . La periodicidad comienza desde el carbonilo de uno de los aminoácidos y desde allí se mantiene cada 3,6 residuos por vuelta, con los ángulos alfa =  $-45^\circ$  y beta =  $-60^\circ$ . El paso es de 0,56 nm [54].



Figura 4. Estructura Secundaria de una proteína Alfa – hélice. Tomada de [http://www.esacademic.com/pictures/eswiki/71/GCN4\\_coiled\\_coil\\_dimer\\_1zik\\_rainbow.pn](http://www.esacademic.com/pictures/eswiki/71/GCN4_coiled_coil_dimer_1zik_rainbow.pn)

La mayoría de las hélices alfa tienen 7 u 11 residuos, que se corresponden con 2 ó 3 vueltas. Además, las hélices alfa tienen momento dipolar, ya que todos los puentes de hidrógeno están alineados hacia el extremo carboxilo de la hélice. Esta distribución de la densidad de carga a lo largo de la hélice permite la unión de ligandos aniónicos a los extremos N-terminales de una hélice alfa de su receptor correspondiente, porque allí los puentes de hidrógeno no tienen con qué enlazarse y están accesibles a una posible interacción. Esta capacidad de unir ligandos puede ser un mecanismo de regulación de la función protéica, y de hecho se observa en receptores de la superficie celular [41,60]

Las representaciones en espiral permiten saber las interacciones y posibles localizaciones de las hélices en la estructura protéica, así como saber hacia dónde se encaran, si hacia dentro o al agua, según sean hidrófobas, hidrófilas o anfipáticas [1][30]. La estructura se puede asimilar a una hélice muy estirada de sólo dos residuos por vuelta y de sólo 1 Å de radio. Los grupos lineales se tienen que asociar formando un ángulo de unos 25° para formar láminas, ya que si no ocurre, no se pueden establecer puentes de hidrógeno entre

ellos [53]. Esta torsión genera que el plano de la lámina alfa tenga concavidad y convexidad, y se pueda hablar de un arriba y un debajo del plano [43,31].

Otra estructura propuesta por Pauling y Corey se conoce como hoja beta; en esta estructura todos los residuos presente una rotación de  $180^\circ$  [47,32,33]. La hoja beta está estabilizada por puentes de hidrógeno entre el grupo amida y el grupo carboxilo de un carbono adyacente, en esta estructura los radicales se disponen alternativamente a uno y otro lado de cadena polipeptídica [1,61].



Figura 5. Estructura Secundaria de una proteína: Hoja Plegada en Beta. Tomada de: [http://www.esacademic.com/pictures/eswiki/49/1m8n\\_Choristoneura\\_fumiferana.png](http://www.esacademic.com/pictures/eswiki/49/1m8n_Choristoneura_fumiferana.png)

### 2.5.3. Estructura terciaria

Está representada por los plegamientos y enrollamientos de la estructura secundaria, constituyendo formas tridimensionales geométricas que se mantienen por enlaces fuertes (puentes disulfuro entre dos cisteínas) y otros débiles (puentes de hidrógeno; fuerzas de Van der Waals; interacciones iónicas e interacciones hidrofóbicas). Desde el punto de vista funcional, esta estructura es la más importante debido a que al alcanzarla es cuando la mayoría de las proteínas adquieren su actividad biológica o función. [48,33,46]

La estructura terciaria de una proteína es la responsable directa de sus propiedades biológicas, ya que la disposición espacial de los distintos grupos funcionales determina su interacción con los diversos ligandos. [65,76,77]. Para las proteínas que constan de una sola cadena polipeptídica (carecen de estructura cuaternaria), la estructura terciaria es la máxima información estructural que se puede obtener. [75] Dentro de esta estructura se distinguen dos grupos: 1) las de tipo fibroso en las que las hélices alfa y las hojas beta pueden mantener su ordenamiento sin recurrir a grandes modificaciones, únicamente introducen ligeras torsiones longitudinales, 2) las de tipo globular en las que se suceden regiones con estructuras alfa y beta. [35,38].

Las hélices alfa presentan tendencias en su conformación las cuales, consisten en la formación natural de acuerdo con residuos [66,74].

Las estructuras beta son los elementos estructurales (descritos por primera vez por Geddes, 1968, y definidos por Venkatachalam, 1968). [67,68]. El posible papel de estas estructuras se ha considerado dos posturas: una que son los elementos pasivos y que se forman como consecuencia de interacciones de largo alcance [70,71]. La otra es que tiene

un papel activo en el plegamiento de zonas distales de cadena y promoviendo las interacciones terciarias. [69,44,72].

Existen formas para representar las estructuras de las proteínas; una de ellas consiste en esquematizar el esqueleto polipeptídico como si fuera una cinta hecha a base de planos que contienen los enlaces peptídicos, con esta representación se pueden ver los elementos de estructura secundaria, la otra forma de representar y la que mejor describe la estructura tridimensional de una proteína es la que representa las hélices alfa por cilindros y los grupos lineales alfa por flechas planas [46,73]. Pero, aunque existen estas representaciones se hace necesario entonces visualizar detalladamente las coordenadas atómicas de cada uno de los átomos que integran la proteína, estas representaciones fueron desarrolladas en la década de los cuarenta por el bioquímico hindú Gopalasamudram Narayana Ramachandran junto con Viswanathan Sasisekharan, desarrollaron una técnica para visualizar todas las combinaciones posibles de los ángulos diedricos  $\Psi$  (psi) contra el  $\Phi$  (phi) en los aminoácidos de un polipéptido [74]. Este diagrama que lleva su apellido (Ramachandran) está dividido en cuadrantes: en el primer cuadrante se ubican las hélices alfa a la izquierda, en el segundo cuadrante se hallan las combinaciones de la hoja beta, en el tercer cuadrante las hélices alfa a la derecha y los giros [81,35,45]. Los valores de libertad permitidos para los giros, se presentan únicamente en el ángulo phi ( $\phi$ ), el cual debe estar entre el CA- N y el ángulo psi ( $\psi$ ), que debe estar entre C-CA, ya que los ángulos omega son rígidos; dichos valores se muestran en la figura 6 [83]

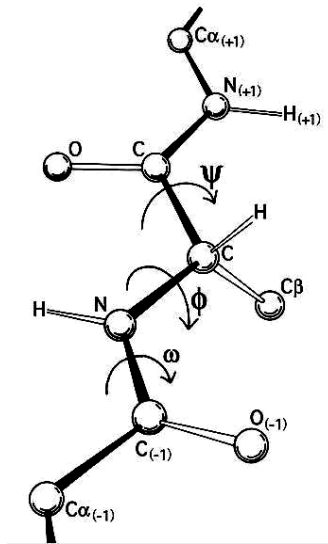


Figura 6. Representación de los ángulos phi y psi descritos por Ramachandran en 1963. Muestra la orientación y los átomos que deben incluirse dentro de cada ángulo Biochemistry. Segunda Edición. Garrett, R. H. Grisham, C. M. Saunders College Publishing, 1999.

#### 2.5.4. Estructura cuaternaria

Cuando una proteína está constituida por más de una cadena polipeptídica; la proteína adquiere estructura cuaternaria. La estructura cuaternaria debe considerar: el número y la naturaleza de las distintas subunidades o monómeros que integran el oligómero y la forma en que se asocian en el espacio para dar lugar al oligómero [33]. Las estructuras cuaternarias se unen por los mismos tipos de uniones químicas encontradas en la estructura terciaria, incluyendo uniones débiles de carácter electrostático y/o puentes de hidrógeno y covalentes como puentes de disulfuro [84, 34].

Cuando varias proteínas con estructura terciaria de tipo globular se asocian para formar una estructura de tipo cuaternario, los monómeros pueden ser: a) iguales, como en el

caso de la fosfoglucoisomerasa o de la hexoquinasa; b) altamente parecidos, como en el caso del lactato deshidrogenasa; c). estructural y funcionalmente distintos, como en el caso del Aspartato transcarbamilasa, un enzima alostérico con seis subunidades con actividad catalítica y seis con actividad reguladora [1,34].

### **3. OBJETIVOS E HIPOTESIS**

#### **3.1. Hipótesis de la investigación**

La información estructural derivada de las coordenadas cartesianas átomo a átomo de los residuos de aminoácidos constituyentes de las cadenas polipeptídicas permitirá establecer relaciones entre las variaciones de las variables internas (longitudes de enlace, ángulos de enlace y ángulos diedros) del aminoácido estudiado y las regiones estructuradas de las proteínas de las cuales hace parte.

#### **3.2. Objetivos de la investigación**

##### **3.2.1. Objetivo General**

Realizar un estudio estructural y conformacional de las estructuras del L-glutamato presentes en los datos reportados experimentalmente por cristalografía de Rayos X y de los estados estructurales preferenciales cuando el glutamato hace parte la cadena aminoacídica de de una región altamente estructurada de una proteína.

##### **3.2.2. Objetivos Específicos**

1. Desarrollar e implementar un algoritmo matemático que permita obtener los valores de las variables internas (distancias de enlace, ángulos de enlace, ángulos



diedros) del aminoácido L-glutamato a partir de sus coordenadas cristalográficas, para las proteínas depositadas en el banco de datos de proteínas PDB (septiembre 2009).

2. Analizar la dependencia conformacional de las interacciones atómicas específicas del aminoácido L-glutamato cuando hace parte de una cadena amino acídica.
3. Generar una análisis estadística entre los valores de las variables internas del amino acido L-glutamato y las regiones altamente estructuradas de las proteínas depositadas en PDB (septiembre 2009).

## 4. METODOLOGIA

Para desarrollar el estudio de las propensiones conformacionales del aminoácido L-glutamato en regiones altamente estructuradas de proteínas resueltas por cristalografía de rayos x se ejecuto una serie de pasos que se agruparon por secciones: 1) en esta se obtuvieron las coordenadas de estructuras resueltas por rayos X, 2) aquí se obtuvo coordenadas de los glutamatos que hacen parte de las regiones estructuradas como hélices alfa y hojas beta, 3) se obtuvo el set de datos completo haciendo uso de los criterios de inclusión, 4) se obtuvo los datos de las variables internas de los glutamatos y 5) se realizó la inferencia estadística.

Un diagrama abreviado de la metodología se presenta en la página 37 figura 7.

### 4.1. **Obtención de las estructuras protéicas de humano resueltas por cristalografía de rayos X disponibles en PDB**

A partir de la información depositada en PDB (Septiembre 2009) se obtuvieron 14371 estructuras protéicas de humanos resueltas por cristalografía de rayos X (Anexo D). Las proteínas fueron incluidas y categorizadas en una base de datos que incluye los identificadores básicos como: PDB ID, resolución cristalográfica, longitud de la secuencia, número de monómeros asociados por estructura obtenida y referencia bibliográfica (Anexo D).

#### **4.2. Abstracción de las coordenadas $x,y,z$ de los glutamatos que hacen parte de proteínas de humano resueltas por cristalografía de rayos X**

A partir de las estructuras obtenidas de las proteínas, se realizó la abstracción de las coordenadas átomo a átomo en los ejes  $x, y, z$  de los glutamatos (Anexo C) para cada proteína y se identificó la posición respectiva en la cadena amino acídica (Anexo D).

#### **4.3. Filtrado de datos para abstraer los glutamatos pertenecientes a proteínas de humano resueltas por cristalografía de rayos X, que hacen parte de las regiones altamente estructuradas como hélices alfa y hojas beta.**

Al tener el archivo de las coordenadas en los ejes  $x, y, z$  de cada uno de los glutamatos pertenecientes a proteínas de humano resueltas por cristalografía de rayos X se filtraron las coordenadas equivalentes a aquellos glutamatos pertenecientes a hélices alfas y betas (Ver anexo F). Para obtener este conjunto de datos se utilizó criterios de inclusión, donde se tomó la ubicación de cada glutamato con sus coordenadas y lo incluyó en el rango de la base de PDB que ubica las hélices alfa y las hojas beta. (Anexo G).

La base de datos filtrada cumplió con los siguientes parámetros de selección: a) Glutamatos de proteínas pertenecientes a humano; b) Glutamatos de proteínas resueltas por cristalografía de rayos X. c) Glutamatos pertenecientes a regiones protéicas altamente estructuradas (hélices alfa y hojas beta) y d) Cadena A cuando correspondía a estructuras multiméricas, e) Glutamatos pertenecientes a proteínas únicamente con los datos cristalográficos completos, f) Glutamatos que no pertenecieran a regiones terminales.

#### **4.4. Cálculo de las variables internas (longitud de enlace, ángulo de enlace y ángulo dihedral) a partir de las coordenadas de los glutamatos.**

Se elaboró y aplicó el algoritmo AMGE “ Aminoacid Geometry Made Easy” (Ver Anexo A), que permitió calcular las variables internas del glutamato (distancias de enlace, ángulos de enlace y ángulos diedricos) a partir de las coordenadas  $x$ ,  $y$ ,  $z$  para las estructuras estudiadas. Este algoritmo fue escrito e implementado con la colaboración del Departamento de Ingeniería de Sistemas de la Pontificia Universidad Javeriana. Para evaluar su eficiencia se validó por medio del software VMD haciendo uso de 100 datos como muestreo probabilístico los cuales corresponden al 0.7% del total de la población homogénea en estudio, obtenido por medio del factor de muestreo aleatorio simple dada la homogeneidad de la población de interés.

#### **4.5. Clustering y análisis de los datos obtenidos.**

Con base en los valores obtenidos para las variables internas a partir de las coordenadas atómicas aplico una prueba de comparación de medias para establecer que los valores de hélices alfa y hojas beta son diferentes, también se aplicó una prueba de intervalo de confianza al 95% para establecer los rangos para alfa y beta en las distancias de enlace, en los ángulos de enlace y en los ángulos diedricos (Anexo H).

#### **4.6. Generación de un análisis estadístico.**

Con los resultados obtenidos a partir de la prueba de comparación de medias se pudo determinar las diferencias en términos de propensiones internas que tiene el L-glutamato cuando hace parte de una estructura protéica en regiones altamente estructuradas como una hoja en alfa y una hoja en beta.

Con la obtención de intervalos de confianza se pudo realizar la inferencia estadística que permite determinar con un 95% de confianza cuando un glutamato pertenece a una estructura en alfa y cuando el L-glutamato pertenece a una hoja en beta.

Estos resultados se reflejaron en gráficos de dispersión que tanto se aleja de la media, para establecer los rangos permitidos en las alfas y betas.

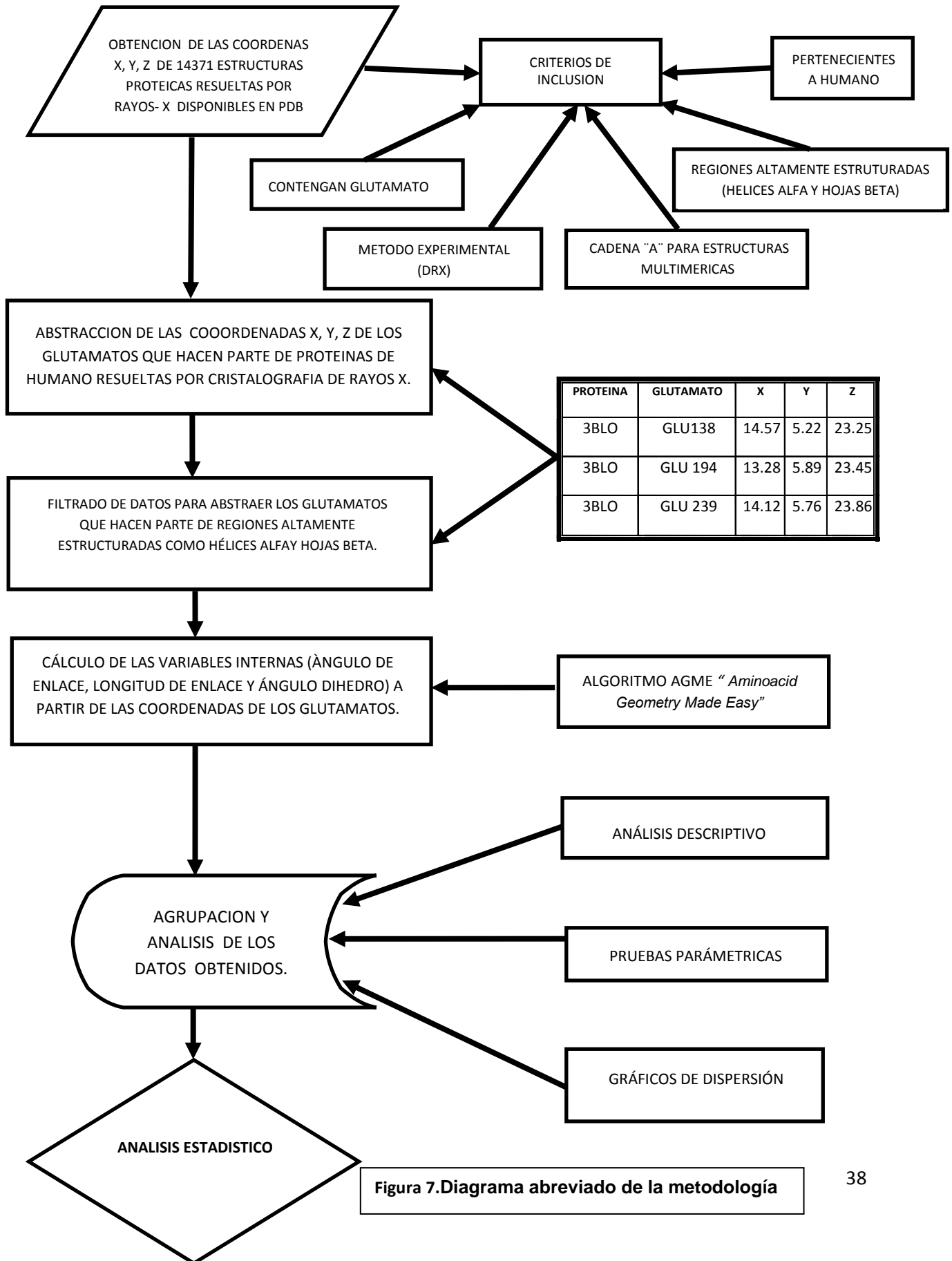


Figura 7. Diagrama abreviado de la metodología

## 5. RESULTADOS Y DISCUSIÓN

A partir del banco de datos de proteínas PDB (Septiembre 2009) y mediante el uso de diferentes algoritmos elaborados e implementados (Anexo C) se obtuvieron 14371 estructuras protéicas pertenecientes a humano y resueltas por cristalografía de rayos X. En el anexo B se presentan los identificadores básicos del conjunto de datos correspondientes a: a) identificador por proteína en la base de datos PDB (PDB ID,) b) resolución cristalográfica, c) longitud de la secuencias, d) número de monómeros asociados por estructura obtenida y f) referencia bibliográfica. La base de datos de dichas proteínas fue filtrada (Anexo C) para obtener exclusivamente las coordenadas x,y,z de los glutamatos de las estructuras protéicas además de su posición en la cadena aminoacídica y su correspondencia o no correspondencias con regiones estructuradas como hélices alfa y hojas beta. El conjunto de datos obtenido, corresponde a un 58% del total de proteínas depositadas en PDB hasta Septiembre de 2009. Es importante anotar que este porcentaje corresponde a la totalidad de las estructuras resueltas por cristalografía; el 42% restante equivale a proteínas u otro tipo de biomoléculas resueltas por otros métodos experimentales como RMN y microscopía electrónica.

El porcentaje de glutamato en el conjunto de estructuras estudiadas se observa un 6.64 % del total de aminoácidos presentes en estructura protéicas (figura 8). Estos valores concuerdan con resultados de estudios preliminares en cuanto a la abundancia relativa de los aminoácidos en vertebrados [18,37,42] (tabla 1, sección 2.3). De acuerdo con lo anterior y considerando el papel biológico del L-glutamato (sección 2.2), éste aminoácido fue escogido como modelo para el análisis estructural de este estudio.

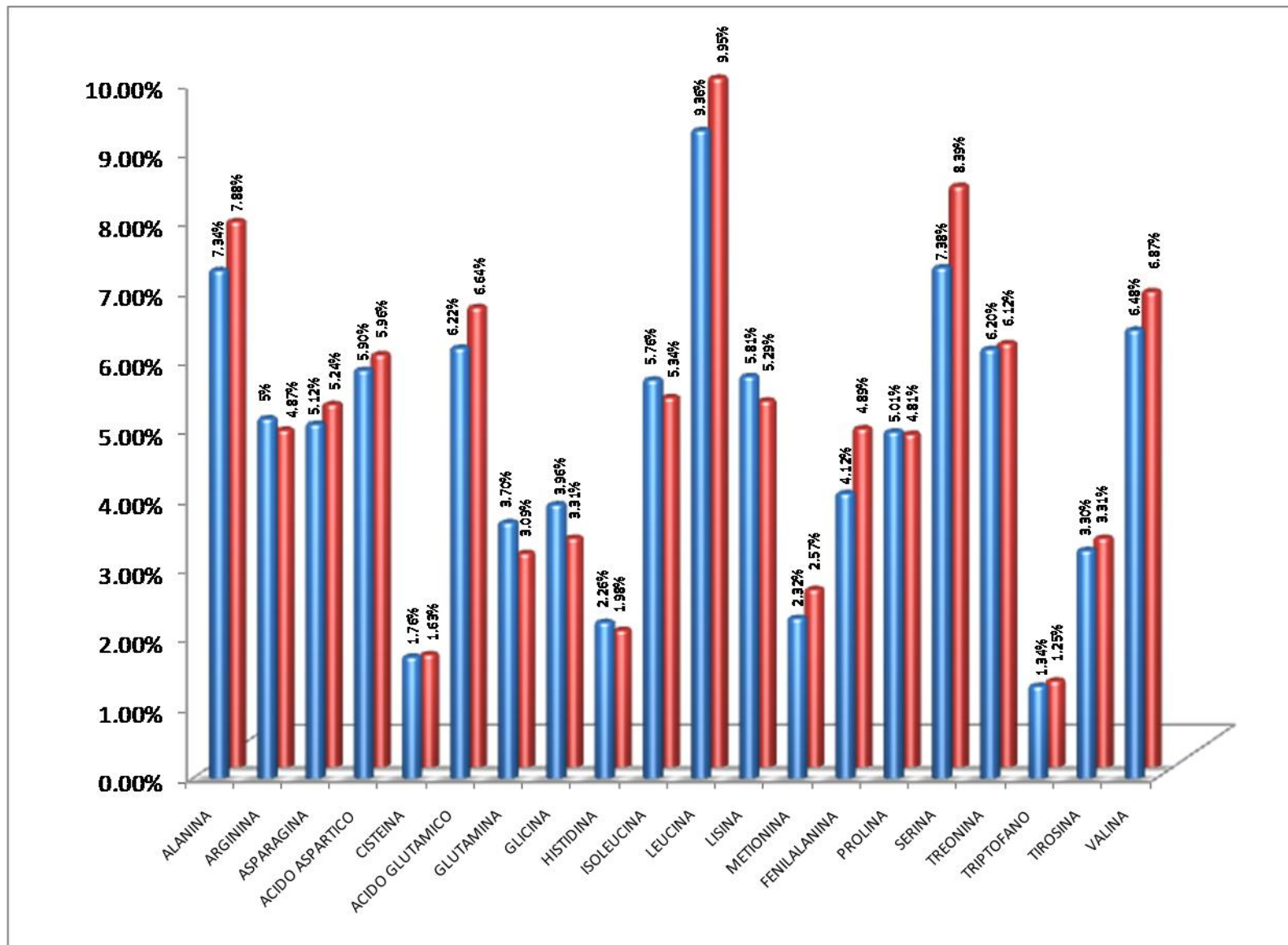


Figura 8. Grafica comparativa que muestra los valores reportados de cada aminoácido en vertebrados (barras azules) y los valores obtenidos en las estructuras proteicas de humano (barras rojas).



Para obtener el cálculo de las variables internas del L-glutamato se utilizó la nomenclatura atómica estandarizada por PDB [80][82]. Con base en esta nomenclatura se definieron las relaciones intermoleculares que se presentan en este estudio. (Tabla 2)

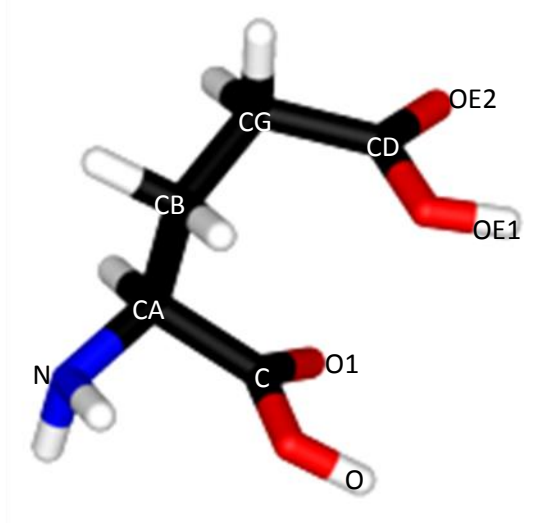


Figura 9. Molécula de Glutamato: nomenclatura PDB. CA: carbono alfa, CB: carbono beta, CG: Carbono Gamma, CD: Carbono delta, N: nitrógeno, O, O1, OE1, OE2: Oxígenos.

Con base en las coordenadas cristalográficas de los glutamatos fue posible obtener los valores de distancias de enlace, ángulos de enlace y ángulos diedros a partir de la generación los respectivos vectores para cada posición atómica, teniendo en cuenta las siguientes conectividades atómicas (Tabla 2):

1. **Dupletas:** longitudes de enlace comprendidas entre dos átomos consecutivos
2. **Tripletas:** ángulos de enlace comprendido entre tres átomos consecutivos.
3. **Cuadrupletas:** ángulos diedricos comprendidos entre cuatro átomos consecutivos

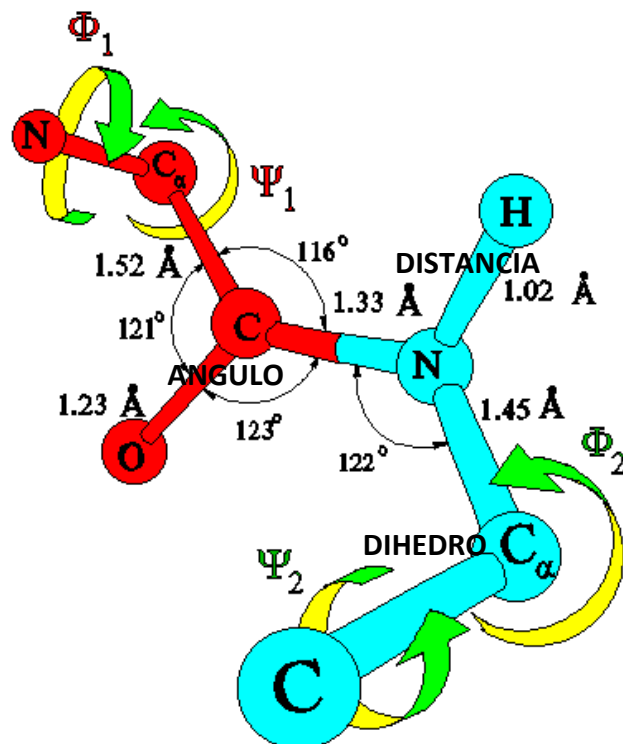


Figura 10. Diagrama que muestra las vectores de las posiciones atómicas.

DISTANCIAS (L)	ANGULOS (A)	DIEDROS (T)
O-C	O-C-O	O-C-CA-CB
C-O1	O1-C-CA	O1-C-CA-N
C-CA	O-C-CA	CA-CB-CG-CD
CA-N	C-CA-N	O1-C-CA-CB
CA-CB	C-CA-CB	CB-CG-CD-OE2
CB-CG	CB-CA-N	CB-CG-CD-OE1
CG-CD	CB-CG-CD	N-CA-CB-CG
CD-OE1	OE2-CD-OE1	C-CA-CB-CG
CD-OE2	CG-CD-OE1	N-CA-C-N
	CA-CB-CG	C-N-CA-C
	CG-CD-OE2	

Tabla 2. Vectores de Posición atómica para estructuras protéicas resueltas por cristalografía de rayos x. En la primera columna se muestran los vectores de las distancias de enlace, en la columna dos se muestran los vectores de los ángulos de enlace y en la tercera columna se muestran los vectores de los ángulos diedros.

Para la obtención de los vectores de cada posición atómica fue escrito e implementado el algoritmo matemático AGME “aminoacid geometry made easy” (Anexo A) <sup>1</sup>. En este algoritmo, se aplica el producto vectorial y el producto punto para determinar las distancias de enlace, y a partir del teorema del seno y el coseno se hallan los ángulos de enlace y ángulos diedros (Anexo A).

Debido al tamaño de la población (300000 aproximadamente), el algoritmo AGME fue validado con base en una muestra aleatoria de 100 datos como menciono anteriormente. La validación consistió en la comparación de la medición de las variables internas tanto con el algoritmo AMG como con VMD (Visual Molecular Dynamics) [83]. VMD es un software de análisis y de visualización molecular, diseñado para los sistemas biológicos como las proteínas, ácidos nucleicos, lípidos bicapa. Entre los programas de gráficos moleculares, VMD es una de los más empleados debido a su capacidad para operar de manera eficiente sobre las trayectorias de la dinámica molecular, haciendo uso de su interoperabilidad con otros paquetes de simulación de dinámica molecular y su integración de estructura y de información de la secuencia [83]. Los resultados obtenidos mediante los sistemas de medición fueron análogos (valores equivalentes), por lo cual se logro validar la eficiencia del algoritmo.

---

<sup>1</sup> AGME “aminoacid geometry made easy fue un esfuerzo colaborativo del entre profesores del grupo biq computacional particularmente Dr, leonardo Lareo q.e.p.d. del departamento del nutrición y biquímica y el ingeniero Fabio Avellaneda del Departamento de Ingeniería de sistemas de la PUJ quienes participaron activamente en el desarrollo e implementación del algoritmo para dar curso a el presente estudio.

Una vez validado el algoritmo, fue ejecutado sobre toda la población de estructuras de L-glutamato. Se hizo un clustering a la información correspondiente a los valores de cada vector de posición atómica, con base en las siguientes variables e indicadores (Anexo F):

1. Datos de distancias atómicas de estructuras asociadas con hélices alfa.
2. Datos de distancias atómicas de estructuras asociadas con hélices beta.
3. Datos de ángulos atómicos de estructuras asociadas con hélices alfa.
4. Datos de ángulos atómicos de estructuras asociadas con hélices beta.
5. Datos de ángulos diedricos de estructuras asociadas con hélices alfa.
6. Datos de ángulos diedricos de estructuras asociadas con hélices beta.

Estos datos fueron sometidos a un análisis estadístico, para garantizar la calidad de los datos se elimino 5 de arriba y 5 de abajo en el total de la población empleada; dicho análisis se desarrollo en dos sesiones: 1) Se aplico una prueba de hipótesis de comparación de medias que consistió en analizar si la diferencia entre las medias de hélices alfa y de hélices beta era igual a cero ( $H_0: \text{diff}=0$ ) v.s. ( $H_a: \text{diff}\neq 0$ ); de esta forma, se estableció que existe diferencia estadística entre las medias, es decir  $MA \neq MB$  (Anexo K). 2) Se establecieron límites de confianza al 95% tanto para hélices alfa y como para hojas beta; en cuanto a distancias de enlace, para ángulos de enlace y para ángulos diedricos. (Tabla 3,4,5). De acuerdo con lo anterior, los resultados indican que por lo menos para las estructuras estudiadas se presentan diferencias entre las distancias, ángulos y diedros cuando el L-glutamato pertenece a una hélice alfa y a una hoja beta.

De los 28057 glutamatos 24645 pertenecen a regiones alfa, para las distancias los promedios están entre 1.0 y 1.3; para los ángulos de enlace los promedios se encuentran entre 109.1° y 121.3° y para los ángulos diedricos los promedios se encuentran entre 116.7° y 139.4°. El restante (3412) corresponden a regiones beta, los promedios de las distancias atómicas en general están entre 1.1 y 1.3; los promedios de los ángulos de enlace se encuentran entre 112.6° y 122.0° y los ángulos diedricos se encuentran entre 111.3° y 177.8° (Tabla 3,4,5).

<b>O-C</b>	ALFA	1,4	1.3	1.3
	BETA	1,4	1.3	1.3
	GENERAL	1,4	1.3	1.3
<b>C-01</b>	ALFA	1,4	1.3	1.3
	BETA	1,4	1.3	1.3
	GENERAL	1,4	1.3	1.3
<b>C-CA</b>	ALFA	1,4	1.3	1.4
	BETA	1,3	1.3	1.3
	GENERAL	1,4	1.3	1.3
<b>CA-N</b>	ALFA	1,3	1.3	1.3
	BETA	1,4	1.3	1.3
	GENERAL	1,3	1.3	1.3
<b>CA-CB</b>	ALFA	1,4	1.3	1.3
	BETA	1,1	1.1	1.1
	GENERAL	1,3	1.3	1.3
<b>CB-CG</b>	ALFA	1,1	1.1	1.1
	BETA	1,1	1.1	1.1
	GENERAL	1,1	1.1	1.1
<b>CG-CD</b>	ALFA	1,2	1.1	1.1
	BETA	1,4	1.3	1.4
	GENERAL	1,2	1.1	1.1
<b>CD-OE1</b>	ALFA	1,2	1.2	1.2
	BETA	1,1	1.1	1.1
	GENERAL	1,2	1.1	1.1
<b>CD-OE2</b>	ALFA	1,1	1.0	1.1
	BETA	1,4	1.3	1.3
	GENERAL	1,1	1.1	1.1

Tabla 3. Intervalos de confianza por variable e indicador, en la primer columna se detalla la variable (distancia de enlace), en la segunda columna se ubica cada indicador (alfa y

beta), en la tercera columna se ubica el valor promedio, en la cuarta columna se ubica el intervalo de confianza el 95% .Léase los valores en Armstrong.

<b>O-C-O1</b>	ALFA	114,0	113.9	114.0
	BETA	109,1	108.9	109.2
	GENERAL	113,4	113.3	113.4
<b>O-C-CA</b>	ALFA	113,8	113.7	113.9
	BETA	108,1	107.8	108.4
	GENERAL	113,1	113.0	113.2
<b>O1-C-C-A</b>	ALFA	114,2	114.1	114.2
	BETA	114,7	114.4	114.8
	GENERAL	114,3	114.1	114.3
<b>C-CA-N</b>	ALFA	113,7	113.6	113.7
	BETA	110,5	110.4	110.5
	GENERAL	113,3	113.2	113.3
<b>CB-CA-N</b>	ALFA	111,7	111.6	111.8
	BETA	113,6	113.5	113.7
	GENERAL	112,0	111.9	112.0
<b>C-CA-CB</b>	ALFA	113,4	113.3	113.4
	BETA	120,4	120.1	120.6
	GENERAL	114,2	114.1	114.2
<b>CA-CB-CG</b>	ALFA	117,4	117.3	117.4
	BETA	120,4	120.0	120.6
	GENERAL	117,7	117.6	117.7
<b>CB-CG-CD</b>	ALFA	116,4	116.2	116.4
	BETA	118,2	117.6	118.8
	GENERAL	116,6	116.4	116.6
<b>CB-CD-OE1</b>	ALFA	119,4	119.2	119.5
	BETA	122,0	121.6	122.4
	GENERAL	119,7	119.6	119.8
<b>CB-CA-N</b>	ALFA	121,4	121.3	121.4
	BETA	114,5	114.4	114.6
	GENERAL	120,6	120.4	120.6
<b>OE2-CD-OE1</b>	ALFA	119,7	119.5	119.7
	BETA	113,5	113.3	113.5
	GENERAL	118,9	118.8	119.0

Tabla 4. . Intervalos de confianza por variable e indicador, en la primer columna se detalla la variable (ángulos de enlace), en la segunda columna se ubica cada indicador (alfa y beta), en la tercera columna se ubica el valor promedio, en la cuarta columna se ubica el intervalo de confianza el 95% .Léase los valores en Armstrong.

DIHEDROS	ESTRUCTURA	PROMEDIO	INTERVALO DE CONFIANZA AL 95%	
O-C-CA-CB	ALFA	134,4	134.0	134.6
	BETA	120,9	120.5	121.1
	GENERAL	132,7	132.4	132.9
O1-C-CA-N	ALFA	132,1	131.8	132.4
	BETA	158,7	157.6	159.8
	GENERAL	135,4	135.0	135.6
CA-CB-CG-CD	ALFA	135,7	135.3	136.0
	BETA	140,4	139.5	141.2
	GENERAL	136,2	135.9	136.5
O1-C-CA-CB	ALFA	158,7	158.3	159.0
	BETA	138,3	137.1	139.5
	GENERAL	156,2	155.8	156.5
CB-CG-CD-OE1	ALFA	148,2	147.8	148.6
	BETA	111,3	110.8	111.8
	GENERAL	143,7	143.3	144.1
CB-CG-CD-OE1	ALFA	116,7	116.4	117.0
	BETA	159,5	158.4	160.4
	GENERAL	121,9	121.5	122.2
O-C-CA-N	ALFA	120,6	120.5	120.7
	BETA	177,8	177.3	178.2
	GENERAL	127,6	127.3	127.8
N-CA-CB-CG	ALFA	139,4	139.1	139.7
	BETA	175,9	175.0	176.6
	GENERAL	143,9	143.5	144.1
phi	ALFA	133,2	132.9	133.5
	BETA	141,2	140.2	142.2
	GENERAL	134,2	133.8	134.4
psi	ALFA	138,8	138.4	139.0
	BETA	133,1	139.1	142.0
	GENERAL	138,1	137.7	138.3

Tabla 5. . Intervalos de confianza por variable e indicador, en la primer columna se detalla la variable (ángulos dihédricos), en la segunda columna se ubica cada indicador (alfa y beta), en la tercera columna se ubica el valor promedio, en la cuarta columna se ubica el intervalo de confianza el 95% .Léase los valores en Armstrong.



La cantidad de glutamatos presentes en hélices alfa en relación con la cantidad de datos presentes en hojas betas se puede explicar a partir de los estudios realizados por Levitt en 1978, quien planteo que las preferencia de los aminoácidos [89] están correlacionadas con su estructura y su estereoquímica, de esta forma el glutamato prefiere conformación de alfa hélices y los aminoácidos como la valina, la isoleucina, el triptófano, la fenilalanina, treonina y la tirosina tienen preferencia por la hélices beta.[87,89].

Con respecto a los resultados obtenidos a partir de la inferencia estadística (tablas 3,4,5) son congruentes con los estudios realizados para la predicción de la estructura secundaria haciendo uso de la información de las distancias y de los ángulos de aminoácidos, dichos estudios fueron reportados por primera vez por Nagano 1973 [86] , Periti et al. 1967, Ptitsyn and Finkelstein 1983 [87], Babal and Wu 1974, Wu and Kabal 1973 [88], aunque sus resultados se derivaron de parámetros que fueron comparativamente pequeños con respecto a la cantidad de estructuras tridimensionales reportadas, fueron pautas para el estudio realizado por Mugilan S. A. et. al. en el 2000 [85]; quienes, haciendo uso de las proteínas reportadas en PDB hasta noviembre 1996, resueltas por cristalografía de rayos X usando la sub-base de datos PDB Select y obtuvieron los parámetros de valor máximo y el valor mínimo de las dupletas y tripletas de los 20 aminoácidos que hacen parte de de la estructura tridimensional de una proteína para predecir su estructura secundaria, obtuvieron valores para las tripletas entre 102° y 132° de libertad para cada uno de los ángulos; valores que están de acuerdo con los reportados en este estudio (108° -123°) (Tabla 4); estos valores se desvían en 9° aproximadamente, esta situación puede deberse a que al 1 de septiembre la base de datos ha sido considerablemente ampliada y los parámetros de selección en este estudio

no se limitan únicamente a las proteínas resueltas por humano y que contengan glutamato, esto pudo haber generado variación en los resultados, además en el presente estudio se analizaron particularmente con el aminoácido glutamato quien hace parte de una estructura protéica en una región altamente estructurada como lo son las hélices alfa y las hojas beta, mientras que en los estudios reportados anteriormente se estudiaron los aminoácidos en conjunto cuando hacen parte de hélices; es decir no se avanza en el estudio de una estructura en particular sino en conjunto.

Con respecto a los valores de los ángulos de phi y psi (Figura 12), es importante considerar que previamente Ramachandran estableció los valores permitidos para dichos ángulos diedricos,(Tabla 6), por medio de un estudio realizado con moléculas resueltas por cristalografía de rayos X [89]. Estos resultados en contraste con los obtenidos en el presente estudio determinan que en el indicador alfa los valores de referencia del ángulo phi difieren en  $9.914^{\circ}$  (Tabla 6) y con respecto del ángulo psi para el mismo indicador los valores difieren en  $5.416^{\circ}$  (Tabla 6). Con respecto a el indicador beta los valores de referencia del ángulo phi difieren en  $1.211^{\circ}$  (Tabla 6) y con respecto del ángulo psi para el mismo indicador los valores difieren en  $4.195^{\circ}$  (Tabla 6). Las variaciones en los datos pueden estar dadas por las imperfecciones que se presentan en los cristales como por ejemplo los defectos puntuales como los isomorfismos que tiene que ver con la presencia de estructuras de igual dimensión y geometría pero que su composición química es diferente.

Estas diferencias son congruentes con los resultados obtenidos de los estudios realizados por Bosco y colaboradores [92] reportan discrepancias con los valores del

mapa estérico estándar de Ramachandran; este estudio consistió en caracterizar las hélices alfa de 500 proteínas homólogas reportadas en PDB con resolución superior a 1,8Å; ellos compararon la distribución estadística de la geometría ideal ( Ramachandran) de una proteína con los datos obtenidos y mostraron variaciones en los valores de las interacciones electrostáticas por medio del estudio de la mecánica molecular; además, dividieron los ángulos phi y psi en cuatro grupos de acuerdo con la dependencia estérica: phi dependientes, phi no dependientes, psi dependientes, psi no dependientes.[90]; encontrando así que las variaciones en el mapa de Ramachandran a lo largo de la hélice alfa esta inducido por la oposición de la dirección C - N, ya que existe una pequeña entropía del amino terminal que actúa como enlace de hidrógeno, produciendo choques y cambiando así la conformación geométrica de la proteína.[91,92].

Otro estudio que muestra diferencia con los ángulos diedricos propuestos por Ramachandran es el estudio de Gunasekaran y colaboradores [90] quienes estudiaron 110 estructuras cristalográficas no homólogas con alta resolución; obtenidas del banco de datos de PDB, sin glicina y encontraron que 66 de estos residuos no estaban dentro de los rangos propuestos por Ramachandran (Tabla 6); estudiaron los ángulos diedricos por medio de algoritmos y la determinación de las desviaciones y demostraron que los residuos que se encontraban fuera de los rangos permitidos, mostraban predominancia de aminoácidos polares cargados [90], así mismo, encontraron que estas variaciones se encuentra relacionados a aminoácidos altamente conservados; tal es el caso del glutamato quien posee estas características. El presente estudio confirma los resultados encontrados por dichos autores; es importante anotar que en este trabajo se trabajo con la totalidad del set de datos de PDB.

Sin embargo, es importante resaltar que la resolución cristalográfica es un factor que permite aproximarse a los valores propuestos por Ramachandran; así como lo corrobora el estudio de Rob y colaboradores en 1997 [92], quienes estudiaron 60 ángulos diedricos extraídos de proteínas con alta resolución y los contrastaron con el estudio de Bernstein en 1977 [93], que consistió en extraer 2897 proteínas de PDB y con una resolución menor a 2.8 Å, y a su vez se comparo con el mapa de Ramachandran; los datos obtenidos demostraron que a una mayor resolución los datos se aproximan mas a el mapa estándar por lo que, los ángulos phi y psi tienen una considerable libertad a nivel conformacional pero con un numero de restricciones importantes [92].

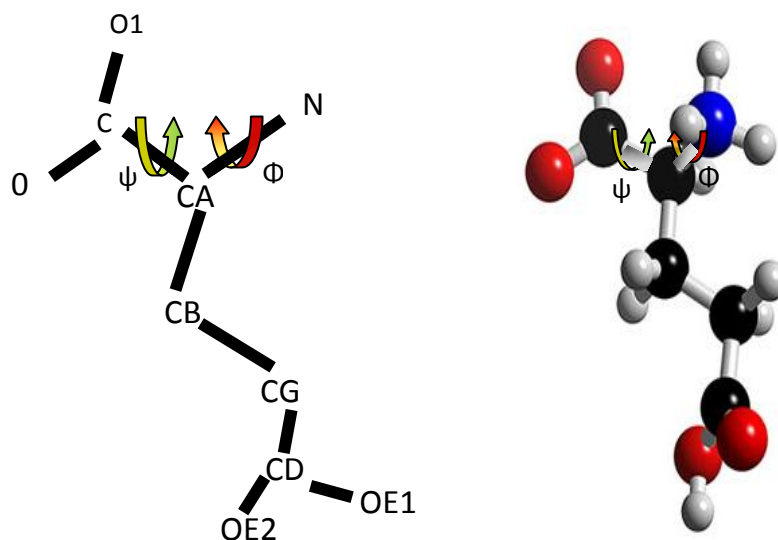


Figura 11 Ángulos diedricos phi y psi en la molécula del glutamato

Estructura	Angulo phi* propuesto por Ramachandran	Angulo phi* propuesto comparación por hipótesis de medias	Angulo psi* propuesto por Ramachandran	Angulo psi* propuesto comparación por hipótesis de medias
Hélice Alfa	120.0-123.0	132.9 - 133.5	130.0 – 133.0	138.4 - 139.0
Hoja Beta	119.0 – 139.0	140.2 - 142.2	113.0 – 135.0	139.1 - 142.0

Tabla 6. Comparación de los ángulos phi\* (N-CA-C-N) y psi\* (C-N-CA-C) propuesto por Ramachandran y propuesto por comparación de hipótesis de medias. Léase las medidas en grados.

Para complementar la inferencia estadística se realizaron graficas de dispersión para cada variable: distancia de enlace, ángulos de enlace y ángulos diedricos (Anexo H, I,J) Lo que permite mostrar de manera visual los valores de libertad en los que para el presente estudio se pueden encontrar los L-glutamatos cuando hacen parte de una proteína en una región altamente estructurada. La figura 11 muestra los gráficos correspondientes a los ángulos phi y psi en la molécula de glutamato.

A continuación (Figura 12) se presenta la grafica que representa los valores (Tabla 6) de los grados de dispersión que presenta tanto el ángulo psi como el ángulo phi para hélices alfa y hojas beta; en esta grafica se muestra que los valores de dispersión para cada ángulo no se cruzan, por ello se delimitan valores para cada caso en particular.

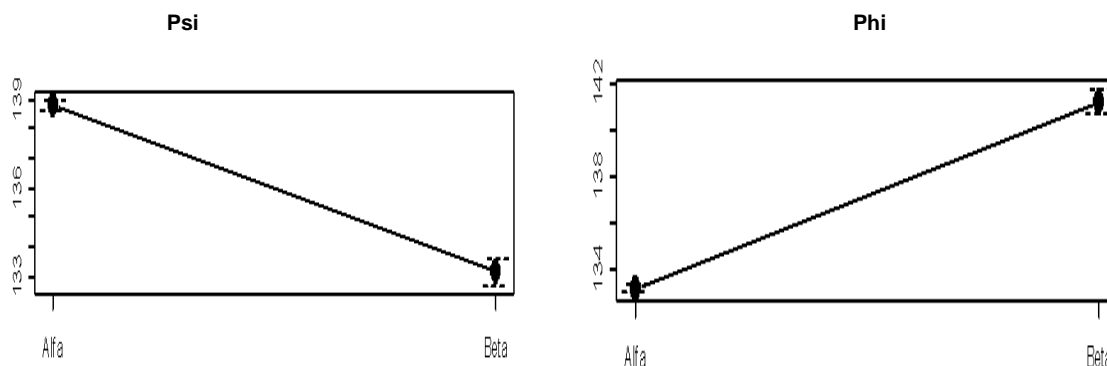


Figura 12. Comparación de los valores de dispersión de la media del ángulos psi y phi para estructuras de L-glutamato pertenecientes a hélices alfa y a hojas beta.

## 5. CONCLUSIONES

Este estudio permitió realizar un estudio estructural y conformacional de las estructuras del L-glutamato presentes en los datos reportados experimentalmente (DRX) y de los estados estructurales preferenciales cuando el glutamato hace parte la cadena aminoacídica de una estructura protéica altamente estructurada.

A partir de minería de datos de coordenadas cristalográficas y análisis estadístico del L-glutamato abstraídas de la base de datos PDB y el desarrollo e implementación del algoritmo "*Aminoacid Geometry Made Easy (AGME)*" se obtuvieron los valores de los vectores de las posiciones atómicas de los glutamatos contenidos en estructuras protéicas en hélices alfa y hojas beta. De esta manera se realiza una predicción en cuanto al comportamiento conformacional cuando dicho aminoácido hace parte de regiones altamente estructuradas. Lo anterior permite que este procedimiento sea aplicado a los demás aminoácidos, permitiendo ampliar la literatura hasta ahora existente.

A continuación se recapitulan las conclusiones relevantes del presente trabajo

1. Los valores de las variables internas (distancias de enlace, ángulos de enlace, ángulos diedros) del aminoácido L-glutamato, pueden derivarse a partir del análisis de sus coordenadas cristalográficas, por medio del desarrollo e implementación del algoritmo (*AGME*), para las proteínas depositadas en el banco de datos de proteínas PDB (septiembre 2009).

2. Existe una dependencia conformacional de las interacciones atómicas específicas del aminoácido L-glutamato cuando hace parte de una cadena aminoácida y la región estructurada (hélice alfa y hoja beta) a la cual pertenece; por lo cual las variables internas presentan diferencias estadísticamente significativas en términos de distancias de enlace, ángulos de enlace y ángulos diedros.
3. Para el caso de las distancias atómicas presenten en la molécula del glutamato, se presenta una variación con un rango pequeño (teniendo en cuenta la diferencia entre el mayor y el menor de los datos), lo que permite evidenciar que ninguna de estas relacionada con la estructura específica para hélice alfa y hoja beta.
4. Los valores obtenidos para los ángulos diedros phi y psi en estructuras protéicas pertenecientes a regiones hélice alfa y hojas beta, con relación a los valores permitidos por el plot de Ramachandran para los mismos indicadores; presentan una desviación de  $9.9^\circ$  y  $5.4^\circ$  respectivamente para el caso de hélices alfa y para el caso de las hojas beta  $1.2^\circ$  y  $4.1^\circ$  respectivamente, dicha situación puede ser a que el L- glutamato es un aminoácido polar cargados y altamente conservado, características que lo hacen preferente para ubicarse en una región no permitida, ya que presenta flexibilidad en su acomodación espacial gracias a las fluctuaciones presentes en las fuerzas estéricas.
5. A partir del análisis estadístico realizado en el presente estudio se estableció que los valores que presentan los L-glutamatos para los ángulos phi que perteneces a hélices alfa están entre  $132.9^\circ$  y  $133.5^\circ$ , los valores de los ángulos phi

pertenecientes a hojas beta están entre  $140.2^\circ$  y  $142.2^\circ$ ; para los ángulos psi pertenecientes a hélices alfa están entre  $138.4^\circ$  y  $139.0^\circ$  y para ángulos psi pertenecientes a hojas beta están entre  $139.1^\circ$  y  $142.0^\circ$ . Valores que no corresponden al plot de Ramachandran ya que el L-glutamato presente características fisicoquímicas particulares que le permiten adoptar flexibilidad en una acomodación espacial, demostrando de esta manera que existen diferencias estadísticas significativas dentro de dichos valores, ya que se aleja de las zonas permitidas propuestas por Ramachandran.



## 6. PERSPECTIVAS

El estudio de las propensiones estructurales de los aminoácidos como constituyente de una estructura protéica ya sea perteneciente a una región alfa o a una región beta nos permitirá aproximarnos al conocimiento de su estructura tridimensional y redundando en un mejor entendimiento de diferentes procesos biológicos como mecanismo de acción de varios cofactores involucrados en diferentes rutas metabólicas en las cuales diferentes moléculas interactúan conservando su identidad química, permitiendo de esta forma comprender la acción de ciertas drogas: con ello, será más factible diseñar y sintetizar nuevas drogas con mayores efectos terapéuticos y pocos o ningún efecto colateral.

El cálculo de las propensiones internas del glutamato como modelo en el presente estudio cuando pertenece a regiones altamente estructuradas, permitió construir un inferencia estadística que puede potencialmente ser aplicado con cada uno de los aminoácidos y predecir su comportamiento estructural tridimensional.

## 7. BIBLIOGRAFIA

1. Branden, C., Tooze, J. Introduction to protein structure. New York : Garland Publishing, 1991.
2. Tomas E. Oreighton. Proteins: Structures and Molecular Principles. MC Graw Hill. 1940.
3. Koehl Patrice & Levitt Michael. Structure-based conformational preferences of aminoacids. October 26, (1999). PNAS 12524-12529.
4. Robert L. Mcfeeters & Robert Oswald. Emerging structural explanation of ionotrópico glutamate receptor function. (2004). FASEB Journal. **18**. 428-438.
5. Mathew Christopher y Van Holde K. E. Bioquimica. Mc Graw Hill. 1992.
6. Lee, K. Y., A. K. Chou, L. C. Yang & H. Buerkle . NMDA receptors offer more than one functionality. (2003). Anesth Analg. **96**, 1553-1534.
7. Hollman, M. Hartley. M, Heimemann, S. Calcium permeability of KA-AMPA-gated glutamate receptor channels depends on subunit composition. (1991). Science **252**. 851-853.
8. Kandel, E. R. & H. J. Schwartz. Principles of Neural Science. Elsevier. New York. 1985.
9. Nicolls, D. G. The release of glutamate from synaptic terminals. (1995). Stone, T. W. (Ed) CNS Neurotransmitters and neuromodulators: Glutamate. CRC.
10. Brose, N. G. W. Huntley, Y. Stern- Bach. G. Sharma, J. H. Morrison & S. F. Heinemann differential Assembly of coexpressed glutamate subunits in neurons of rat cerebral cortex. (1994). J. Biol. Chem. **269**. 16780-16784.
11. Paul H. Brookes. Baltimore: Huttenlocher, P.R., & a.S. Dabholkar. Regional differences in synaptogenesis in human cerebral cortex. (1997). Journal of Comparative Neurology. **367**. 167-178.

12. Hua X, Malarkey EB, Sunjara V, Rosenwald SE, Li WH, Parpura. Calcium dependent glutamate release involves two classes of endoplasmic reticulum, calcium stores in astrocytes. (2004). *J. Neurosci Res* **76**: 86-97.
13. Felipo, V. & R. F. Butterworth. Neurobiology of ammonia. (2002). *Prog. Neurobiol.* **67**. 259-279.
14. González, J. Identificación In silico de los factores involucrados en la interacción del glutamato y sus receptores. (2009). PUJ.
15. Yeaman, C. K. K. Grindstaff, J. R. Wright W. J. Nelson complexes on trans-Golgi network and plasma membrane regulate late stages of exocytosis in mammalian cells. (2001). *J. Cell Biol.* **155**, 593-604.
16. M. Mar Albà, Mauro F. Santibáñez-Koref, John M. Hancock. Amino Acid Reiteration in Yeast Are Overrepresented in Particular Classes of Proteins and Show Evidence of a Slippage-Like Mutational Process. (1999). *J Mol Evol* **49**: 789-797.
17. Dyer, K. F. The quiet revolution: A new synthesis of biological knowledge. (1971). *Journal of Biological Education* **5**:15-24
18. Robert B. Russell, Matthew J. Betts & Michael 0052. Barnes. Amino acids Properties. (2003). *Bioinformatics for Geneticists*, M.R. Barnes, I.C. Gray eds, Wiley,
19. Lee-Wei Yang, Eran Eyal, Chakra Chennubhotla, JunGoo Jee, Angela M. Gronenborn, and Ivet Bahar Structure. Author manuscript Insights into Equilibrium Dynamics of Proteins from Comparison of NMR and X-Ray Data with Computational Predictions; (2007) *PMC Structure.* **10**:10-16.
20. Wyss DF, Wang YS, Eaton HL, Strickland C, Voigt JH, Zhu Z, Stamford AW.. Combining NMR and X-ray Crystallography in Fragment-Based Drug Discovery: Discovery of Highly Potent and Selective BACE-1 Inhibitors. (2009). *JMol.* PMID: 21647837.
21. Kuzmanic A, Kruschel D, van Gunsteren WF, Pannu NS, Zagrovic B. Dynamics May Significantly Influence the Estimation of Interatomic Distances in Biomolecular X-ray Structures. (2011). *JMol.* PMID: 21645520.

22. Baranova EV, Weeks SD, Beelen S, Bukach OV, Gusev NB, Strelkov SV. Three-Dimensional Structure of  $\alpha$ -Crystallin Domain Dimers of Two Human Small Heat Shock Proteins, HSPB1 and HSPB6'. (2008). *JMol*, PMID: 21641913.
23. Acevedo, O. E.; Lareo, Leonardo R. Comparación entre las estructuras tridimensionales de proteínas obtenidas por resonancia magnética nuclear y por difracción de rayos X. (1998). *Revista de la Asociación Colombiana de Ciencias Biológicas (Bogotá)*, **10**, No. 01-02.
24. Jiménez, M.A., Nieto, J.L., Rico, M. Determinación estructural de péptidos. (1997). *Péptidos en biología & biomedicine. CSIC*.
25. Wright, P.E., Dyson, H.J., Lerner, R.A. Conformation of peptide fragments of proteins in aqueous solution: implications for initiation of protein folding. (1988) *Biochemistry* **27**, 7167-7175.
26. Pioto, M., Saudek, V., Sklenar, V.J. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. (1992). *Biomol. NMR* **6**, 661-665.
27. Hore, P.J. *Nuclear Magnetic Resonance*. New York : Oxford University Press, 1995.
28. Wüthrich, K. *NMR of Protein and Nucleic Acids*. New York : Willey, 1986.
29. Bradford H. Carl. L. Schmidt, C. The history of discovery of the amino acids. (1931). *Chem. Rev.* **9**: 169-318.
30. Sheehan, J.C., Hess, G.P. A new method of forming peptide bonds. (1955). *J. Am. Chem. Soc.* **77**, 1067-1068.
31. Avbelj Franc. Amino Acid Conformational Preferences and Solvation of Polar Backbone Atoms in Peptides and Proteins. (2000). *J. Mol. Biol.*;1335-1359.
32. Avbelj, F. & Moulton, J.. The conformation of folding initiation sites in proteins determined by computer simulation. (1995). *Proteins: Struct. Funct. Genet.* **23** , 129 - 141.

33. Crippen, G.M., Havel, T.F. Distance Geometry and Molecular Conformation. Taunton, (England) : Research Studies Press, 1988.
34. Robert J. Anderson Main-chain conformational tendencies of Amino Acids. (1996). J. Mol. Biol. 2005; **60**:679-689.
35. Ramachandran GN, Ramakrishnan C. Stereochemistry of polypeptide chain configurations.(1991). J Mol Biol 1963; **7**:95–99.
36. Nicolls, D. Stone, T.W. (Ed) CR. Press. Boca Raton. 1995.
37. Miller, S. L.. Prebiotic synthesis in atmospheres containing CH<sub>4</sub>, CO, and CO<sub>2</sub> Orig. (1974). Life **5**, 139-151.
38. Bradford, H. Carl, L., Shmidt, C. The history of the discovery of the amino acids.(1931). Chem. Rev. **9**:169-318.
39. Harada, K.. Synthesis of amino acids and peptides under possible prebiotic conditions. (1914). Chemistry and Biochemistry of amino acids, peptides and proteins. **2**: 297-351.
40. Young VR, Ajami AM. Glutamate: An amino acid of particular distinction. (2000). J Nutr **130**: 892S-900S.
41. Hill, A., Bohler, C., Orgel, L. Polimerization on the rocks: negarively charged alpha-amino acids. (1988). Origins Life Evol. Biosph **3**: 35-49.
42. Nakanishi S. Molecular diversity of glutamate receptors and implications for brain function. (1992). Science **258**: 597-603.
43. Helen M. Berman<sup>1,2</sup>, John Westbrook<sup>1,2</sup>, Zukang Feng<sup>1,2</sup>, Gary Gilliland<sup>1,3</sup>, T. N. Bhat<sup>1,3</sup>, Helge Weissig<sup>1,4</sup>, Ilya N. Shindyalov<sup>4</sup> and Philip E. Bourne<sup>1,4,5,6</sup>. The (2000). Protein Data Bank. Nucleic Acids Research, **28**, No. 1.

44. Hill, A. V.. The mode of action of nicotine and curare determined by the form of the concentration curve and the method of temperature coefficients. (1909). *J. Physiol.* **39**, 361-373.
45. Torda, A.E., Gunsteren, W.F. *Reviews in Computational Chemistry, Vol III.* New York : VCH Publishers, 1992.
46. Walther D, Cohen FE. Conformational attractors on the Ramachandran map . *Acta Crystallogr D Biol Crystallogr* (1999); **55**(2):506-517
47. Pal D, Chakrabarti P. Cis peptide bonds in proteins: residues involved , their conformations, interactions and locations.(1999). *J Mol. Biol*; **294**:271–288.
48. Jiménez, M.A., Bruix, M., Gonzalez, C., Blanco, F.J., Nieto, J.L., Herranz, J., Rico, M. CD and <sup>1</sup>H-NMR studies on the conformational properties of peptide fragments from the C-terminal domain of thermolysin. (1993). *Eur. J. Biochem.* **211**, 569-581.
49. Mayer Mark. Glutamate Receptors at atomic resolution. (2006). *Nature.* **440.** 456-462.
50. Meeker Rick. Capitulo 11. Glutamate receptor Autoradiography and in situ Hybridization. North Carolina. Chapel Hill.1996.
51. Nicolas J. Maragakis & Jeffrey D. Rothstein. Glutamate transporter: animal models to neurologic disease. (1998). *Neurobiology of Disease.* **15:** 461-473.
52. Emanuel Jaspard. A computational analysis of three isoforms of glutamate dehydrogenase reveals structural features of the isoform EC 1.4.1.4. supporting a key role in ammonium assimilation by plants. *Biology Direct* (2006): *J. Physiol.* **1:**38-48.
53. Corredor, C. Motifs Elements in Protein Structure and Function: A Proposal.Rev. (1998). *Acad. Ciencias. Exactas, Físicas y Naturales:***12;** 519-525.

54. Harada, K.. Synthesis of amino acids and peptides under possible prebiotic conditions. *Chemistry and Biochemistry of amino acids, peptides and proteins.* (1914). *J. Physiol.* **2**, 297-351.
55. Danbolt, N. Glutamate uptake. (2001). *Prog Neurobiol.* **65**:1-105.
56. Dingle, R., MacBain, C.R.. Glutamate and aspartate. *Basic Neurochemistry: Molecular, Cellular and Medical Aspects.*, 6th Ed., Lippincott-Raven Publishers, Philadelphia. 315-333. 1999.
57. Meldrum B. y Garthwaite J. *Trends . Pharmacol*, (1970): *J. Physiol.* **11**, 379-387.
58. Smith, J.L., Hendrickson, W.A., Honzatko, R.B., Sheri, S. Structural heterogeneity in protein crystals. *Biochemistry.* (1986.). *J.Physiol.* **25** 5018-5027.
59. Gonzalez, J., Corredor, C., Lareo, LR. Theoretical Studies of Glutamate Enantiomers at Different pH Conditions. (2007). *Astrobiology*; **7** : 487-488.
60. Nakamura, H. Roles of electrostatic interactions in proteins. (1996). *Quart. Rev.Biophys.*, **29**: 12- 23.
61. Sheinerman, F., Norel, R., Honing, B. Electrostatic aspects of proteinprotein interactions. *Curr.* (2000). *Opin. Struct. Biol.* **10**:153-159.
62. Burkley, SK., Petsko, G. Amino-aromatic interactions in proteins, (1986). *FEBS Lett.* **203**:139-143.
63. Lareo, LR; Gonzalez, J. Intramolecular Excited Energy Transfer Pathways in Proteins. (2008) .*Journal of Theoretical and Computational Chemistry.* **7**: 91-102.
64. Trevor P. Creamer and George D. Rose. Alpha helix forming propensities in peptides and proteins.(2000). *Proteins: structure, function and genetics* **19**: 85-97.
65. Leszczynski, J. F. and Rose, G. D. *Science of Life.* 1986.

66. Ring, C. S., Kneller, D. O., Langridge, R. and Cohen, E. E. Glutamate in the receptors. (1992) *J. Mol. Biol.*, 224,685.
67. Richardson, and Richardson, D. C. EGPred: prediction of Eukariotic Genes Using Ab Initio Methods After combining with sequence similarity Approaches. (1989) *J. Mol.* **14**, 304 -317.
68. Rose, G. D., Winters, R.H., Wetlaufer, D.B. Recent advances in gene structure prediction.(1976) *J. Mol* **63**, 10 -26
69. Wilmot, C. M. and Thornton, M. The role of glutamate in neurodegenerative diseases and potential opportunities for intervention. (1988).*J Mol. Biol.* **23**, 203 - 221.
70. Lewis, P. N., Momany, F. A. and Seheraga, S. A. (1971) *Proc. Nat. Acad. Sci. USA*, **68**, 2293.
71. Skolnick, J. and Kolinski, A. (1990) *Science*, 250, 1121.
72. Tobias, D, Sneddon, J. F. and Brooks III, C. L. Groningen Molecular Simulation. (1987). *J. Mol. Biol.* **8**, 216, 783.
73. Padmanabhan, S. Marqusee, S. Ridgeway, T, Laue, T. M. Baldwin, R. L. Starlinght – chain non polar amino acids are good helix formers in water. (1997). *J Mol. Biol.*219:135-137.
74. Zimm, B. H., Bragg, J. K. Theory of phase transition between helix and random coil in polypeptide chains.(1959) *J. Chem. Phys.* **31**:536-535.
75. O'Neil, K. T., DeGrado, W.F. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids.(1990). *Science* **250**: 646-651.
76. Blaber, M. Zhang, X. J., Mathews, B. W. Structural basis of amino acid alpha helix propensity.(1993) *Science* **260**: 1637-1640.
77. Blaber, M. Zhang, X. J., Lindstrom, J. D. Pepoit, S.D. Baase, W. A. Mathews, B. M. Determination of alpha helix propensity within the context of a folded protein: Sites 44 and 131 to bacteriophage T4 lysozyme. (1994). *J. Mol. Biol.* **235**: 600 – 624.



78. Acevedo Sarmiento, Orlando E. Estudio Comparativo entre estructuras helicoidales tipo alfa para determinar su estabilidad. PUJ. 2006.
79. Varela Juan de Dios. Elementos geométricos de la Cristalografía. Academia Colombiana de ciencias Exactas, físicas y Naturales. Universidad Nacional de Colombia. No. 9. 2000.
80. Helen M. Berman and et. Al. The Protein Data Bank. 2000. Nucleic Acids Research. **28**. No.1. 235-242.
81. Gelbin, A. Scheinder B., Clowey L. Hsieh S. H. Olson W.K. and Berman, H.M. (1996).J. Am. Chem Soc. **118**. 519-528.
82. IUPAC-IUB Join Commision on Biochemical Nomenclature. (1983) Eur. J. Biochem. **131**. 9-15.
83. Hurphey William, Dalke Andrew and Schulten Klaus. Visual Molecular Dynamics. (1996). Journal of Molecular Graphics. **14**: 33 – 38.
84. Robert B. Russell, Matthew J. Betts & Michael R. Barnes. Amino acids Properties. (2003). Bioinformatics for Geneticists, **20**: 567-589.
85. Mugilan S. A. & Veluraja K. Generation of desviation parameters for aminoacid singlets, doublets and triplets from the three-dimentional struture of proteins and its implications for secondary struture prediction Fromm amino acid sequence.(2004). J.Phisiol. **5**: 678-698.
86. Kabat E. A. and Wu T.T. Further comparison of predict experimentally determined structure of adenylate kinase: Proc. (1974). Natl. Acad. Sci. USA. **71**: 4217- 4220.
87. Nagano K. Logical Analysis of the mechanism of protein folding. (1973). J. Mol. Biol. **75**. 401- 420.
88. Pittsyn O. B. and Finkelstein A. V. Theory of protein secondary structure and algorithm of this prediction (1983.) Biopolymers. **22**: 15 -25.

89. Levitt Michael. Conformation Preferences of Amino Acids in Globular Proteins?. (1978). *Biochemistry*. **20**, 4275 – 4285.
90. K. Gunasekaran, C. Ramakrishnan and P. Balaram. Disallowed Ramachandran Conformations of Amino Acid Residues in Protein Structures. (1996). *J. Biol. Chem.* **271**, 191-195.
91. Rob W.W.Hoof, Chris Sander and Gerrit Vriend. Objectively judging the quality of a protein structure from a Ramachandran plot. (1997). *Biochemistry*.: **13**, 4: 425-430.
92. Bosco K. Ho, Annick Thomas & Brasseur Robert. Revisiting the Ramachandran plot: Hard-sphere repulsion, electrostatics and H-bonding in the alpha helix. (2003). *Protein Science*. **12**: 2508 – 2522.
93. Hoof, Rob, Sander Chris & Vriend Gerrit. Objectively judging the quality of a protein structure from Ramachandran plot. (1997). *J. Physiol.* **13**. 425 – 430.

## 8. ANEXOS

### Anexo A. ALGORITMO AGM (Aminoacid Geometry Made Easy)

- Leer archivo de cada modelo. leer columnas de coordenadas x,y,z.
- Generar vector (x,y,z), (rmn)
- Generar las parejas de vectores: ver tabla de L,s
- Generar el producto escalar luego calcular sen de 2 y de 1.
- Generar tripletes de vectores: ver tablas de A's
- Generar  $\frac{r_{ij} r_{ki}}{(r_{ij}r_{ij})^{1/2}(r_{kj}r_{kj})^{1/2}}$  luego calcular  $\cos^{-1}$  de 2.
- Generar las cuadrupletas de vectores: ver tabla T's
- Generar  $\frac{r_{ij}(r_{ik}r_{ik})(r_{jk}r_{jk})^{1/2}}{|r_{ij}r_{jk}||r_{jk}r_{jk}|}$  luego calcular  $\sin^{-1}$  de 3
- Agrupar en las tablas individuales los datos de 1,2, y 3 por cada tipo.
- Calcular todas las estadísticas descriptivas de statistic 6.0 para cada tabla.

Datos

1HG-CG-CD-OE2

2HG-CG-CO-OE1

CG-CD-OE1-OE2

CG-CD-OE1-HE

HE-OE1-CD-OE2

$$Lr_{ij}(r_{ij}r_{ij})^{1/2}(x_{ij}^2 + y_{ij}^2 + z_{ij}^2)^{1/2}$$

$$[(x_i x_j)^2 + (y_i y_j)^2 + (z_i z_j)^2]^{1/2}$$

$$A_{ijk} = \cos^{-1} \left( \frac{r_{ji}r_{jk}}{r_{ij}r_{ik}} \right) = \cos^{-1} \left[ \frac{x_{ij}x_{jk} + y_{ji}y_{jk} + k_{ji}k_{jk}}{r_{ij}r_{ik}} \right]$$

$$= \cos^{-1} \left[ \frac{x_{ij}x_{jk} + y_{ji}y_{jk} + k_{ji}k_{jk}}{(x_{ji}^2 + y_{ij}^2 + z_{ij}^2)^{1/2}(x_{jk}^2 + y_{jk}^2 + z_{jk}^2)^{1/2}} \right]$$

$$= \cos^{-1} \left[ \frac{(x_j - x_i)(x_j - x_k) + (y_j - y_i)(y_j - y_k) + (z_j - z_i)(z_j - z_k)}{\left[ (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \right]^{1/2} \left[ (x_j - x_k)^2 + (y_j - y_k)^2 + (z_j - z_k)^2 \right]^{1/2}} \right]$$

Programar los cálculos de las variables internas: x,y,z para cada átomo.

Longitud del enlace L, Ángulos de enlace, A Ángulos de torsión T

Anexo B.DATOS BASICOS DE LAS PROTEINAS DE HUMANO RESUELTAS POR CRISTALOGRAFIA DE RAYOS X.

PDB ID	RESOLUCION CRISTALOGRAFICA	LONGITUD DE LA SECUENCIA	# MONOMEROS	REFERENCIA BIBLIOGRAFICA
2ZQS	1.90 Å	6812	8	Not Available.
3EGS	3.60 Å	5873	7	peptide on the HIV-1 gp41 epitope recognized by the broadly neutralizing antibody 2F5. (2009) J.Phys.Chem.B 113: 13626-13637
3EZZ	2.90 Å	6881	9	Crystal structure of the catalytic domain of human MKP-2 reveals a 24-mer assembly. (2009) Proteins 76: 763-767
3FEC	1.49 Å	5873	8	Structural insight into the evolutionary and pharmacologic homology of glutamate carboxypeptidases II and III (2009) Febs J. 276: 4448-4462
3FED	1.29 Å	5819	4	Structural insight into the evolutionary and pharmacologic homology of glutamate carboxypeptidases II and III (2009) Febs J. 276: 4448-4462
3FFE	1.37 Å	5855	6	pharmacologic homology of glutamate carboxypeptidases II and III (2009) Febs J. 276: 4448-4462
3FM8	2.30 Å	7165	6	Phosphorylation-independent dual-site binding of the FHA domain of KIF13 mediates phosphoinositide transport via centaurin alpha1. (2010) Proc.Natl.Acad.Sci.USA 107: 20346-20351
3H9O	2.30 Å	2258	10	Benzo[c][2,7]naphthyridines as inhibitors of PDK-1 (2009) Bioorg.Med.Chem.Lett. 19: 5225-5228

## Anexo C. COORDENADAS x,y,z.

En la página de PDB en la sección de PDB File se encuentran cada una de las coordenadas x,y,z para cada uno de los glutamatos contenidos en cada proteína, dicha información fue abstraída haciendo uso de un script que permite tener con precisión cada una de las coordenadas.

ATOM	385	CE2	PHE	A	493	-11.951	4.532	13.152	1.00	40.92	C
ATOM	386	CZ	PHE	A	493	-13.332	4.603	13.173	1.00	41.20	C
ATOM	387	N	GLU	A	494	-8.800	5.196	7.731	1.00	38.92	N
ATOM	388	CA	GLU	A	494	-7.716	4.767	6.852	1.00	38.75	C
ATOM	389	C	GLU	A	494	-7.457	3.295	7.121	1.00	38.11	C
ATOM	390	O	GLU	A	494	-7.147	2.915	8.253	1.00	38.08	O
ATOM	391	CB	GLU	A	494	-6.433	5.555	7.103	1.00	38.96	C
ATOM	392	CG	GLU	A	494	-6.394	6.929	6.456	1.00	40.27	C
ATOM	393	CD	GLU	A	494	-5.003	7.553	6.478	1.00	41.64	C
ATOM	394	OE1	GLU	A	494	-4.004	6.799	6.359	1.00	42.42	O
ATOM	395	OE2	GLU	A	494	-4.916	8.797	6.602	1.00	42.11	O
ATOM	396	N	LEU	A	495	-7.584	2.475	6.080	1.00	37.50	N
ATOM	397	CA	LEU	A	495	-7.430	1.024	6.202	1.00	36.95	C
ATOM	398	C	LEU	A	495	-6.160	0.637	6.951	1.00	36.22	C
ATOM	399	O	LEU	A	495	-6.215	-0.181	7.865	1.00	36.18	O
ATOM	400	CB	LEU	A	495	-7.435	0.353	4.821	1.00	36.91	C
ATOM	401	CG	LEU	A	495	-7.286	-1.177	4.801	1.00	37.12	C
ATOM	402	CD1	LEU	A	495	-8.284	-1.832	5.748	1.00	37.30	C
ATOM	403	CD2	LEU	A	495	-7.446	-1.727	3.381	1.00	37.14	C
ATOM	404	N	GLU	A	496	-5.031	1.228	6.566	1.00	35.45	N
ATOM	405	CA	GLU	A	496	-3.742	0.856	7.155	1.00	35.04	C
ATOM	406	C	GLU	A	496	-3.678	1.156	8.649	1.00	34.22	C
ATOM	407	O	GLU	A	496	-3.367	0.264	9.436	1.00	34.32	O
ATOM	408	CB	GLU	A	496	-2.521	1.464	6.404	1.00	35.27	C
ATOM	409	CG	GLU	A	496	-2.655	2.902	5.858	1.00	36.21	C
ATOM	410	CD	GLU	A	496	-2.839	2.970	4.333	1.00	37.19	C
ATOM	411	OE1	GLU	A	496	-3.386	2.017	3.732	1.00	37.45	O
ATOM	412	OE2	GLU	A	496	-2.424	3.990	3.736	1.00	37.85	O
ATOM	413	N	LYS	A	497	-3.983	2.390	9.040	1.00	33.23	N
ATOM	414	CA	LYS	A	497	-3.908	2.779	10.453	1.00	32.60	C
ATOM	415	C	LYS	A	497	-4.856	1.925	11.324	1.00	32.06	C
ATOM	416	O	LYS	A	497	-4.512	1.556	12.453	1.00	31.78	O
ATOM	417	CB	LYS	A	497	-4.219	4.274	10.631	1.00	32.59	C
ATOM	418	CG	LYS	A	497	-3.216	5.228	9.970	1.00	32.47	C
ATOM	419	CD	LYS	A	497	-3.843	6.587	9.609	1.00	32.97	C
ATOM	420	CE	LYS	A	497	-3.775	7.613	10.737	1.00	32.93	C
ATOM	421	NZ	LYS	A	497	-2.626	8.550	10.555	1.00	33.00	N
						<b>X</b>	<b>Y</b>	<b>Z</b>			

## Anexo D. SCRIPTS EMPLEADOS PARA LA MINERIA DE DATOS.

### **SCRIPT PARA OBTENER 14371 ESTRUCTURAS PROTEÍCAS RESUELTAS POR RAYOS X.**

El script realiza la secuencia de órdenes que consiste en traer desde la base de datos de PDB, las coordenadas de cada proteína que fue resuelta por cristalografía de rayos X, leerlos, extraer a partir de allí únicamente los datos de las coordenadas x,y,z de cada proteína.

```
wget (http://www.pdb.org/pdb/files/xray.pdb)|catpdb/files/xray.pdb|grep"$4,$7$9">xray
```

### **SCRIPT PARA OBTENER LAS COORDENADAS DE x,y,z DE LOS GLUTAMATOS QUE HACEN PARTE DE DE LAS 14371 PROTEINAS RESUELTAS POR RAYOS X.**

El script realiza la secuencia de órdenes que consiste en traer desde la base de datos de PDB, las coordenadas X,Y,Z de cada glutamato perteneciente a las proteínas resueltas por rayos X, leerlos, extraer a partir de allí únicamente los datos de las coordenadas X,Y,Z de cada glutamato perteneciente a las proteínas resueltas por rayos X.

```
wget (xray.txt)|cat xray.txt|grep"$4,$7$9">glutamatos
```

### **SCRIPT PARA OBTENER LAS COORDENADAS DE LA PROTEINA Y DE LOS GLUTAMATOS QUE HACEN PARTE DE REGIONES ALTAMENTE ESTRUCTURADAS**

El script realiza la secuencia de órdenes que consiste en traer desde la base de datos de PDB, las coordenadas de cada proteína con sus respectivos glutamatos, leerlos, extraer a partir de allí únicamente los datos de las coordenadas de los glutamatos que hacen parte de regiones hélice alfa.

```
wget (http://www.pdb.org/pdb/files/i.pdb)|catpdb/files/i.pdb|grep"HELIX">helices
```

Anexo E.TABLA DE DISTANCIAS, ANGULOS Y DIEDROS DE LOS MONOMEROS ENCONTRADOS

1A65

GLU A2

Distance Bet	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6	MODEL 7	MODEL 8	MODEL 9	MODEL 10	MODEL 11
C-CA	153,021	153,014	152,962	153,005	15,301	153,039	152,99	153,07	153,01	152,978	153,017
C-H	313,058	316,427	315,084	29,543	318,384	312,787	295,94	311,18	318,55	311,794	317,856
C-HA	21,596	215,931	215,876	215,867	215,828	21,589	215,96	215,99	215,9	215,919	215,899
CA-CB	153,006	152,929	153,029	152,984	152,985	153,045	153,05	152,98	153,02	153,064	153,015
CA-N	145,334	145,34	145,32	14,527	145,21	145,24	145,29	145,38	145,3	145,322	14,534
CB-CG	153,006	153,05	152,98	153,036	152,92	153,03	153,07	153,02	153	153,026	152,973
CD-OE1	123,974	124,01	123,97	124,039	123,97	1,239	123,99	124	124	123,998	124,035
CD-OE2	128,952	129	128,89	129,018	129,04	129,03	128,99	129,09	128,99	129,034	128,925
CG-CD	152,963	153,06	153,03	152,975	152,98	153,03	152,98	15,296	153,06	153,016	152,952
O-C	122,991	123	123,03	122,995	123,07	123,03	123	122,97	123	1213,02	123,015

Anexo F.SELECCIÓN DE HOJAS ALFA Y BETA.

RANGO DE BUSQUEDA DE LOS GLUTAMATOS  
PERTENECIENTES A HOJAS ALFA

HELIX	1	1	GLY A	45	GLU A	58	1
HELIX	2	2	SER A	62	LEU A	91	1
HELIX	3	3	ALA A	99	GLY A	112	1
HELIX	4	4	THR A	120	LYS A	132	1
HELIX	5	5	ASP A	133	ASN A	141	1
HELIX	6	6	SER A	151	ASN A	157	1
HELIX	7	7	ASP A	158	ALA A	163	1
HELIX	8	8	THR A	170	ARG A	176	1
HELIX	9	9	GLU A	207	GLU A	216	1
HELIX	10	10	SER A	228	LEU A	232	5
HELIX	11	11	ASN A	241	ASN A	255	1
HELIX	12	12	LYS A	270	ILE A	278	1
HELIX	13	13	PRO A	282	CYS A	286	5
HELIX	14	14	THR A	295	ASP A	309	1
HELIX	15	15	ASP A	328	LYS A	345	1
HELIX	16	16	GLY B	45	GLU B	58	1
HELIX	17	17	SER B	62	LEU B	91	1
HELIX	18	18	ALA B	99	GLY B	112	1
HELIX	19	19	THR B	120	LYS B	132	1
HELIX	20	20	ASP B	133	ASN B	141	1
HELIX	21	21	SER B	151	ASN B	157	1
HELIX	22	22	ASP B	158	ALA B	163	1
HELIX	23	23	THR B	170	ARG B	176	1
HELIX	24	24	GLU B	207	GLU B	216	1
HELIX	25	25	SER B	228	LEU B	232	5
HELIX	26	26	ASN B	241	ASN B	255	1
HELIX	27	27	LYS B	270	ILE B	278	1
HELIX	28	28	PRO B	282	CYS B	286	5
HELIX	29	29	THR B	295	ASP B	309	1

SHEET	1	AA	6	VAL A	185	PHE A	191	0									
SHEET	2	AA	6	LEU A	194	ASP A	200	-1	O	LEU A	194	N	PHE A	191			
SHEET	3	AA	6	GLU A	33	GLY A	40	1	O	VAL A	34	N	LYS A	197			
SHEET	4	AA	6	ALA A	220	ALA A	226	1	O	ALA A	220	N	LEU A	37			
SHEET	5	AA	6	SER A	263	ASN A	269	1	O	SER A	263	N	ILE A	221			
SHEET	6	AA	6	ILE A	319	PHE A	323	1	O	TYR A	320	N	LEU A	266			
SHEET	1	BA	6	VAL B	185	PHE B	191	0									
SHEET	2	BA	6	LEU B	194	ASP B	200	-1	O	LEU B	194	N	PHE B	191			
SHEET	3	BA	6	GLU B	33	GLY B	40	1	O	VAL B	34	N	LYS B	197			
SHEET	4	BA	6	ALA B	220	ALA B	226	1	O	ALA B	220	N	LEU B	37			
SHEET	5	BA	6	SER B	263	ASN B	269	1	O	SER B	263	N	ILE B	221			
SHEET	6	BA	6	ILE B	319	PHE B	323	1	O	TYR B	320	N	LEU B	266			

RANGO DE BUSQUEDA DE LOS GLUTAMATOS  
PERTENECIENTES A HOJAS BETA



Anexo G. COMPOSICION DEL SET DE DATOS

**DISTANCIAS DE GLUTAMATOS PERTENECIENTES A REGIONES ALFA**

DATOS GRALES		DISTANCIAS					
PROTEINA	UBICACION	O-C	C-01	C-CA	CA-N	CA-CB	CB-CG
9PAI	148	1.5	1.5	1.5	1.4	1.5	1.2
9PAI	309	1.5	1.5	1.5	1.4	1.5	1.2
9PAI	71	1.5	1.5	1.5	1.4	1.5	1.2
9PAI	99	1.5	1.5	1.5	1.4	1.5	1.2
121P	153	1.5	1.5	1.5	1.4	1.5	1.2
121P	162	1.5	1.5	1.5	1.4	1.5	1.2
121P	91	1.5	1.5	1.5	1.4	1.5	1.2
12CA	14	1.4	1.5	1.5	1.4	1.5	1.3
12CA	221	1.5	1.5	1.5	1.4	1.5	1.3
12GS	112	1.5	1.5	1.5	1.4	1.5	1.2
12GS	130	1.5	1.5	1.5	1.4	1.5	1.2
12GS	163	1.5	1.5	1.5	1.4	1.5	1.2
12GS	197	1.5	1.5	1.5	1.4	1.5	1.2
12GS	36	1.5	1.5	1.5	1.4	1.5	1.2
12GS	40	1.5	1.5	1.5	1.4	1.5	1.2
12GS	85	1.5	1.5	1.5	1.4	1.5	1.2
12GS	97	1.5	1.5	1.5	1.4	1.5	1.2
14GS	112	1.5	1.5	1.5	1.4	1.5	1.2
14GS	130	1.5	1.5	1.5	1.4	1.5	1.2
14GS	163	1.5	1.5	1.5	1.4	1.5	1.2
14GS	197	1.4	1.5	1.5	1.4	1.5	1.2
14GS	85	1.5	1.5	1.5	1.4	1.5	1.2
14GS	97	1.5	1.5	1.5	1.4	1.5	1.2
17GS	112	1.5	1.5	1.5	1.4	1.5	1.2
17GS	130	1.5	1.5	1.5	1.4	1.5	1.2
17GS	163	1.5	1.5	1.5	1.4	1.5	1.2
17GS	197	1.4	1.5	1.5	1.4	1.4	1.2
17GS	36	1.5	1.5	1.5	1.4	1.5	1.2
17GS	40	1.5	1.5	1.5	1.4	1.5	1.2
17GS	85	1.5	1.5	1.5	1.4	1.5	1.2
17GS	97	1.5	1.5	1.5	1.4	1.5	1.2
18GS	112	1.5	1.5	1.5	1.4	1.5	1.2
18GS	130	1.5	1.5	1.5	1.4	1.5	1.2

## ÁNGULOS DE GLUTAATOS PERTENECIENTES A REGIONES ALFA

DATOS GRALES		ANGULOS			
PROTEINA	UBICACION	O-C-O1	O-C-CA	01-C-CA	C-CA-N
9PAI	148	113.3	119.3	118.3	115.6
9PAI	309	113.5	119.5	118.0	110.5
9PAI	71	112.8	118.6	117.8	112.1
9PAI	99	113.2	118.8	118.5	120.1
121P	153	116.5	118.8	114.8	114.9
121P	162	114.3	118.8	118.1	109.8
121P	91	116.7	118.8	116.0	113.0
12CA	14	127.3	116.4	121.8	110.6
12CA	221	115.0	110.9	121.8	111.9
12GS	112	112.8	118.3	121.8	115.7
12GS	130	112.7	118.4	118.4	120.1
12GS	163	112.7	117.6	119.0	120.1
12GS	197	111.9	118.2	119.1	110.5
12GS	36	112.1	118.2	118.5	110.3
12GS	40	112.3	118.4	118.4	107.3
12GS	85	112.9	119.5	117.8	109.7
12GS	97	113.0	119.1	118.6	111.8
14GS	112	113.4	118.5	118.5	116.0
14GS	130	113.6	119.0	118.1	107.8
14GS	163	112.9	118.8	118.2	120.1
14GS	197	112.1	119.1	117.8	110.9
14GS	85	113.8	119.4	117.6	110.0
14GS	97	112.3	119.1	116.8	112.4
17GS	112	113.7	119.1	118.0	120.1
17GS	130	113.9	119.6	118.6	106.2
17GS	163	113.3	118.9	120.1	120.1
17GS	197	111.3	118.8	118.1	112.2
17GS	36	114.4	120.2	118.0	108.5
17GS	40	112.7	118.0	118.8	108.5
17GS	85	113.5	119.0	118.0	111.4
17GS	97	113.8	119.4	118.1	111.9
18GS	112	112.5	118.4	118.2	115.6
18GS	130	113.3	118.9	118.3	120.1
18GS	163	113.7	119.6	117.8	107.6

## ANGULOS DE GLUTAMATOS PERTENECIENTES A REGIONES ALFA

DATOS GRALES		DIEDROS			
PROTEINA	UBICACION	O-C-CA-O1	O-C-CA-CB	O1-C-CA-N	CA-CB-CG-CD
9PAI	148	120.3	122.2	103.0	129.9
9PAI	309	120.8	122.4	176.1	121.9
9PAI	71	120.2	123.5	111.7	124.7
9PAI	99	120.9	122.6	111.7	121.7
121P	153	123.6	122.6	111.7	121.7
121P	162	122.0	122.6	120.1	121.7
121P	91	121.2	122.6	120.1	121.7
12CA	14	120.2	121.4	109.5	124.3
12CA	221	119.7	124.4	105.3	125.7
12GS	112	120.4	120.1	154.8	128.2
12GS	130	121.2	123.1	124.0	117.2
12GS	163	121.0	123.2	163.5	120.7
12GS	197	120.8	122.6	115.5	121.5
12GS	36	120.8	123.2	179.3	121.7
12GS	40	120.8	123.1	108.3	120.1
12GS	85	121.0	122.5	162.0	121.4
12GS	97	120.5	122.2	134.5	123.8
14GS	112	120.2	122.8	147.9	120.1
14GS	130	120.9	120.1	122.1	118.5
14GS	163	120.2	122.8	141.6	120.6
14GS	197	120.9	122.9	113.0	120.9
14GS	85	120.6	122.8	156.0	122.1
14GS	97	120.1	123.9	114.4	124.4
17GS	112	120.9	122.7	166.9	127.4
17GS	130	120.9	121.7	117.9	116.6
17GS	163	120.5	122.3	139.7	120.5
17GS	197	120.8	122.9	106.4	122.0
17GS	36	121.2	121.6	171.9	120.1
17GS	40	120.7	123.0	116.2	120.5
17GS	85	120.0	122.9	164.3	124.5
17GS	97	120.2	122.3	131.6	124.4
18GS	112	120.1	123.3	168.9	127.5
18GS	130	121.2	122.6	121.8	120.1
18GS	163	120.9	120.1	120.1	120.1

## DISTANCIAS DE GLUTAMATOS PERTENECIENTES A REGIONES BETA

DATOS		DISTANCIAS								
PROTEINA	ubicación	O-C	C-01	C-CA	CA-N	CA-CB	CB-CG	CG-CD	CD-OE1	CD-OE2
121P	143	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
121P	3	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
121P	37	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
121P	49	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
12CA	117	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
12GS	30	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
12GS	31	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
14GS	30	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
14GS	31	1.5	1.5	1.4	1.5	1.3	1.2	1.5	1.2	1.5
17GS	30	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
17GS	31	1.5	1.5	1.5	1.5	1.2	1.2	1.5	1.2	1.5
18GS	30	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
18GS	31	1.5	1.5	1.5	1.5	1.3	1.2	1.5	1.2	1.5
1A0L	107	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A12	73	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A1M	198	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A1N	198	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A1X	35	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A1X	36	1.5	1.5	1.4	1.5	1.3	1.2	1.5	1.2	1.5
1A31	348	1.4	1.5	1.4	1.5	1.2	1.2	1.2	1.2	1.5
1A31	403	1.5	1.5	1.4	1.5	1.2	1.2	1.4	1.2	1.5
1A31	526	1.5	1.5	1.4	1.5	1.2	1.2	1.2	1.2	1.5
1A36	348	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A36	403	1.4	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A36	526	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A3Q	245	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A3Q	259	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A3Q	264	1.5	1.5	1.4	1.5	1.2	1.2	1.5	1.2	1.5
1A3Q	44	1.5	1.5	1.4	1.5	1.3	1.2	1.5	1.2	1.5
1A42	117	1.5	1.5	1.5	1.5	1.2	1.2	1.5	1.2	1.5

## ANGULOS DE GLUTAMATOS PERTENECIENTES A REGIONES BETA

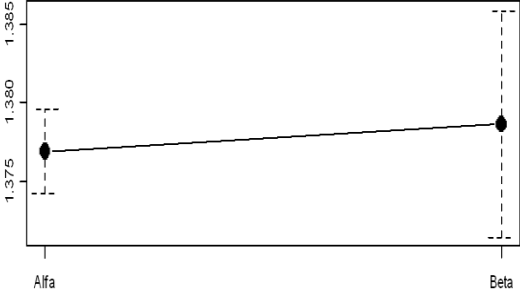
DATOS		ANGULOS					
PROTEINA	ubicación	O-C-O1	O-C-CA	01-C-CA	C-CA-N	CB-CA-N	C-CA-CB
121P	143	108.8	110.3	107.0	111.3	116.0	121.3
121P	3	110.0	106.3	111.7	110.8	1175.8	120.6
121P	37	108.0	103.9	114.4	112.85	116.9	120.8
121P	49	108.9	109.4	113.1	110.0	112.8	119.0
12CA	117	109.3	107.9	124.3	115.0	112.8	121.6
12GS	30	109.4	107.2	114.1	116.0	112.6	118.0
12GS	31	110.1	104.1	112.8	117.0	112.2	118.6
14GS	30	109.9	107.4	115.2	118.0	114.2	119.7
14GS	31	109.9	106.1	115.2	109.9	111.8	118.9
17GS	30	108.2	109.5	115.2	110.5	113.1	120.1
17GS	31	108.2	104.4	113.5	111.7	112.7	119.0
18GS	30	108.5	109.0	115.0	111.4	113.3	119.5
18GS	31	110.0	103.7	113.4	110.8	111.9	118.4
1A0L	107	109.9	106.3	112.5	111.8	113.5	119.4
1A12	73	108.5	112.7	115.3	110.2	113.2	118.8
1A1M	198	109.3	109.3	115.3	111.6	113.2	120.6
1A1N	198	108.3	112.4	115.1	111.6	112.9	118.8
1A1X	35	108.3	104.6	110.6	110.5	113.3	117.9
1A1X	36	109.9	104.6	114.5	108.1	111.9	118.9
1A31	348	109.9	108.3	114.5	109.0	113.4	118.8
1A31	403	109.6	108.0	114.0	109.0	112.0	117.8
1A31	526	108.7	108.0	114.3	110.2	112.0	118.1
1A36	348	109.5	108.0	113.5	110.1	113.3	119.9
1A36	403	108.4	108.3	114.6	109.1	113.3	118.6
1A36	526	110.1	107.9	115.3	109.0	111.9	118.0
1A3Q	245	110.1	114.0	115.8	109.2	112.1	119.1
1A3Q	259	108.6	110.4	115.1	110.2	111.8	118.7
1A3Q	264	111.2	103.5	114.3	110.5	113.9	119.7
1A3Q	44	107.9	108.6	113.8	110.5	113.1	119.5
1A42	117	109.0	110.1	119.6	110.5	114.1	119.0

## DIEDROS DE GLUTAMATOS PERTENECIENTES A REGIONES BETA

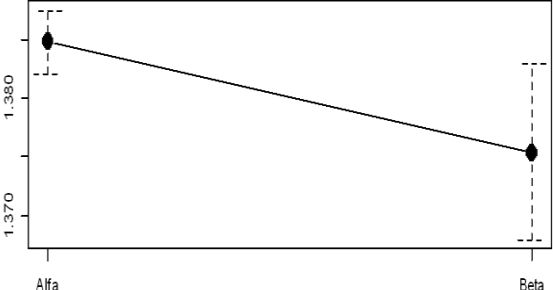
DATOS		DIEDROS				
PROTEINA	UBICACION	O-C-CA-O1	O-C-CA-CB	O1-C-CA-N	CA-CB-CG-CD	CD-CG-OE1-OE2
121P	143	168.1	121.9	155.7	172.9	173.4
121P	3	174.3	117.3	163.4	175.8	174.1
121P	37	106.6	114.3	172.0	113.1	115.6
121P	49	119.4	120.0	170.2	127.5	115.6
12CA	117	115.2	120.0	170.2	159.7	157.4
12GS	30	174.1	118.0	170.2	175.3	175.7
12GS	31	174.1	118.0	170.2	112,4	112.3
14GS	30	176.5	118.0	170.2	179.3	177.6
14GS	31	176.5	118.0	153.8	179.3	177.6
17GS	30	168.7	119.9	100.5	174.7	175.8
17GS	31	151.7	115.8	113.8	174.7	175.8
18GS	30	170.0	121.5	100.9	178.7	178.5
18GS	31	163.7	114.2	104.7	108.1	108.0
1A0L	107	121.4	117.8	166.6	140.4	140.4
1A12	73	168.1	124.2	111.7	137.1	138.1
1A1M	198	171.5	121.1	151.6	106.6	108.3
1A1N	198	161.6	123.8	179.0	106.6	133.7
1A1X	35	172.7	116.4	169.2	140.1	140.2
1A1X	36	116.7	128.5	179.9	152.8	152.3
1A31	348	116.7	118.0	102.1	155.0	154.2
1A31	403	129.6	118.1	152.9	158.2	158.0
1A31	526	109.2	122.8	152.9	157.9	156.7
1A36	348	172.7	123.2	152.9	126.3	125.4
1A36	403	140.5	117.9	172.6	168.2	167.7
1A36	526	117.2	118,0	179.3	153.6	153.8
1A3Q	245	140.4	127.3	176.5	153.6	153.8
1A3Q	259	162.1	127.3	178.0	142.7	140.2
1A3Q	264	109.3	110.5	158.2	140.0	140.4
1A3Q	44	123.0	110.5	158.2	103.2	104.1

Anexo H. GRÁFICOS DE MEDIAS DE DISTANCIAS

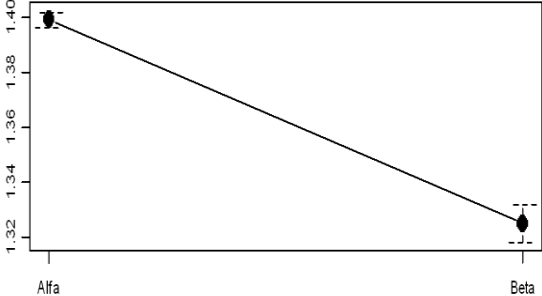
DISTANCIA O-C



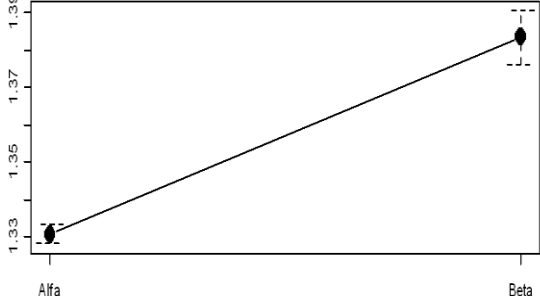
DISTANCIA C-01



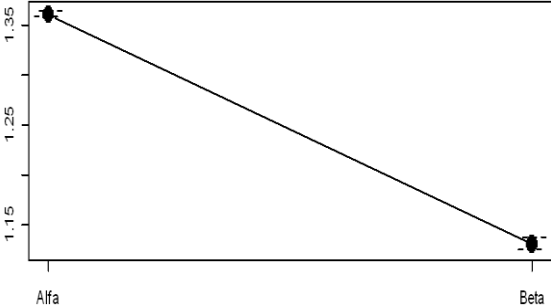
DISTANCIA C-CA



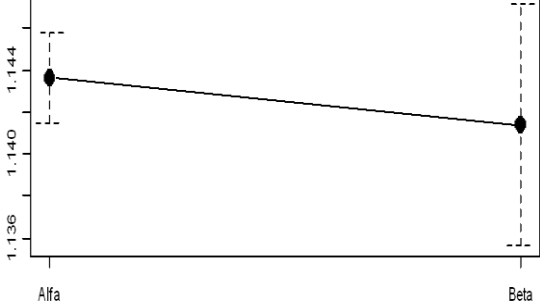
DISTANCIA CA-N



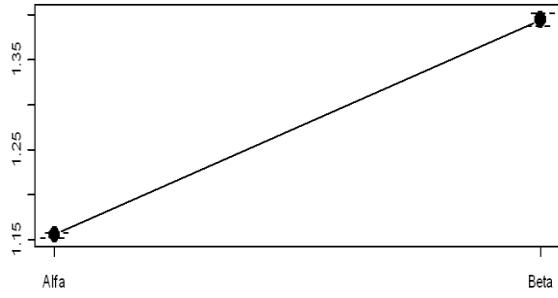
DISTANCIA CA-CB



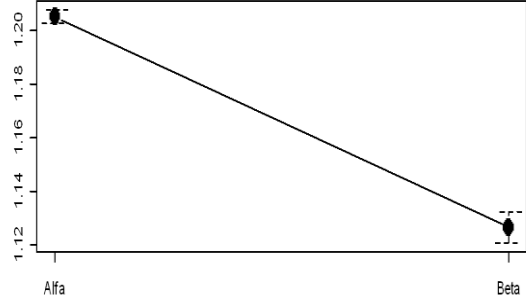
DISTANCIA CB-CG



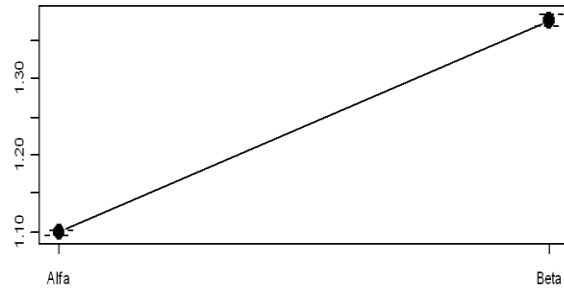
**DISTANCIA CG-CD**



**DISTANCIA CD-OE1**



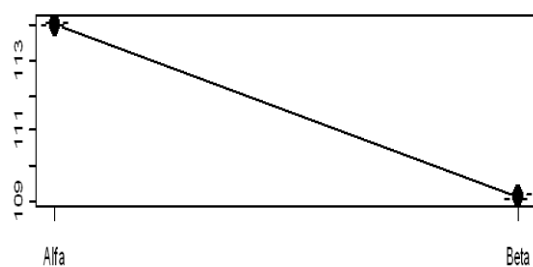
**DISTANCIA CD-OE2**



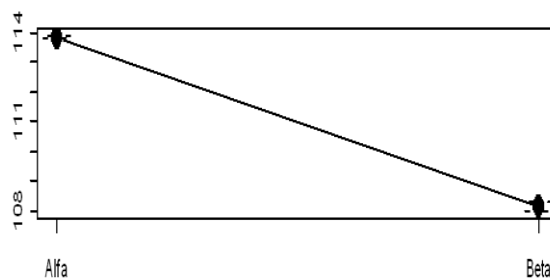


# Anexo I.GRÁFICOS DE MEDIAS DE ÁNGULOS DE ENLACE

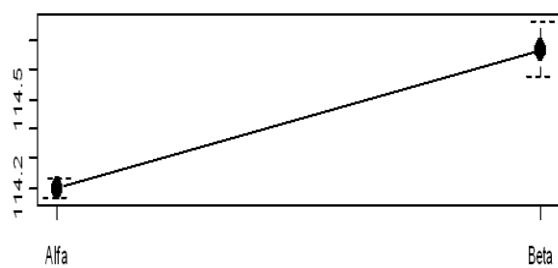
### ÁNGULO O-C-O1



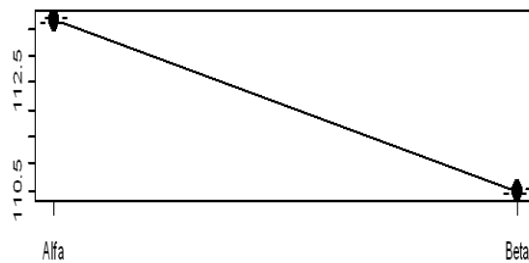
### ÁNGULO O-C-CA



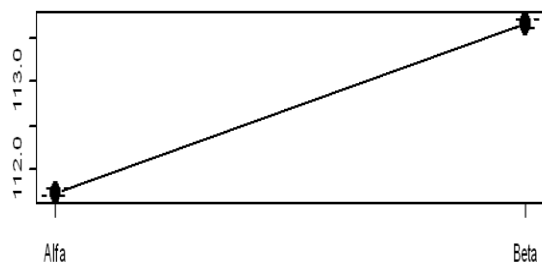
### ÁNGULO O1-C-CA



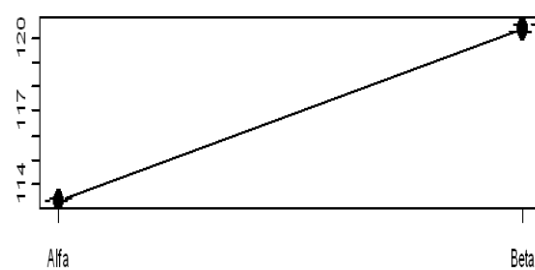
### ÁNGULO C-CA-N



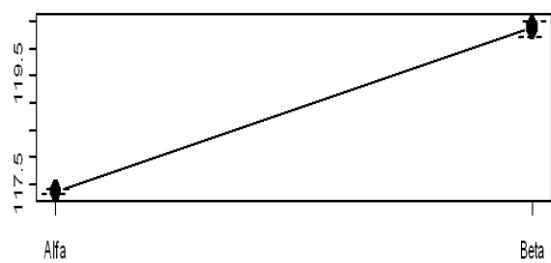
### ÁNGULO CB-CA-N



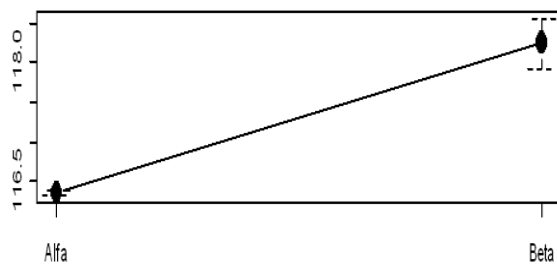
### ÁNGULO C-CA-CB



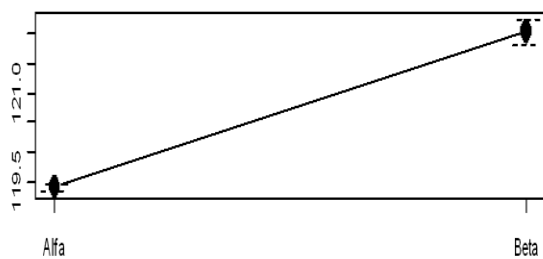
ÁNGULO CA-CB-CG



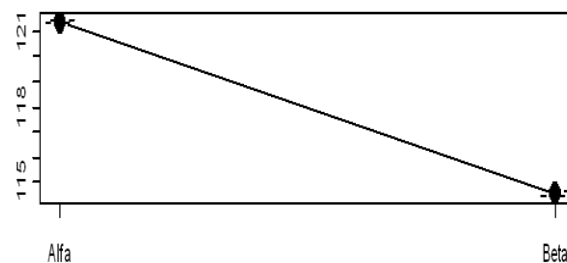
ÁNGULO CB-CG-CD



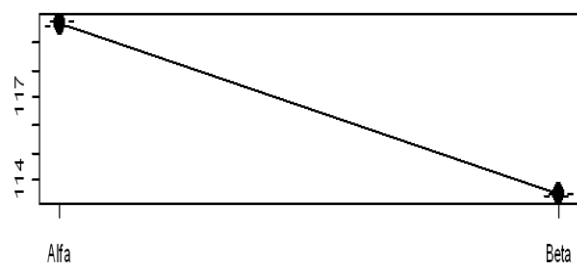
ÁNGULO CB-CD-OE1



ÁNGULO CG-CO-OE2

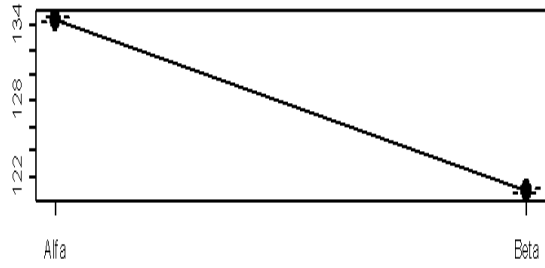


ÁNGULO OE2-CD-OE1

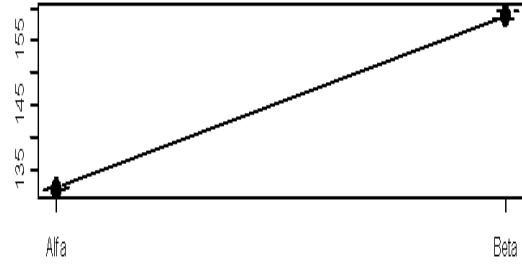


Anexo J. GRÁFICOS DE MEDIAS DE ÁNGULOS DIEDROS

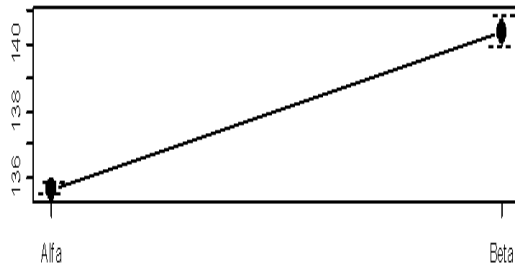
DIHEDRO O-C-CA-CB



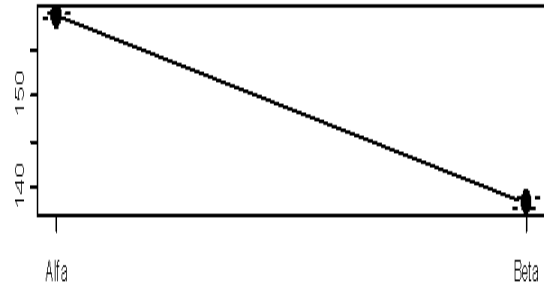
DIHEDRO O1-C-CA-N



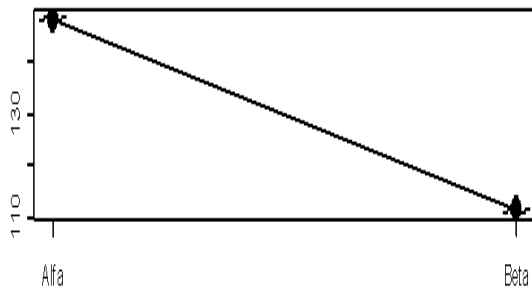
DIHEDRO CA-CB-CG-CD



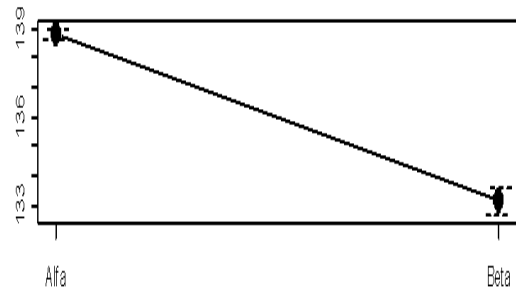
DIHEDRO O1-C-CA-CB



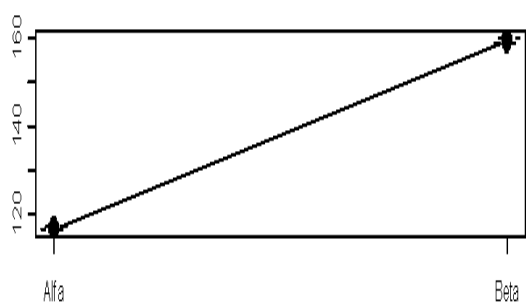
DIHEDRO CB-CG-CD-OE1



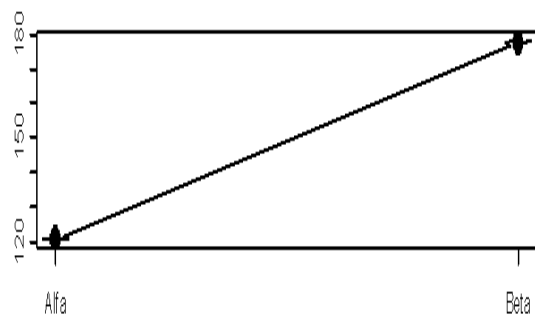
DIHEDRO N-CA-CB-CG



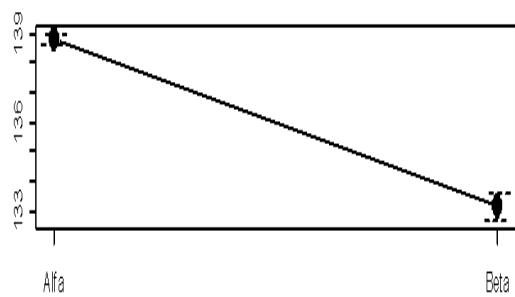
**DIHEDRO CB-CG-CD-OE2**



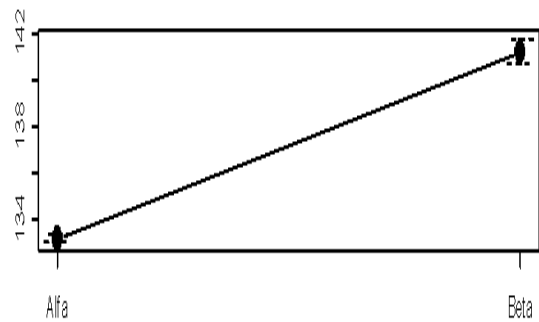
**DIHEDRO O-C-CA-N**



**Psi**



**Phi**



## Anexo K. ESTADÍSTICA POR INDICADOR

### Estadística Distancia O-C.

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645     1.3769    0.002      0.42        1.371585    1.382216
Beta  |     3412     1.378661  0.007      0.42        1.36446    1.392862
-----+-----
combined |    28057     1.377114  .0025398   .4254233    1.372136    1.382093
-----+-----
diff  |              -.0017602  .007734              -.0169228   .0134024
-----

diff = mean(Alfa) - mean(Beta)                                t = -0.2276
Ho: diff = 0                                                    Satterthwaite's degrees of freedom = 4422.49

Ha: diff < 0                                                    Ha: diff != 0                                                    Ha: diff > 0
Pr(T < t) = 0.4100        Pr(|T| > |t|) = 0.8200        Pr(T > t) = 0.5900

```

### Estadística Distancia C-01

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645     1.384755  .0026794   .4206289    1.379503    1.390007
Beta  |     3412     1.375419  .0074506   .4352054    1.360811    1.390027
-----+-----
combined |    28057     1.38362  .0025219   .4224315    1.378677    1.388563
-----+-----
diff  |              .009336   .0079177              -.0061867   .0248588
-----

diff = mean(Alfa) - mean(Beta)                                t = 1.1791
Ho: diff = 0                                                    Satterthwaite's degrees of freedom = 4340.27

```

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0  
Pr(T < t) = 0.8808                      Pr(|T| > |t|) = 0.2384                      Pr(T > t) = 0.1192

### Estadística Distancia C-CA

```
-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa |   24645   1.399283   .0027426   .4305581   1.393908   1.404659
Beta |    3412   1.325029   .0067673   .395291   1.311761   1.338298
-----+-----
combined |  28057   1.390253   .0025499   .4271084   1.385256   1.395251
-----+-----
diff |           .0742541   .0073019           .0599388   .0885693
-----
```

diff = mean(Alfa) - mean(Beta)                      t = 10.1691  
Ho: diff = 0                      Satterthwaite's degrees of freedom = 4606.35

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0  
Pr(T < t) = 1.0000                      Pr(|T| > |t|) = 0.0000                      Pr(T > t) = 0.0000

### Estadística Distancia CA-N

```
-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa |   24645   1.330804   .0025499   .4003   1.325806   1.335802
Beta |    3412   1.383417   .0070553   .4121188   1.369584   1.39725
-----+-----
combined |  28057   1.337202   .0024007   .4021163   1.332497   1.341908
-----+-----
diff |           -.0526135   .007502           -.0673212   -.0379058
-----
```

```

diff = mean(Alfa) - mean(Beta)                                t = -7.0133
Ho: diff = 0                                                  Satterthwaite's degrees of freedom = 4350.01

Ha: diff < 0                                                  Ha: diff != 0                                               Ha: diff > 0
Pr(T < t) = 0.0000      Pr(|T| > |t|) = 0.0000      Pr(T > t) = 1.0000

```

### Estadística Distancia CA-CB

```

-----+-----
Group |   Obs   Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa |  24645  1.361137  .00271117  .4257044    1.355821    1.366452
Beta |   3412  1.130627  .0059662   .3484988    1.11893     1.142325
-----+-----
combined | 28057  1.333104  .0025303   .4238239    1.328145    1.338064
-----+-----
diff |           .2305093  .0065535           .2176615  .2433572
-----+-----

```

```

diff = mean(Alfa) - mean(Beta)                                t = 35.1733
Ho: diff = 0                                                  Satterthwaite's degrees of freedom = 4936.72

Ha: diff < 0                                                  Ha: diff != 0                                               Ha: diff > 0
Pr(T < t) = 1.0000      Pr(|T| > |t|) = 0.0000      Pr(T > t) = 0.0000

```

### Estadística Distancia CB-CG

```

-----+-----
Group |   Obs   Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa |  24645  1.143622  .0021566   .3385511    1.139395    1.147849
Beta |   3412  1.141372  .0057473   .3357123    1.130103    1.15264
-----+-----
combined | 28057  1.143349  .0020191   .338202     1.139391    1.147306
-----+-----

```

```

diff |          .0022506   .0061386           -.0097841   .0142853
-----
diff = mean(Alfa) - mean(Beta)                      t =    0.3666
Ho: diff = 0                      Satterthwaite's degrees of freedom = 4426.99

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0
Pr(T < t) = 0.6430                Pr(|T| > |t|) = 0.7139                Pr(T > t) = 0.3570

```

### Estadística Distancia CG-CD

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa |   24645   1.154282   .0022499   .3532032   1.149872   1.158692
Beta |    3412   1.394505   .0067613   .3949435   1.381248   1.407761
-----+-----
combined |  28057   1.183495   .0021912   .3670276   1.1792    1.18779
-----+-----
diff |          -.2402231   .0071258           -.2541934   -.2262527
-----
diff = mean(Alfa) - mean(Beta)                      t =  -33.7117
Ho: diff = 0                      Satterthwaite's degrees of freedom = 4201.08

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0
Pr(T < t) = 0.0000                Pr(|T| > |t|) = 0.0000                Pr(T > t) = 1.0000

```

### Estadística Distancia CD-OE1

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa |   24645   1.204809   .0024534   .3851514   1.2    1.209617
Beta |    3412   1.126307   .0056213   .328355   1.115286   1.137329
-----+-----

```



```

combined | 28057 1.195262 .002266 .3795625 1.190821 1.199704
-----+-----
diff | .0785015 .0061334 .0664773 .0905258
-----+-----
diff = mean(Alfa) - mean(Beta) t = 12.7990
Ho: diff = 0 Satterthwaite's degrees of freedom = 4810.08

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 1.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 0.0000

```

### Estadística Distancia CD-OE2

```

-----+-----
Group | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]
-----+-----
Alfa | 24645 1.099259 .0025717 .4037219 1.094218 1.1043
Beta | 3412 1.374818 .0074502 .4351855 1.360211 1.389426
-----+-----
combined | 28057 1.13277 .0024925 .4174999 1.127884 1.137655
-----+-----
diff | -.2755592 .0078816 -.2910112 -.2601072
-----+-----
diff = mean(Alfa) - mean(Beta) t = -34.9624
Ho: diff = 0 Satterthwaite's degrees of freedom = 4263.89

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 1.0000

```

### Estadística Angulo O-C-O1

```

-----+-----
Group | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]
-----+-----
Alfa | 24645 114.023 .03322 5.215116 113.9579 114.0881

```

Beta		3412	109.1378	.0804991	4.702139	108.98	109.2957
-----+-----							
combined		28057	113.4289	.0322203	5.396969	113.3658	113.4921
-----+-----							
diff			4.885139	.0870843		4.714413	5.055866
-----							
diff = mean(Alfa) - mean(Beta)						t = 56.0967	
Ho: diff = 0			Satterthwaite's degrees of freedom = 4653.04				
Ha: diff < 0		Ha: diff != 0			Ha: diff > 0		
Pr(T < t) = 1.0000		Pr( T  >  t ) = 0.0000			Pr(T > t) = 0.0000		

### Estadística Angulo O-C-CA

Group		Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
-----+-----							
Alfa		24645	113.8371	.0367973	5.776705	113.765	113.9093
Beta		3412	108.1444	.1486907	8.685367	107.8529	108.4359
-----+-----							
combined		28057	113.1448	.0386652	6.476514	113.0691	113.2206
-----+-----							
diff			5.69273	.1531762		5.392416	5.993045
-----							
diff = mean(Alfa) - mean(Beta)						t = 37.1646	
Ho: diff = 0			Satterthwaite's degrees of freedom = 3839.61				
Ha: diff < 0		Ha: diff != 0			Ha: diff > 0		
Pr(T < t) = 1.0000		Pr( T  >  t ) = 0.0000			Pr(T > t) = 0.0000		

## Estadística Angulo 01-C-CA

---

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
-----+						
Alfa	24645	114.199	.0335477	5.266556	114.1332	114.2647
Beta	3412	114.6716	.0914615	5.342481	114.4922	114.8509
-----+						
combined	28057	114.2565	.0315101	5.278012	114.1947	114.3182
-----+						
diff		-.4725895	.09742		-.663582	-.281597

---

diff = mean(Alfa) - mean(Beta) t = -4.8511

Ho: diff = 0 Satterthwaite's degrees of freedom = 4379.59

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0

Pr(T < t) = 0.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 1.0000

## Estadística Angulo C-CA-N

---

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
-----+						
Alfa	24645	113.6629	.031014	4.868804	113.6021	113.7237
Beta	3412	110.4857	.0371777	2.171638	110.4128	110.5586
-----+						
combined	28057	113.2765	.0283022	4.740687	113.2211	113.332
-----+						
diff		3.177245	.0484154		3.08234	3.27215

---

diff = mean(Alfa) - mean(Beta) t = 65.6247

Ho: diff = 0 Satterthwaite's degrees of freedom = 9194.08

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0  
Pr(T < t) = 1.0000                  Pr(|T| > |t|) = 0.0000                  Pr(T > t) = 0.0000

### Estadística Angulo CB-CA-N

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Alfa	24645	111.7394	.0348789	5.475541	111.671	111.8078
Beta	3412	113.6423	.0513649	3.000343	113.5416	113.743
combined	28057	111.9708	.031487	5.274148	111.9091	112.0325
diff		-1.90288	.0620878		-2.02459	-1.781169

diff = mean(Alfa) - mean(Beta)                      t = -30.6482  
Ho: diff = 0                      Satterthwaite's degrees of freedom = 7073.66

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0  
Pr(T < t) = 0.0000                  Pr(|T| > |t|) = 0.0000                  Pr(T > t) = 1.0000

### Estadística Angulo C-CA-CB

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Alfa	24645	113.3751	.0307656	4.829813	113.3148	113.4354
Beta	3412	120.403	.1191297	6.958645	120.1695	120.6366
combined	28057	114.2298	.0335885	5.626153	114.1639	114.2956
diff		-7.027928	.1230383		-7.269154	-6.786703

```

diff = mean(Alfa) - mean(Beta)                                t = -57.1199
Ho: diff = 0                                                  Satterthwaite's degrees of freedom = 3878.78

Ha: diff < 0                                                  Ha: diff != 0                                                Ha: diff > 0
Pr(T < t) = 0.0000      Pr(|T| > |t|) = 0.0000      Pr(T > t) = 1.0000

```

### Estadística Angulo CA-CB-CG

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645    117.3612   .0311292    4.886893    117.3002    117.4222
Beta  |     3412    120.3511   .1312203    7.664884    120.0939    120.6084
-----+-----
combined |  28057    117.7248   .0321914    5.392135    117.6617    117.7879
-----+-----
diff  |              -2.989919   .1348622                -3.254328   -2.725509
-----

```

```

diff = mean(Alfa) - mean(Beta)                                t = -22.1702
Ho: diff = 0                                                  Satterthwaite's degrees of freedom = 3804.06

Ha: diff < 0                                                  Ha: diff != 0                                                Ha: diff > 0
Pr(T < t) = 0.0000      Pr(|T| > |t|) = 0.0000      Pr(T > t) = 1.0000

```

### Estadística Angulo CB-CG-CD

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645    116.3626   .0351898    5.524353    116.2936    116.4316
Beta  |     3412    118.2347   .3025221    17.67102    117.6416    118.8279
-----+-----
combined |  28057    116.5902   .0481863    8.07131    116.4958    116.6847
-----+-----

```

```

diff |          -1.872149   .3045619          -2.469285   -1.275012
-----
diff = mean(Alfa) - mean(Beta)          t =  -6.1470
Ho: diff = 0          Satterthwaite's degrees of freedom = 3503.84

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0000    Pr(|T| > |t|) = 0.0000    Pr(T > t) = 1.0000

```

### Estadística Angulo CB-CD-OE1

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa |   24645   119.4044   .0600809    9.431939    119.2866    119.5222
Beta |    3412   122.0195   .1979755   11.56421    121.6313    122.4077
-----+-----
combined | 28057   119.7224   .0582294    9.753555    119.6083    119.8366
-----+-----
diff |          -2.615082   .2068913          -3.020703   -2.209462
-----
diff = mean(Alfa) - mean(Beta)          t = -12.6399
Ho: diff = 0          Satterthwaite's degrees of freedom = 4063.45

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0000    Pr(|T| > |t|) = 0.0000    Pr(T > t) = 1.0000

```

### Estadística Angulo CG-CO-OE2

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa |   24645   121.382    .040873    6.416543    121.3019    121.4621
Beta |    3412   114.5455   .0701508    4.097669    114.408    114.6831
-----+-----

```

```

combined | 28057 120.5506 .0392388 6.572595 120.4737 120.6275
-----+-----
diff | 6.836455 .0811895 6.677294 6.995615
-----+-----
diff = mean(Alfa) - mean(Beta) t = 84.2037
Ho: diff = 0 Satterthwaite's degrees of freedom = 6023.91

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 1.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 0.0000

```

### Estadística Angulo OE2-CD-OE1

```

-----+-----
Group | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]
-----+-----
Alfa | 24645 119.6782 .0601853 9.448333 119.5602 119.7961
Beta | 3412 113.4753 .044289 2.587023 113.3885 113.5621
-----+-----
combined | 28057 118.9238 .0545006 9.128977 118.817 119.0307
-----+-----
diff | 6.202873 .0747248 6.056405 6.34934
-----+-----
diff = mean(Alfa) - mean(Beta) t = 83.0096
Ho: diff = 0 Satterthwaite's degrees of freedom = 18777.9

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 1.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 0.0000

```

### Estadística Dihedro O-C-CA-CB

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645    134.3636   .1516523    23.80748    134.0664    134.6609
Beta  |     3412    120.8609   .167586    9.789088    120.5323    121.1895
-----+-----
combined |    28057    132.7216   .1373107    22.99985    132.4524    132.9907
-----+-----
diff  |              13.50271   .2260166              13.05968    13.94575
-----

diff = mean(Alfa) - mean(Beta)                t = 59.7421
Ho: diff = 0                Satterthwaite's degrees of freedom = 10326.3

Ha: diff < 0                Ha: diff != 0                Ha: diff > 0
Pr(T < t) = 1.0000        Pr(|T| > |t|) = 0.0000        Pr(T > t) = 0.0000

```

### Estadística Dihedro O1-C-CA-N

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645    132.1481   .1401121    21.99581    131.8735    132.4227
Beta  |     3412    158.7356   .5617988    32.81598    157.6341    159.8371
-----+-----
combined |    28057    135.3814   .150016    25.12801    135.0874    135.6754
-----+-----
diff  |             -26.58749   .5790072             -27.72268    -25.4523
-----

diff = mean(Alfa) - mean(Beta)                t = -45.9191
Ho: diff = 0                Satterthwaite's degrees of freedom = 3846.46

```



Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0  
Pr(T < t) = 0.0000                  Pr(|T| > |t|) = 0.0000                  Pr(T > t) = 1.0000

### Estadística Dihedro CA-CB-CG-CD

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Alfa	24645	135.675	.1684072	26.43779	135.3449	136.0051
Beta	3412	140.4003	.4569263	26.69013	139.5044	141.2962
combined	28057	136.2497	.1582853	26.51315	135.9394	136.5599
diff		-4.725263	.4869729		-5.679975	-3.77055

diff = mean(Alfa) - mean(Beta)                      t = -9.7033  
Ho: diff = 0                      Satterthwaite's degrees of freedom = 4389.43

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0  
Pr(T < t) = 0.0000                  Pr(|T| > |t|) = 0.0000                  Pr(T > t) = 1.0000

### Estadística Dihedro C-CA-CB-N

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Alfa	24645	139.4419	.1553926	24.39466	139.1373	139.7465
Beta	3412	175.8725	.3983319	23.26749	175.0915	176.6535
combined	28057	143.8722	.161338	27.02448	143.556	144.1885
diff		-36.43056	.4275689		-37.26881	-35.59232

```

-----
diff = mean(Alfa) - mean(Beta)                                t = -85.2040
Ho: diff = 0                                                  Satterthwaite's degrees of freedom = 4513.73

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0000    Pr(|T| > |t|) = 0.0000    Pr(T > t) = 1.0000

```

### Estadística Dihedro 01-C-CA-CB

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |   24645   158.7113   .1752629   27.51404   158.3678   159.0548
Beta  |    3412   138.3145   .609641   35.61055   137.1193   139.5098
-----+-----
combined |  28057   156.2309   .1754403   29.38664   155.887   156.5747
-----+-----
diff  |           20.39675   .6343337           19.1531   21.6404
-----

```

```

-----
diff = mean(Alfa) - mean(Beta)                                t = 32.1546
Ho: diff = 0                                                  Satterthwaite's degrees of freedom = 3994.35

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 1.0000    Pr(|T| > |t|) = 0.0000    Pr(T > t) = 0.0000

```

### Estadística Dihedro CB-CG-CD-OE1

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645    148.2253   .2043477    32.07999    147.8248    148.6259
Beta  |     3412    111.3372   .2447224    14.2948    110.8574    111.817
-----+-----
combined |    28057    143.7394   .1956665    32.77458    143.3559    144.1229
-----+-----
diff  |              36.88814   .3188213              36.26318    37.51311
-----

diff = mean(Alfa) - mean(Beta)                                t = 115.7016
Ho: diff = 0                                                    Satterthwaite's degrees of freedom = 9206.48

Ha: diff < 0                Ha: diff != 0                Ha: diff > 0
Pr(T < t) = 1.0000          Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000

```

### Estadística Dihedro CB-CG-CD-OE1

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645    116.7326   .1609007    25.25937    116.4172    117.048
Beta  |     3412    159.4599   .4968665    29.02313    158.4857    160.4341
-----+-----
combined |    28057    121.9286   .1748602    29.28948    121.5859    122.2714
-----+-----
diff  |              -42.72733   .5222694              -43.75125    -41.7034
-----

diff = mean(Alfa) - mean(Beta)                                t = -81.8109
Ho: diff = 0                                                    Satterthwaite's degrees of freedom = 4157.58

Ha: diff < 0                Ha: diff != 0                Ha: diff > 0

```

Pr(T < t) = 0.0000      Pr(|T| > |t|) = 0.0000      Pr(T > t) = 1.0000

### Estadística Dihedro O-C-CA-N

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Alfa	24645	120.6241	.0577852	9.071543	120.5108	120.7373
Beta	3412	177.8083	.2434436	14.22011	177.331	178.2856
combined	28057	127.5782	.126107	21.1232	127.331	127.8254
diff		-57.18425	.2502078		-57.6748	-56.69369

diff = mean(Alfa) - mean(Beta)      t = -2.3e+02

Ho: diff = 0      Satterthwaite's degrees of freedom = 3804.53

Ha: diff < 0      Ha: diff != 0      Ha: diff > 0

Pr(T < t) = 0.0000      Pr(|T| > |t|) = 0.0000      Pr(T > t) = 1.0000

### Estadística Dihedro phi C-N-CA-C

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Alfa	24645	133.2178	.1549811	24.33006	132.914	133.5216
Beta	3412	141.2358	.5224051	30.5149	140.2116	142.2601
combined	28057	134.1929	.1510367	25.29898	133.8968	134.4889
diff		-8.018051	.5449094		-9.086374	-6.949728

diff = mean(Alfa) - mean(Beta)      t = -14.7145

Ho: diff = 0      Satterthwaite's degrees of freedom = 4033.52

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0  
 Pr(T < t) = 0.0000                  Pr(|T| > |t|) = 0.0000                  Pr(T > t) = 1.0000

### Estadística Dihedro psi N-CA-C-N

```

-----
Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
Alfa  |    24645    138.7576   .1741393    27.33765    138.4162    139.0989
Beta  |     3412    133.1231   .4731399    27.6372    139.1954    142.0507
-----+-----
combined |   28057    138.0724   .1637927    27.43564    137.7513    138.3934
-----+-----
diff  |              5.634506   .5041685              4.646081    6.622931
-----
  
```

diff = mean(Alfa) - mean(Beta)    t = 11.1758  
 Ho: diff = 0    Satterthwaite's degrees of freedom = 4386.57

Ha: diff < 0                      Ha: diff != 0                      Ha: diff > 0  
 Pr(T < t) = 1.0000                  Pr(|T| > |t|) = 0.0000                  Pr(T > t) = 0.0000