



**UNIVERSIDADE DA BEIRA INTERIOR**  
**Ciências**

**PREVISÃO POR MÉTODOS INFORMÁTICOS DA ESTRUTURA  
SECUNDÁRIA DE PROTEÍNAS E DE ALGUMAS  
PROPRIEDADES ASSOCIADAS**

---

**DISSERTAÇÃO DE MESTRADO EM BIOQUÍMICA  
(MASTER DISSERTATION OF BIOCHEMISTRY)**

**Nuno José Costa Carvalho**

Co-orientador - Professor António Mendonça  
Departamento de Química

**COVILHÃ, 2010**



2010

Bioinformática, o passo seguinte



Todo o conteúdo desta tese é da exclusiva  
responsabilidade do seu autor

(Nuno José da Costa Carvalho)

## AGRADECIMENTOS

Seria falso pensar-se que ao longo da nossa carreira, a caminhada é feita de maneira solitária como se fosse uma travessia do deserto, na qual estaríamos única e exclusivamente dependentes de nós próprios. Na realidade, cada decisão ou rumo que tomamos, tem uma influência directa ou indirecta de uma pessoa ou de um conjunto de pessoas que nos rodeiam. O fenómeno está de tal ordem bem feito, que na maior parte das vezes não nos damos conta da imensa ajuda que o familiar, amigo ou até mesmo um docente nos deu, e consequentemente esquecemo-nos de lhes agradecer.

A decisão de tirar o mestrado foi longamente ponderada. As dúvidas foram muitas e as dificuldades esperadas, prefiro nem pensar nelas. Mas eu só pude decidir avançar com tal empreendimento, sabendo que tinha familiares, amigos e docentes competentes a auxiliar todos os meus passos. Por isso não, não posso esquecer todas estas pessoas que tanto fizeram por mim. Só receio que a minha falta de destreza no malabarismo das palavras, não ilustre ao seu justo valor todo o apoio que elas me deram.

A ordem pela qual agradeço, aos meus olhos não tem grande importância. Prefiro enfatizar o apoio de cada um, porque como todas as pessoas são diferentes, cada uma delas auxiliou-me à sua maneira.

Agradeço a toda a minha família pelo apoio moral, em particular os meus pais, que tiveram de aturar as minhas angústias e preocupações ao longo deste ano. Agradeço igualmente à professora Cândida Tomaz que teve certamente alguns suores frios quando lhe fiz parte da minha vontade em acrescentar mais um “tijolo” ao meu currículo académico. Agradeço o profissionalismo exemplar do professor Mendonça, que se armou de paciência e apoiou-me em tudo. Devo dizer que graças à sua dedicação, tomei gosto pelo projecto que *a priori* não me parecia muito aliciante. Um grande bem-haja à minha fiel amiga Catarina Nunes por ter disponibilizado alguns dos seus preciosos dias para me ajudar a transcrever para a dissertação os apontamentos que eu tinha feito ante e pelo apoio sempre construtivo. O mesmo digo às minhas amigas Vera Cortez e Madalena Silva que também elas disponibilizaram um fim-de-semana para me ajudar. Não me posso esquecer de agradecer com a mesma árdua, as minhas amigas Cheila Neves; Laura Canelo e Paula Miranda.

Todas estas pessoas e outras que provavelmente não me recordo e pelas quais peço desculpa, formaram o elenco de um palco importante, mais um palco da minha vida. Com este projecto, não só enriqueci o meu conhecimento científico como também aprendi que os verdadeiros amigos existem e sacrificam-se se tal for necessário para nos ajudar. A todos eles, muito obrigado.

## RESUMO

A bioinformática é uma disciplina inovadora que conheceu um grande desenvolvimento nestes últimos anos. Ela não só permite uma real economia para a indústria farmacêutica como também evita parte do sacrifício das espécies de animais habitualmente usadas nos laboratórios de investigação científica.

Com o intuito de prever a estrutura secundária (hélices- $\alpha$ ; folhas- $\beta$  e voltas) das proteínas a partir de escalas de hidrofobicidade dos aminoácidos, recorreu-se a ferramentas informáticas, (Folhas de cálculo). Mais concretamente, utilizou-se vários métodos de cálculo como a escala de hidrofobicidade de Kyte & Doolittle entre outras; método de Chou & Fasman e método de Lupas. Também foi estudada a presença de zonas de maior alergenicidade nas proteínas. Estes estudos foram efectuados com as seguintes proteínas: lisozima C; ovalbumina; ovomucina; ovomucóide;  $\alpha$ -caseína s1;  $\alpha$ -caseína s2;  $\beta$ -caseína;  $\beta$ -lactoglobulina;  $\beta$ -2-microglobulina; osteopontina e ubiquitina. Os cálculos foram também efectuados para a BSA.

A interpretação dos resultados obtidos não foi tão fácil quanto o esperado. Se o método de Chou & Fasman e o método de Lupas não deixam margem de dúvida quanto à possível existência de estruturas como: as hélices- $\alpha$ ; folhas- $\beta$  e voltas (*turns*) ou ainda a super-hélices- $\alpha$  (*coiled-coil*), não se pode esquecer que se trata de métodos baseados em cálculos de probabilidade. Por essa razão, não permitem dizer com toda a certeza qual a estrutura secundária das proteínas em estudo. Os resultados gráficos obtidos através dos cálculos baseados nas escalas de hidrofobicidade, revelaram uma falta de clareza e evidência sendo mesmo por vezes sujeitos a interpretações subjectivas. No estudo da alergenicidade, também ele baseado numa escala de hidrofobicidade (escala de Hopp & Wood), todos os picos obtidos podem ser interpretados como potenciais zonas de ligação ou seja, potenciais epítomos visto que têm valores positivos.

De um ponto de vista geral, pode dizer-se que todos os métodos utilizados, encontraram aplicação em todas as proteínas estudadas. No entanto, é preciso ter-se em conta que todos os resultados são apresentados como se as proteínas estivessem totalmente desenroladas, o que não é conforme à realidade.

Palavras-chave: escalas de hidrofobicidade; estrutura secundária; epítomos; alergenicidade

## ABSTRAT

Bioinformatics is an innovative discipline that has been significantly developed in the past few years. It not only allows a real economy for the pharmaceutical industry, but also avoids the sacrifice of the animal species usually used in scientific investigation laboratories.

Informatic tools were used with the objective of predicting the secondary structure ( $\alpha$ -helices,  $\beta$ -sheets and turns) of proteins from aminoacids hydrophobicity scales. Specifically, several calculation methods were used like the Kyte & Doolittle hydrophobicity scale, among other scales, Chou & Fasman method and Lupas method. The presence of larger allergenicity areas in proteins was also studied. These studies were performed using the following proteins: lysozyme C, ovalbumin, ovomucin, ovomucoid,  $\alpha$ -casein s1,  $\alpha$ -casein s2,  $\beta$ -casein,  $\beta$ -lactoglobulin,  $\beta$ -2-microglobulin, osteopontin and ubiquitin. Calculations were also preformed with BSA.

The interpretation of the obtained results was not as easy as expected. If Chou & Fasman and Lupas methods did not raise any doubts to the possible existence of structures like  $\alpha$ -helices,  $\beta$ -sheets and turns or even coiled-coil  $\alpha$ -helices, it cannot be disregard that these are probability based methods. For this reason it is not possible to predict with absolute certainty the secondary structure of the studied proteins. The graphic results obtained by calculations based on the hydrophobicity scales reveal a lack of clarity and evidence, and were sometimes object of incorrect interpretations. In the allergenicity study, also based in the Hopp & Wood hydrophobicity scale, all the obtained peaks could be interpreted as potential binding areas, or epitopes, since they have positive values.

In a broad point of view, it can be concluded that all the used methods find application in the studied proteins. However, the study was made considering that the proteins are totally unwind, which does not correspond to reality.

Keywords: hydrophobicity scales, secondary structure, epitopes, allergenicity.

# ÍNDICE GERAL

## 1 Conteúdo

AGRADECIMENTOS.....	I
RESUMO .....	II
ABSTRAT .....	III
ÍNDICE GERAL .....	IV
LEGENDAS DAS ILUSTRAÇÕES, TABELAS E GRÁFICOS .....	VI
LISTA DE ABREVIATURAS .....	XVIII
<b>1. INTRODUÇÃO TEÓRICA .....</b>	<b>- 1 -</b>
1.1. BIOINFORMÁTICA .....	- 1 -
1.2. OS AMINOÁCIDOS.....	- 1 -
<b>1.5.3. Estrutura geral</b> .....	- 1 -
<b>1.5.3. Propriedades iónicas</b> .....	- 2 -
<b>1.5.3. Solubilidade dos aminoácidos</b> .....	- 3 -
1.3. AS PROTEÍNAS .....	- 5 -
<b>1.5.3. Estrutura primária</b> .....	- 6 -
<b>1.5.3. Estrutura secundária</b> .....	- 7 -
<b>1.5.3. Estrutura terciária</b> .....	- 10 -
<b>1.3.1. Estrutura quaternária</b> .....	- 11 -
<b>1.3.5. Localização dos aminoácidos na estrutura tridimensional</b> .....	- 11 -
<b>1.3.6. Solubilidade das proteínas</b> .....	- 13 -
1.4. PROTEÍNAS EM ESTUDO .....	- 14 -
<b>1.4.7. BSA</b> .....	- 14 -
<b>1.4.2. Lisozima C</b> .....	- 16 -
<b>1.4.3. Ovalbumina</b> .....	- 16 -
<b>1.4.4. Ovomucina</b> .....	- 17 -
<b>1.4.5. Ovomucóide</b> .....	- 17 -
<b>1.4.6. Caseínas</b> .....	- 18 -
<b>1.4.7. B-lactoglobulina</b> .....	- 20 -
<b>1.4.8. B-2-microglobulina</b> .....	- 20 -
<b>1.4.9. Osteopontina</b> .....	- 20 -
<b>1.4.10. Ubiquitina</b> .....	- 21 -
1.5. REACÇÕES ALÉRGICAS .....	- 22 -
<b>1.5.1. Epítomos</b> .....	- 22 -
<b>1.5.2. Imunoglobulinas (Igs)</b> .....	- 23 -
<b>1.5.3. Mecanismos celulares e moleculares da Alergia Alimentar</b> .....	- 23 -
1.6. MÉTODOS DE ESTUDO.....	- 25 -
<b>1.6.1. As escalas de hidrofobicidade</b> .....	- 26 -
<b>1.6.2. Composição da proteína</b> .....	- 29 -

1.6.3. <i>Determinação das cargas</i> .....	- 32 -
1.6.4. <i>Método de Chou &amp; Fasman</i> .....	- 34 -
1.6.5. <i>Estruturas super-enroladas “Coiled-coil”</i> .....	- 37 -
2. <b>OBJECTIVO</b> .....	- 43 -
3. <b>MATERIAIS E MÉTODOS</b> .....	- 44 -
3.1. <b>MATERIAIS</b> .....	- 44 -
3.1.1. <i>Material informático</i> .....	- 44 -
3.1.2. <i>Sequências das proteínas</i> .....	- 44 -
3.2. <b>PROTOCOLO</b> .....	- 45 -
4. <b>REGISTO E DISCUSSÃO DOS RESULTADOS</b> .....	- 47 -
4.1. <b>HIDROFOBICIDADE E ALERGENICIDADE</b> .....	- 47 -
4.2. <b>ESCOLHA DA ESCALAS DE HIDROFOBICIDADE</b> .....	- 49 -
4.2.1. <i>BSA</i> .....	- 53 -
4.2.2. <i>Proteínas do ovo</i> .....	- 61 -
4.2.3. <i>Proteínas do leite</i> .....	- 102 -
4.2.4. <i>Proteínas do organismo</i> .....	- 128 -
5. <b>CONCLUSÃO</b> .....	- 146 -
6. <b>FUTURAS INVESTIGAÇÕES</b> .....	- 148 -
7. <b>BIBLIOGRAFIA</b> .....	- 149 -
8. <b>CIBERGRAFIA</b> .....	- 153 -
9. <b>ANEXO</b> .....	- 154 -
<i>Proteína de referência</i> .....	- 154 -
<i>Proteínas do ovo</i> .....	- 156 -
<i>Proteínas do leite</i> .....	- 164 -
<i>Proteínas do organismo</i> .....	- 172 -

# LEGENDAS DAS ILUSTRAÇÕES, TABELAS E GRÁFICOS

## Ilustrações

Figura 1	Estrutura elementar dos aminoácidos	1
Figura 2	Equilíbrio protónico do ácido aspártico. Adaptado de Murray <i>et al.</i> , 2002.	2
Figura 3	Estrutura primária das proteínas, sequência de aminoácidos. (Adaptado de Goldsby <i>et al.</i> , 2002)	6
Figura 4	Representação das ligações peptídicas. Num par de aminoácidos ligados, seis átomos ( $C_{\alpha}$ , C, O, N e $C_{\alpha}$ ), estão no mesmo plano. As cadeias laterais, representadas a verde, situam-se noutra plano.	7
Figura 5	Estrutura secundária, representação de hélice- $\alpha$ e folhas- $\beta$	7
Figura 6	Estrutura terciária, configuração das proteínas constituídas pelas diferentes estruturas: hélices- $\alpha$ ; folhas- $\beta$ e voltas.	10
Figura 7	Estrutura quaternária. As proteínas exibem quatro níveis de organização estrutural, (subunidades).	11
Figura 8	Solubilidade da $\beta$ -lactoglobulina em função do pH, a concentrações de NaCl diferentes.	13
Figura 9	Estrutura tridimensional da albumina bovina serosa. Representação dos três domínios e seis subdomínios.	14
Figura 10	Localização das ligações de dissulfureto.	15
Figura 11	Representação da albumina serosa com os resíduos básicos coloridos a azul, os resíduos ácidos coloridos a vermelho e os resíduos neutros coloridos em amarelo.	15
Figura 12	Estrutura cristalina da lisozima C da clara de ovo de galinha	16
Figura 13	Representação da $\beta$ -lactoglobulina (Nativa).	20
Figura 14	Estrutura geral da IgE. Representação das cadeias leves (rosa) e das cadeias pesadas (amarelo). As pontes dissulfuretos estão representadas por linhas pretas finas.	24
Figura 15	Mecanismo molecular e celular da alergia alimentar. As letras B e T representam respectivamente os linfócitos B e T. Th1 e Th2 representam respectivamente as células T helper 1 e 2. As diferentes interleucinas estão representadas através de (IL-4 e IL-13). IFN- $\gamma$ representa o interferão gama (Interferon- $\gamma$ ). IgE e IgM representam respectivamente as imunoglobulinas E e M. A letra (M) representa uma célula ao acaso da membrana gastrointestinal.	25
Figura 16	Análise do receptor TRH do rato para o $H_7$ , o $H_{\alpha}$ e o $H_{Ag}$ .	30

## Tabelas

Tabela 1	Valores do pK dos principais 20 aminoácidos para o ácido carboxílico, o grupo amina e o grupo R. assim como para o C-terminal e N-terminal.	3
Tabela 2	Classificação dos aminoácidos L das proteínas, baseada na hidrofobicidade relativa e na hidrofiliicidade relativa.	4
Tabela 3	Valores da solubilidade dos principais 20 aminoácidos.	5
Tabela 4	Os aminoácidos estão agrupados de acordo com a sua preferência (a negrito) em formar hélices- $\alpha$ , folhas- $\beta$ e voltas. A arginina não mostra preferência significativa por qualquer uma das estruturas.	12
Tabela 5	Representação de várias escalas de hidrofobicidade.	28
Tabela 6	Representação da tabela de cálculos da hidrofobicidade para $H_7$ ; $H_\alpha$ e $H_{Ag}$ , de uma amostra dos 20 primeiros aminoácidos da BSA. HYKD (escala hidrofóbica de <i>Kyte &amp; Doolittle</i> ); HYHW (escala hidrofóbica de <i>Hopp &amp; Wood</i> ).	29
Tabela 7	Representação da composição real e relativa da BSA em aminoácidos, assim como a percentagem de cada um deles. Nesta tabela também está representada a média real e relativa dos aminoácidos das proteínas em geral.	32
Tabela 8	Representação dos cálculos de uma amostra de 28 aminoácidos da $\alpha$ -caseína s1.	34
Tabela 9	Matriz de probabilidades através da qual se aplica o algoritmo de Chou & Fasman para a elaboração da tabela 9. Esta matriz é idêntica para todas as proteínas.	35
Tabela 10	Representação da folha de cálculos a partir da qual se constrói o gráfico de Chou & Fasman para as diferentes estruturas: hélice- $\alpha$ , folha- $\beta$ e volta. Amostra dos primeiros 9 aminoácido da BSA.	37
Tabela 11	Resumo das diferentes funções que permitem calcular os valores das diferentes colunas da tabela 9.	37
Tabela 12	Frequência com a qual, cada aminoácido é encontrado em cada uma das sete posições possíveis em proteínas super-enroladas conhecidas.	39
Tabela 13	Representação da folha de cálculos a partir da qual se calcula os valores das 7 referências relativas às 7 posições (a, b, c, d, e, f e g) e para cada aminoácido da cadeia polipeptídica. Amostra dos primeiros 31 aminoácidos.	40
Tabela 14	Representação da folha de cálculos a partir da qual se calcula os valores que servirão à elaboração do gráfico das estruturas super-enroladas "Coiled coil". Amostra dos primeiros 30 aminoácidos da BSA.	41
Tabela 15	Resumo das diferentes funções que permitem calcular os valores das diferentes colunas da tabela 12 e 13.	43
Tabela 16	Características do equipamento e softwares utilizados.	45
Tabela 17	Proteínas em estudo, da clara do ovo, do leite, do organismo humano e BSA.	46
Tabela 18	Diferentes escalas de hidrofobicidade.	49

## Gráficos

Gráfico 1	Representação gráfica das curvas de hidrofobicidade $H_7$ para a escala de Kyte &	50
-----------	---	----

	Doolittle (preto) e a escala de Black (roxo) na BSA (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).	
Gráfico 2	Representação gráfica das curvas de índice de hélice- $\alpha$ (Ha) para a escala de Kyte & Doolittle (preto) e a escala de Black (roxo) na BSA (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	50
Gráfico 3	Representação gráfica das curvas de hidrofobicidade $H_7$ para a escala de Sweet <i>et al.</i> (verde) e a escala de Eisenberg <i>et al.</i> (azul) na BSA (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).	51
Gráfico 4	Representação gráfica das curvas de hidrofobicidade $H_7$ para as escalas de Kyte & Doolittle (preto), de Sweet <i>et al.</i> (verde) e de Eisenberg <i>et al.</i> (azul) na BSA (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).	51
Gráfico 5	Representação gráfica das curvas de índice de hélice- $\alpha$ (Ha) para as escalas de Kyte & Doolittle (preto), de Sweet <i>et al.</i> (verde) e de Eisenberg <i>et al.</i> (azul) na BSA (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	52
Gráfico 6	Representação gráfica das curvas de hidrofobicidade $H_7$ para as escalas de Welling <i>et al.</i> (castanho), de Sweet <i>et al.</i> (verde) e de Eisenberg <i>et al.</i> (azul) na BSA, (numa amostra de 150 aminoácidos). Linha de referência de 1,6 (vermelho).	52
Gráfico 7	Representação gráfica das curvas de hidrofobicidade $H_7$ para as escalas de Kyte & Doolittle (preto) e de Welling <i>et al.</i> (castanho) na BSA (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).	53
Gráfico 8	Representação gráfica das curvas de hidrofobicidade $H_7$ para as escalas de Kyte & Doolittle (preto), de Sweet <i>et al.</i> (verde), de Welling <i>et al.</i> (castanho), de Eisenberg <i>et al.</i> (azul) e de Black (roxo), na BSA (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).	53
Gráfico 9	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a BSA (607 resíduos de aminoácidos) usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho). (a) resíduos 1-201; (b) resíduos 203-404 e (c) resíduos 405-605.	54
Gráfico 10	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a BSA, usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro). (a) resíduos 1-202; (b) resíduos 203-404 e (c) resíduos 405-607.	55
Gráfico 11	Representação gráfica das zonas antigénicas, para a BSA, usando a escala de Hopp & Wood (azul). (a) resíduos 1-202; (b) resíduos 203-404 e (c) resíduos 405-607. Zonas antigénicas possíveis (setas azuis).	56
Gráfico 12	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na BSA. Cada pico corresponde à presença de um possível elemento. (a) resíduos 3-202; (b) resíduos 203-404 e (c) resíduos 405-603.	58
Gráfico 13	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na BSA. A cada pico corresponder a possível presença dessa estrutura.	59
Gráfico 14	Representação gráfica da quantidade real de cada aminoácido na BSA (607 resíduos de aminoácido).	60
Gráfico 15	Representação gráfica das cargas da BSA. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos. (a) resíduos 1-202; (b) resíduos 203-404 e (c) resíduos 405-580.	61
Gráfico 16	Representação gráfica do cálculo da hidrofobicidade $H_7$ com base na escala de Kyte & Doolittle (preto) para a lisozima C (147 resíduos de aminoácido). Linha de referência de 1,6 (vermelho).	62
Gráfico 17	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a lisozima C, usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	63
Gráfico 18	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a lisozima C, usando	64

	a escala de Hopp & Wood (azul). Zonas antigénicas possíveis (setas azuis).	
Gráfico 19	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na lisozima C. Cada pico corresponde à presença de um possível elemento.	64
Gráfico 20	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na lisozima C.	65
Gráfico 21	Representação gráfica da quantidade real de cada aminoácido na lisozima C (147 resíduos de aminoácido).	65
Gráfico 22	Representação gráfica da percentagem de cada aminoácido na lisozima C, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína (147 resíduos de aminoácido).	66
Gráfico 23	Representação gráfica da percentagem relativa dos aminoácidos da lisozima C, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	67
Gráfico 24	Representação gráfica das cargas da lisozima C. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	68
Gráfico 25	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a ovalbumina (386 resíduos de aminoácido) usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho). (a) resíduos 1-193 e (b) resíduos 194-386.	69
Gráfico 26	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a ovalbumina, usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro). (a) resíduos 1-193 e (b) resíduos 194-386.	70
Gráfico 27	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a ovalbumina, usando a escala de Hopp & Wood (azul). (a) resíduos 1-193 e (b) resíduos 194-386. Zonas antigénicas possíveis (setas azuis).	71
Gráfico 28	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na ovalbumina. Cada pico corresponde à presença de um possível elemento estrutural. (a) resíduos 1-193 e (b) resíduos 194-383.	72
Gráfico 29	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na ovalbumina. A cada pico corresponder a possível presença dessa estrutura.	73
Gráfico 30	Representação gráfica da quantidade real de cada aminoácido na ovalbumina (386 resíduos de aminoácido).	73
Gráfico 31	Representação gráfica da percentagem de cada aminoácido na ovalbumina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína (386 resíduos de aminoácido).	74
Gráfico 32	Representação gráfica da percentagem relativa dos aminoácidos da ovalbumina, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão à direita e os mais hidrofóbicos estão à esquerda. O círculo a tracejado vermelho representa a média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	75
Gráfico 33	Representação gráfica das cargas da ovalbumina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos. (a) resíduos 1-193 e (b) resíduos 194-359.	75
Gráfico 34	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a subunidade $\alpha$ da ovomucina (2108 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho). (a) resíduos 1-264; (b) resíduos 265-528; (c) resíduos 529-792; (d) resíduos 793-1056; (e) resíduos 1057-1320; (f) resíduos 1321-1584; (g) resíduos 1585-1848 e (h) resíduos 1849-2108.	76
Gráfico 35	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a ovomucina (2108	79

	resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro). (a) resíduos 1-264; (b) resíduos 265-528; (c) resíduos 529-792; (d) resíduos 793-1056; (e) resíduos 1057-1320; (f) resíduos 1321-1584; (g) resíduos 1585-1848 e (h) resíduos 1849-2108.	
Gráfico 36	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a ovomucina, usando a escala de Hopp & Wood (azul). (a) resíduos 1-264; (b) resíduos 265-528; (c) resíduos 529-792; (d) resíduos 793-1056; (e) resíduos 1057-1320; (f) resíduos 1321-1584; (g) resíduos 1585-1848 e (h) resíduos 1849-2108. Zonas antigénicas possíveis (setas azuis).	81
Gráfico 37	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na ovomucina. Cada pico corresponde à presença de um possível elemento. (a) resíduos 1-269; (b) resíduos 270-534; (c) resíduos 535-799; (d) resíduos 800-1064; (e) resíduos 1065-1329; (f) resíduos 1330-1595; (g) resíduos 1596-1859 e (h) resíduos 1860-2104.	84
Gráfico 38	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na ovomucina.	86
Gráfico 39	Representação gráfica da quantidade real de cada aminoácido na ovomucina (2108 resíduos de aminoácido).	87
Gráfico 40	Representação gráfica da percentagem de cada aminoácido na ovomucina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína (2108 resíduos de aminoácido).	88
Gráfico 41	Representação gráfica da percentagem relativa dos aminoácidos da ovomucina, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão situados à direita e os mais hidrofóbicos estão situados à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	89
Gráfico 42	Representação gráfica das cargas da ovomucina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos. (a) resíduos 1-264; (b) resíduos 265-528; (c) resíduos 529-792; (d) resíduos 793-1056; (e) resíduos 1057-1320; (f) resíduos 1321-1584; (g) resíduos 1585-1848; (h) resíduos 1849-2081.	90
Gráfico 43	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a ovomucóide (210 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).	93
Gráfico 44	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a ovomucóide (210 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	93
Gráfico 45	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a ovomucóide, usando a escala de Hopp & Wood (azul). Zonas antigénicas possíveis (setas azuis).	94
Gráfico 46	Representação gráfica dos cálculos da hidrofobicidade $H_7$ para a ovomucóide, usando a escala de Kyte & Doolittle para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Esta mutação consiste em substituir a glicina 32 (G32) e a fenilalanina 37 (F37) pela metionina. Linha de referência de 1,6 (vermelho).	95
Gráfico 47	Representação gráfica pormenorizada da hidrofobicidade $H_7$ para a ovomucóide, usando a escala de Kyte & Doolittle para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Janela de apenas 56 resíduos de aminoácido (Terceiro domínio completo). Linha de referência de 1,6 (vermelho).	95
Gráfico 48	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a ovomucóide, usando a escala de Kyte & Doolittle para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Esta mutação consiste em substituir a glicina 32 (G32) e a fenilalanina 37 (F37) pela metionina. (210 resíduos de aminoácido). Linha de referência de 0,8 (verde claro).	96

Gráfico 49	Representação gráfica pormenorizada do índice de hélice- $\alpha$ para a ovomucóide, usando a escala de Kyte & Doolittle para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Janela de apenas 56 resíduos de aminoácido (Terceiro domínio completo). Linha de referência de 0,8 (verde claro).	96
Gráfico 50	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a ovomucóide, usando a escala de Hopp & Wood para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Esta mutação consiste em substituir a glicina 32 (G32) e a fenilalanina 37 (F37) pela metionina.	97
Gráfico 51	Representação gráfica pormenorizada da antigenicidade $H_{Ag}$ para a ovomucóide, usando a escala de Hopp & Wood para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Janela de apenas 56 resíduos de aminoácido (Terceiro domínio completo). Zonas de divergência (setas azuis).	97
Gráfico 52	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na ovomucóide. Cada pico corresponde à presença de um possível elemento estrutural.	98
Gráfico 53	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na ovomucóide.	99
Gráfico 54	Representação gráfica da quantidade real de cada aminoácido na ovomucóide (210 resíduos de aminoácido).	99
Gráfico 55	Representação gráfica da percentagem de cada aminoácido na ovomucóide, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 19 aminoácidos presentes na proteína.	100
Gráfico 56	Representação gráfica da percentagem relativa dos aminoácidos da ovomucóide, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 19 aminoácidos presentes na proteína.	101
Gráfico 57	Representação gráfica das cargas da ovomucóide. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	102
Gráfico 58	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a $\alpha$ -caseína s1 (214 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).	103
Gráfico 59	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a $\alpha$ -caseína s1 (214 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	104
Gráfico 60	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a $\alpha$ -caseína s1, usando a escala de Hopp & Wood (azul). Epítomos imunodominantes (setas verdes), ligações à IgE (setas vermelhas) e ligação antigénica provável (seta azul).	104
Gráfico 61	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na $\alpha$ -caseína s1. Cada pico corresponde à presença de um possível elemento.	105
Gráfico 62	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na $\alpha$ -caseína s1. A cada pico corresponder a possível presença dessa estrutura.	105
Gráfico 63	Representação gráfica da quantidade real de cada aminoácido na $\alpha$ -caseína s1 (214 resíduos de aminoácido).	106
Gráfico 64	Representação gráfica da percentagem de cada aminoácido na $\alpha$ -caseína s1, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes	107

	na proteína.	
Gráfico 65	Representação gráfica da percentagem relativa dos aminoácidos da $\alpha$ -caseína s1, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão situados à direita e os mais hidrofóbicos estão situados à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	108
Gráfico 66	Representação gráfica das cargas da $\alpha$ -caseína s1. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	109
Gráfico 67	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a $\alpha$ -caseína s2 (222 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).	109
Gráfico 68	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a $\alpha$ -caseína s2 (222 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	110
Gráfico 69	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a $\alpha$ -caseína s2, usando a escala de Hopp & Wood (azul). Ligações fracas à IgE (setas azuis) e ligações fortes à IgE (setas vermelhas).	111
Gráfico 70	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na $\alpha$ -caseína s2. Cada pico corresponde à presença de um possível elemento.	112
Gráfico 71	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na $\alpha$ -caseína s2. A cada pico corresponder a possível presença dessa estrutura.	112
Gráfico 72	Representação gráfica da quantidade real de cada aminoácido na $\alpha$ -caseína s2 (222 resíduos de aminoácido).	113
Gráfico 73	Representação gráfica da percentagem de cada aminoácido na $\alpha$ -caseína s2 (222 resíduos de aminoácido), relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	114
Gráfico 74	Representação gráfica da percentagem relativa dos aminoácidos da $\alpha$ -caseína s2, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa a média das percentagens relativas da proteína $\alpha$ -caseína s2.	115
Gráfico 75	Representação gráfica das cargas da $\alpha$ -caseína s2. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	116
Gráfico 76	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a $\beta$ -caseína (224 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).	116
Gráfico 77	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a $\beta$ -caseína (224 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	117
Gráfico 78	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a $\beta$ -caseína, usando a escala de Hopp & Wood (azul). Zonas de ligação à IgE (setas vermelhas) e possíveis zonas de ligação antigénicas (setas azuis).	118
Gráfico 79	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na $\beta$ -caseína. Cada pico corresponde à presença de um possível elemento.	118
Gráfico 80	Representação gráfica do prognóstico de Lupas para as super-hélice- $\alpha$ na $\beta$ -caseína. A cada pico corresponde a possível presença dessa estrutura.	119
Gráfico 81	Representação gráfica da quantidade real de cada aminoácido na $\beta$ -caseína (224 resíduos de aminoácido).	119

Gráfico 82	Representação gráfica da percentagem de cada aminoácido na $\beta$ -caseína, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína. Isto facilita a observação da composição relativa desta proteína.	120
Gráfico 83	Representação gráfica da percentagem relativa dos aminoácidos da $\beta$ -caseína, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	121
Gráfico 84	Representação gráfica das cargas da $\beta$ -caseína. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	122
Gráfico 85	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a $\beta$ -lactoglobulina (178 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).	123
Gráfico 86	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a $\beta$ -lactoglobulina (178 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	123
Gráfico 87	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a $\beta$ -lactoglobulina, usando a escala de Hopp & Wood (azul). Zonas antigénicas conhecidas (setas vermelhas) e zona antigénica possível (seta azul).	124
Gráfico 88	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na $\beta$ -lactoglobulina. Cada pico corresponde à presença de um possível elemento.	124
Gráfico 89	Representação gráfica do prognóstico de Lupas para as super-hélice- $\alpha$ na $\beta$ -lactoglobulina. A cada pico corresponder a possível presença dessa estrutura.	125
Gráfico 90	Representação gráfica da quantidade real de cada aminoácido na $\beta$ -lactoglobulina (178 resíduos de aminoácido).	125
Gráfico 91	Representação gráfica da percentagem de cada aminoácido na $\beta$ -lactoglobulina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína. Isto facilita a observação da composição relativa desta proteína.	126
Gráfico 92	Representação gráfica da percentagem relativa dos aminoácidos da $\beta$ -lactoglobulina, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	127
Gráfico 93	Representação gráfica das cargas da $\beta$ -lactoglobulina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	128
Gráfico 94	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a $\beta$ -2-microglobulina (118 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).	129
Gráfico 95	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a $\beta$ -2-microglobulina (118 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	129
Gráfico 96	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a $\beta$ -2-microglobulina, usando a escala de Hopp & Wood (azul). Zonas antigénicas possíveis (setas azuis).	130
Gráfico 97	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na $\beta$ -2-microglobulina. Cada pico corresponde à presença de um possível elemento.	130
Gráfico 98	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na $\beta$ -2-microglobulina.	131

Gráfico 99	Representação gráfica da quantidade real de cada aminoácido na $\beta$ -2-microglobulina (118 resíduos de aminoácido).	131
Gráfico 100	Representação gráfica da percentagem de cada aminoácido na $\beta$ -2-microglobulina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na $\beta$ -2-microglobulina. Isto facilita a observação da composição relativa desta proteína.	132
Gráfico 101	Representação gráfica da percentagem relativa dos aminoácidos da $\beta$ -2-microglobulina, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na $\beta$ -2-microglobulina.	133
Gráfico 102	Representação gráfica das cargas da $\beta$ -2-microglobulina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	133
Gráfico 103	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a osteopontina (278 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).	134
Gráfico 104	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a osteopontina (278 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	135
Gráfico 105	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a osteopontina, usando a escala de Hopp & Wood. Zonas antigénicas possíveis (setas azuis).	135
Gráfico 106	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na osteopontina. Cada pico corresponde à presença de um possível elemento.	136
Gráfico 107	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na osteopontina.	136
Gráfico 108	Representação gráfica da quantidade real de cada aminoácido na osteopontina (278 resíduos de aminoácido).	137
Gráfico 109	Representação gráfica da percentagem de cada aminoácido na osteopontina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na osteopontina.	138
Gráfico 110	Representação gráfica da percentagem relativa dos aminoácidos da osteopontina, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão situados à direita e os mais hidrofóbicos estão situados à esquerda O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na osteopontina.	139
Gráfico 111	Representação gráfica das cargas da osteopontina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	139
Gráfico 112	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a ubiquitina (76 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).	140
Gráfico 113	Representação gráfica do cálculo do índice de hélice- $\alpha$ para a ubiquitina (76 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).	141
Gráfico 114	Representação gráfica do cálculo da antigenicidade $H_{Ag}$ para a ubiquitina, usando a escala de Hopp & Wood (azul). Zonas antigénicas possíveis (setas azuis).	141
Gráfico 115	Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$ (azul); folha- $\beta$ (vermelho) e volta (verde), na ubiquitina. Cada pico corresponde à presença de um possível elemento.	142

Gráfico 116	Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$ na ubiquitina.	142
Gráfico 117	Representação gráfica da quantidade real de cada aminoácido na ubiquitina (76 resíduos de aminoácido).	143
Gráfico 118	Representação gráfica da percentagem de cada aminoácido na ubiquitina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	144
Gráfico 119	Representação gráfica da percentagem relativa dos aminoácidos da ubiquitina, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão à direita e os mais hidrofóbicos estão à esquerda O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.	145
Gráfico 120	Representação gráfica das cargas da ubiquitina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.	146
Gráfico 121	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a BSA (607 resíduos de aminoácidos). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	155
Gráfico 122	Representação gráfica do índice de hélice- $\alpha$ $H_\alpha$ para a BSA (607 resíduos de aminoácidos). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	156
Gráfico 123	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a lisozima C (147 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	157
Gráfico 124	Representação gráfica do índice de hélice- $\alpha$ $H_\alpha$ para a lisozima C (147 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	158
Gráfico 125	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a ovalbumina (386 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	159
Gráfico 126	Representação gráfica do índice de hélice- $\alpha$ $H_\alpha$ para a ovalbumina (386 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	160
Gráfico 127	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a ovomucina (2108 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	161
Gráfico 128	Representação gráfica do índice de hélice- $\alpha$ $H_\alpha$ para a ovomucina (2108 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	162
Gráfico 129	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a ovomucóide (210 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	163
Gráfico 130	Representação gráfica do índice de hélice- $\alpha$ $H_\alpha$ para a ovomucóide (210 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	164
Gráfico 131	Representação gráfica do cálculo da hidrofobicidade $H_7$ para a $\alpha$ -caseína s1 (214	165

	resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	
Gráfico 132	Representação gráfica do índice de hélice- $\alpha$ H $\alpha$ para a $\alpha$ -caseína s1 (214 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	166
Gráfico 133	Representação gráfica do cálculo da hidrofobicidade H $_7$ para a $\alpha$ -caseína s2 (222 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	167
Gráfico 134	Representação gráfica do índice de hélice- $\alpha$ H $\alpha$ para a $\alpha$ -caseína s2 (222 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	168
Gráfico 135	Representação gráfica do cálculo da hidrofobicidade H $_7$ para a $\beta$ -caseína (224 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	169
Gráfico 136	Representação gráfica do índice de hélice- $\alpha$ H $\alpha$ para a $\beta$ -caseína (224 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	170
Gráfico 137	Representação gráfica do cálculo da hidrofobicidade H $_7$ para a $\beta$ -lactoglobulina (178 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	171
Gráfico 138	Representação gráfica do índice de hélice- $\alpha$ H $\alpha$ para a $\beta$ -lactoglobulina (178 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	172
Gráfico 139	Representação gráfica do cálculo da hidrofobicidade H $_7$ para a $\beta$ -2-microglobulina (118 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	173
Gráfico 140	Representação gráfica do índice de hélice- $\alpha$ H $\alpha$ para a $\beta$ -2-microglobulina (118 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	174
Gráfico 141	Representação gráfica do cálculo da hidrofobicidade H $_7$ para a osteopontina (278 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	175
Gráfico 142	Representação gráfica do índice de hélice- $\alpha$ H $\alpha$ para a osteopontina (278 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	176
Gráfico 143	Representação gráfica do cálculo da hidrofobicidade H $_7$ para a ubiquitina (76 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	177
Gráfico 144	Representação gráfica do índice de hélice- $\alpha$ H $\alpha$ para a ubiquitina (76 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).	178



## LISTA DE ABREVIATURAS

<b>Abreviatura</b>	<b>Designação</b>
MHC	Main Histocompatibility Complex
Th1	Linfócitos T Helper 1
Th2	Linfócitos T Helper 2
BSA	Bovin serum albumin
PI	Ponto isoeléctrico

# 1. INTRODUÇÃO TEÓRICA

## 1.1. BIOINFORMÁTICA

Estes últimos anos, a bioquímica conheceu inúmeras mudanças. As técnicas de análise e de dosagem das biomoléculas mostram progressos significativos. Essas técnicas permitem um estudo cada vez mais minucioso das biomoléculas e em particular as macromoléculas informativas tais como os ácidos nucleicos e as proteínas. O desenvolvimento da bioquímica, para o estudo dos fenómenos biológicos, integrando biologia molecular e a genética, originou a necessidade de desenvolver novos utensílios informáticos (Coutouly *et al.*, 2006).

A Bioinformática é uma disciplina de análise “*in Silico*” da informação biológica contida nas sequências de nucleótidos e nas sequências de aminoácidos. O seu aparecimento, nos anos oitenta, coincide com os primeiros bancos de dados (EMBL e GenBank). A partir da década de noventa, a bioinformática tornou-se indispensável com a acumulação dos dados da sequenciação de genomas completos. (Coutouly *et al.*, 2006).

Uma das principais utilizações da bioinformática, é no estudo de proteínas, seja ao nível da catálise, seja ao nível da interacção com outras moléculas ou biomoléculas (Bujnicki, 2008).

## 1.2. OS AMINOÁCIDOS

### 1.5.3. Estrutura geral

Um  $\alpha$ -aminoácido é constituído por um átomo central de carbono, chamado carbono- $\alpha$ , ligado a um grupo amina ( $-\text{NH}_2$ ), pelo menos um grupo carboxilo ( $-\text{COOH}$ ), um átomo de hidrogénio e a um grupo R diferenciado, sendo este último referido como uma cadeia lateral, (Figura 1) (Berg et al., 2002; Chang, 1995). Os  $\alpha$ -aminoácidos são moléculas quirais, isto significa que um dos isómeros é a imagem oposta do outro, (imagem no espelho). Essas duas formas são chamadas de isómero L e isómero D. Nas proteínas apenas existe o isómero L para cada aminoácido que as compõe. Quase todos os aminoácidos L têm uma configuração absoluta S (do latim *sinister* que significa “esquerda”) (Berg et al., 2002).

Entre as múltiplas funções que os aminoácidos cumprem nas células, há uma que consiste em fornecer as unidades monoméricas a partir das quais as cadeias polipeptídicas são

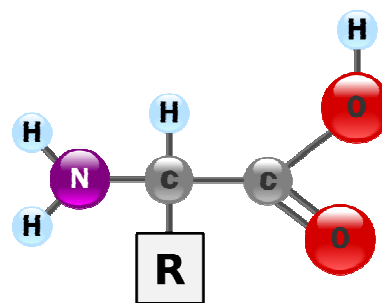


Figura 1: Estrutura elementar dos aminoácidos

construídas (Murray *et al.*, 2002). Deste modo, pode afirmar-se que os aminoácidos são as unidades estruturais básicas das proteínas (Berg *et al.*, 2002).

### 1.5.3. Propriedades iónicas

Os aminoácidos em solução com pH neutro são predominantemente iões bipolares, também conhecidos como “*zwitterions*”. Na forma bipolar, a amina é protonada ( $-\text{NH}_3^+$ ) e o carboxilo é desprotonado ( $-\text{COO}^-$ ). O estado de ionização de um aminoácido varia com o pH, (Figura 2). Em solução ácida, o grupo amina está protonada ( $-\text{NH}_3^+$ ) e o grupo carboxilo sem dissociação ( $-\text{COOH}$ ). À medida que o pH aumenta, o ácido carboxílico é o primeiro grupo a perder um protão, porque o seu  $\text{pK}_a$  está próximo de 2. A forma bipolar persiste até o pH se aproximar de 9, quando a amina protonada perde um protão (Berg *et al.*, 2002; Voet *et al.*, 2004).

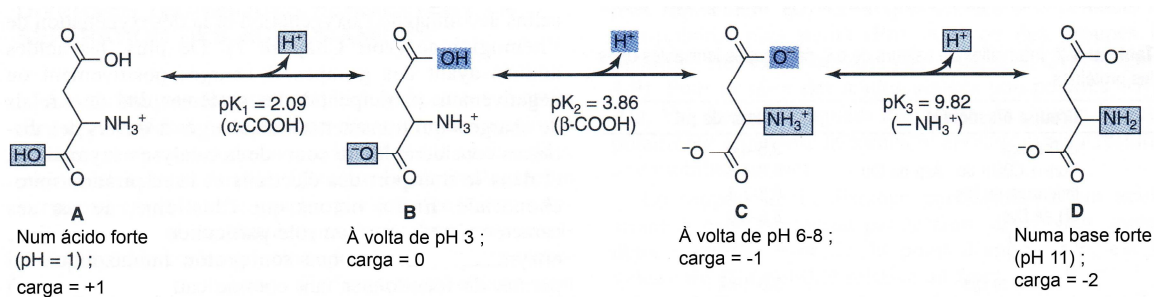


Figura 2: Equilíbrio protónico do ácido aspártico. Adaptado de Murray *et al.*, 2002.

Os valores de  $\text{pK}$  dos 20  $\alpha$ -aminoácidos “standard” das proteínas, estão tabelados na tabela 1. O  $\text{pK}_1$  e  $\text{pK}_2$ , referem-se especificamente ao ácido carboxílico e ao grupo amina. O  $\text{pK}_R$  refere-se às cadeias laterais com as propriedades ácido-base (Voet *et al.*, 2004). Sete dos vinte aminoácidos têm cadeias laterais facilmente ionizáveis. Eles tanto podem doar como aceitar protões para facilitar reacções ou formar ligações iónicas. É o caso da tirosina, da cisteína, da arginina, da lisina, da histidina e dos ácidos aspártico e glutâmico, sem esquecer o grupo amina  $\alpha$  terminal e o grupo carboxilo  $\alpha$  terminal (Berg *et al.*, 2002). O valor característico de  $\text{pK}_a$  destes sete aminoácidos é usado no cálculo das cargas eléctricas, no estudo de proteínas (Fibly, 1997).

Na escala de pH fisiológico, tanto o ácido carboxílico como o grupo amina dos ácidos  $\alpha$ -amina estão completamente ionizados (Voet *et al.*, 2004).

**Tabela 1:** Valores do pK dos principais 20 aminoácidos para o ácido carboxílico, o grupo amina e o grupo R. assim como para o C-terminal e N-terminal. Adaptado de Murray et al., 2002; Berg et al., 2002

Aminoácidos	pK <sub>1</sub> ( $\alpha$ -COOH)	pK <sub>2</sub> ( $\alpha$ -NH <sub>3</sub> <sup>+</sup> )	pK <sub>R</sub> (Grupo R)
Glicina	2,4	9,8	
Alanina	2,4	9,9	
Valina	2,2	9,7	
Leucina	2,3	9,7	
Isoleucina	2,3	9,8	
Serina	2,2	9,2	≈ 13
Treonina	2,1	9,1	≈ 13
Tirosina	2,2	9,1	10,1
Cisteína	1,9	10,8	8,3
Metionina	2,1	9,3	
Aspartato	2,0	9,9	3,9
Asparagina	2,1	8,8	
Glutamato	2,1	9,5	4,1
Glutamina	2,2	9,1	
Arginina	1,8	9,0	12,5
Lisina	2,2	9,2	10,8
Histidina	1,8	9,3	6,0
Fenilalanina	2,2	9,2	
Triptofano	2,4	9,4	
Prolina	2,0	10,6	
C-termina			3,1
N-termina			8,0

### 1.5.3. Solubilidade dos aminoácidos

Tendo em conta que numerosos grupos carregados estão presentes nos aminoácidos, estes são facilmente solvatados por solventes polares, sendo por essa razão alguns aminoácidos muito solúveis por exemplo na água e no etanol. No sentido inverso, eles são totalmente insolúveis em solventes apolares tais como o benzeno, o hexano ou o éter (Murray *et al.*, 2002; Nelson e Cox, 2004).

As cadeias laterais alifáticas maiores são hidrofóbicas, isto é, tendem a aglomerarem-se para evitar o contacto com a água. A estrutura tridimensional de proteínas hidrossolúveis é estabilizada por esta tendência dos grupos hidrofóbicos em agruparem-se (Berg *et al.*, 2002).

A prolina tem uma cadeia lateral alifática, mas difere dos restantes dezanove aminoácidos mais comuns que compõem as proteínas porque a sua cadeia lateral tanto pode ser ligada ao azoto como ao átomo de carbono. Ela influencia pronunciadamente a arquitectura das proteínas porque a sua estrutura em anel faz com que ela tenha maior restrição de conformação do que os outros aminoácidos. A

fenilalanina é totalmente hidrofóbica, enquanto que a tirosina e o triptofano são um pouco menos, devido aos seus grupos hidroxilos ( $-OH$ ) e aminas ( $-NH_2$ ). Os grupos hidroxilo da serina e treonina conferem-lhes um comportamento muito mais hidrofílico e reactivo em comparação com a alanina e a valina (Berg et al., 2002; Murray *et al.*, 2002). Em resumo, podemos dizer que o grupo alifático R da alanina; da leucina; da valina e da isoleucina, é hidrofóbico, assim como o grupo aromático R da fenilalanina; da tirosina e do triptofano (Murray *et al.*, 2002). Na tabela 2 observa-se para os vinte aminoácidos mais comuns uma classificação em hidrofóbicos e hidrofílicos. Na tabela 3, observa-se os valores de solubilidade dos mesmos vinte aminoácidos.

**Tabela 2:** Classificação dos aminoácidos L das proteínas, baseada na hidrofobicidade relativa e na hidrofiliicidade relativa. Adaptado de Murray et al., 2002.

Hidrofóbicos	Hidrofílicos	
Alanina	Arginina	Serina
Isoleucina	Asparagina	Treonina
Leucina	Aspartato	
metionina	Cisteína	
Fenilalanina	Glutamato	
Prolina	Glutamina	
Triptofano	Glicina	
Tirosina	Histidina	
Valina	Lisina	

**Tabela 3:** Valores da solubilidade dos principais 20 aminoácidos. (Adaptado de Lide, 2009)

Aminoácidos	Solubilidade a 25° de aminoácidos (g/kg de água)	Aminoácidos	Solubilidade a 25° de aminoácidos (g/kg de água)
Tirosina	0,51	Histidina	43,5
Aspartato	5,04	Metionina	56
Lisina	5,8	Valina	88
Glutamato	8,6	Treonina	90,6
Triptofano	13,2	Alanina	166,9
Leucina	23,8	Arginina	182,6
Asparagina	25,1	Glicina	239
Fenilalanina	26,9	Serina	250
Isoleucina	34,2	Prolina	1625
Glutamina	42	Cisteína	Muito solúvel

### 1.3. AS PROTEÍNAS

As proteínas são as macromoléculas mais versáteis dos sistemas vivos e servem para funções cruciais em quase todos os processos biológicos (Berg *et al.*, 2002). Apesar de não existir um sistema de classificação internacional, as proteínas podem ser classificadas com base na sua solubilidade, forma, funções biológicas ou da sua estrutura tridimensional (Murray *et al.*, 2002). Funcionam como catalisadores, transportam e armazenam outras moléculas, tais como o oxigénio ou o dióxido de carbono (CO<sub>2</sub>). Fornecem apoio mecânico e protecção imunitária, transmitem impulsos nervosos e controlam o crescimento, assim como a diferenciação (Berg *et al.*, 2002; Chang, 1995). Elas são polímeros constituídos de unidades monoméricas denominadas por aminoácidos (Berg *et al.*, 2002). A maioria das proteínas contém, em proporções variáveis, os mesmos 20  $\alpha$ -L aminoácidos. Para além disso, muitas proteínas específicas contêm aminoácidos  $\alpha$ -L, derivados de um dos vinte resíduos de base através de processos que decorrem após a formação do esqueleto polipeptídico (Murray *et al.*, 2002).

A função de uma proteína está directamente dependente da sua estrutura tridimensional (Berg *et al.*, 2002). Os tipos de aminoácidos, a ordem pela qual eles se ligam e as suas orientações espaciais, ditam a estrutura tridimensional das proteínas, assim com as suas propriedades biológicas (Murray *et al.*, 2002). As proteínas são a personificação da transição do mundo unidimensional das sequências de aminoácidos para o mundo tridimensional das moléculas capazes de diversas actividades (Berg *et al.*, 2002).

As proteínas podem não só interagir umas com as outras como também podem associar-se a outras macromoléculas biológicas, formando estruturas complexas. Algumas proteínas têm uma estrutura rígida, enquanto que outras apresentam flexibilidade limitada (Berg *et al.*, 2002).

Um dos maiores objectivos da análise da sequência proteica é compreender a relação entre a sequência de aminoácidos e a estrutura tridimensional das proteínas. Se esta relação for conhecida, a estrutura da proteína poderá ser prevista com segurança a partir da sequência de aminoácidos. Infelizmente, a relação em causa não é assim tão simples, contudo estão já a ser desenvolvidos métodos que visam a categorização das proteínas com base na sua estrutura e sequência, como a modelização com recurso a métodos informáticos (Mount, 2004).

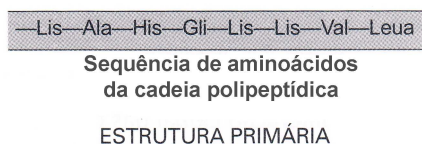
### 1.5.3. Estrutura primária

Uma cadeia peptídica é constituída por uma parte que se repete regularmente, à qual se dá o nome de cadeia principal (Figura 3) e uma parte variável que corresponde às cadeias laterais (Berg *et al.*, 2002). A estrutura primária de uma cadeia polipeptídica de uma proteína, é a ordem pela qual

os aminoácidos estão ligados uns aos outros. Isto inclui a localização das ligações dissulfureto (Murray *et al.*, 2002). A cadeia principal é rica em potencial de formação de pontes de hidrogénio. Cada unidade contém um grupo carboxilo, que é um bom aceitador para a formação de pontes de hidrogénio, com excepção da prolina que tem um grupo amina e que é um bom doador para formar este mesmo tipo de pontes. Estes grupos interagem uns com os outros e com grupos funcionais de cadeias laterais estabilizando determinadas estruturas (Berg *et al.*, 2002).

A biossíntese de ligações peptídicas requer uma entrada de energia livre. No entanto, as ligações peptídicas têm cinéticas estáveis. O tempo de vida de uma ligação peptídica em solução aquosa, na ausência de um catalisador, aproxima-se de mil anos (Berg *et al.*, 2002).

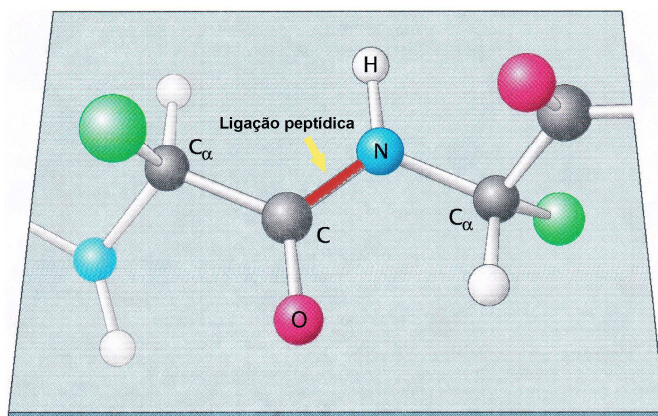
Uma cadeia peptídica tem polaridade porque as suas extremidades são diferentes, com um grupo amina  $\alpha$  numa ponta e um carboxilo  $\alpha$  na outra (Berg *et al.*, 2002). Em algumas proteínas, a cadeia peptídica linear está interligada. Dessas interligações, as mais comuns são as pontes dissulfureto, cada uma formada pela oxidação de um par de radicais de cisteína (Berg *et al.*, 2002).



**Figura 3:** Estrutura primária das proteínas, sequência de aminoácidos. (Adaptado de Goldsby *et al.*, 2000)

As proteínas têm sequências particulares de aminoácidos que são especificadas por genes. A sequência de nucleótidos no DNA, específica uma sequência complementar de nucleótidos no RNA, que por sua vez determina a sequência de aminoácidos numa proteína. Cada um dos vinte aminoácidos é codificado por uma ou mais sequências específicas de três nucleótidos (Berg *et al.*, 2002).

A ligação peptídica é rígida e plana. Para um par de aminoácidos unidos por uma ligação peptídica, seis átomos situam-se no mesmo plano (co-planar) como se pode ver na (Figura 4). O carbono  $\alpha$  com o grupo ( $-\text{CO}$ ) do primeiro aminoácido e o grupo ( $-\text{NH}$ ) com o carbono  $\alpha$  do segundo aminoácido (Berg *et al.*, 2002).



**Figura 4:** Representação das ligações peptídicas. Num par de aminoácidos ligados, seis átomos (C $\alpha$ , C, O, N e C $\alpha$ ), estão no mesmo plano. As cadeias laterais, representadas a verde, situam-se noutro plano. (Adaptado de Berg *et al.*, 2002)

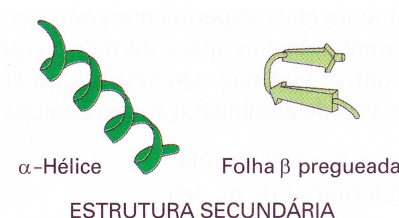
consequências importantes na organização da estrutura proteica ao nível primário. No entanto, em volta das restantes ligações do esqueleto polipeptídico, existe um elevado grau de rotação. É o que acontece por exemplo nas ligações entre o carbono  $\alpha$  (C $\alpha$ ) e o carbonilo ( $-\text{CO}$ ) do primeiro resíduo e entre o carbono  $\alpha$  (C $\alpha$ ) e o grupo ( $-\text{NH}$ ) do segundo resíduo. A ligação C $\alpha$ -CO denomina-se ângulo psi ( $\Psi$ ) e a ligação C $\alpha$ -NH, denomina-se ângulo fi ( $\Phi$ ) (Murray *et al.*, 2002).

A natureza da ligação química dentro de um péptido explica esta preferência geométrica (Berg *et al.*, 2002). Os péptidos são representados por uma ligação simples entre o carbono  $\alpha$  e a amina, esta ligação possui na realidade um carácter de dupla ligação parcial. Não existe liberdade de rotação em volta da ligação que une estes dois átomos. Esta rigidez parcial tem

### 1.5.3. Estrutura secundária

As cadeias peptídicas podem dobrar-se em estruturas regulares, tais como: hélice- $\alpha$ ; folha- $\beta$ ; voltas (turns) e alças (Figura 5). A distribuição destas estruturas ao longo de uma cadeia proteica é muitas vezes designada como sendo a estrutura secundária (Berg *et al.*, 2002).

Inúmeras interações não covalentes, fracas mas em número elevado, estabilizam a conformação das proteínas. Entre essas forças constam as pontes de hidrogénio, as interações hidrofóbicas, as interações electrostáticas e as forças de van der Waals (Murray *et al.*, 2002).



**Figura 5:** Estrutura secundária, representação de hélice- $\alpha$  e folhas- $\beta$ . (Adaptado de Goldsby *et al.*, 2000)

Os resíduos de aminoácidos que possuem grupos R polares, encontram-se geralmente à superfície das proteínas globulares onde formam pontes de hidrogénio, essencialmente com as moléculas de água. Por outro lado, os resíduos de aminoácidos que possuem grupos R apolares, encontram-se geralmente no interior das proteínas, devido às interacções hidrofóbicas (Murray *et al.*, 2002).

A hélice- $\alpha$  é uma estrutura em espiral estabilizada por pontes de hidrogénio e por forças de van der Waals, dentro da cadeia polipeptídica (Berg *et al.*, 2002). Ela forma-se de maneira espontânea visto que a sua configuração é estável e a sua cadeia polipeptídica tem baixa energia (Murray *et al.*, 2002).

As pontes de hidrogénio estabelecem-se entre o grupo ( $-\text{NH}$ ) e ( $-\text{CO}$ ) da cadeia principal. O grupo ( $-\text{CO}$ ) de cada aminoácido forma uma ponte de hidrogénio com o ( $-\text{NH}$ ) do aminoácido que está situado a quatro unidades à frente na sequência linear. Com excepção aos aminoácidos que estão perto das pontas de uma hélice- $\alpha$ , todos os grupos ( $-\text{CO}$ ) e ( $-\text{NH}$ ) formam pontes de hidrogénio. Cada aminoácido relaciona-se com o seguinte por um deslocamento de  $1,5\text{\AA}$  ao longo do eixo da hélice- $\alpha$  e por uma rotação de  $100^\circ$ , o que dá  $3,6$  radicais de aminoácidos por volta de hélice. Isto significa que aminoácidos que estão a três e quatro unidades, estão espacialmente bem próximos uns dos outros na hélice- $\alpha$ . Por outro lado, aminoácidos à distância de duas unidades na sequência linear encontram-se situados em lados opostos da hélice- $\alpha$  o que é muito improvável que façam contacto (Berg *et al.*, 2002).

O sentido da rotação de uma hélice pode ser para a direita (sentido horário “dextrógera”), ou para a esquerda (sentido anti-horário “levógera”). Tanto a hélice- $\alpha$  direita como a esquerda têm conformações permitidas. No entanto, as hélices dextrógera são energeticamente mais favoráveis porque há menos impedimentos estereoquímicos entre as cadeias laterais e a principal. Todas as hélices- $\alpha$  encontradas nas proteínas são dextrógeras. Em esquemas de proteínas, as hélices- $\alpha$  são representadas por cadeias retorcidas e/ou bastões (Berg *et al.*, 2002).

Os resíduos de aminoácidos da alanina, da glicina, da leucina e da metionina são mais frequentemente encontrados nas hélices- $\alpha$  do que a prolina, a serina e a tirosina. Esta tendência é no entanto pouco útil para os prognósticos estruturais. Tendo em conta que o átomo de azoto da prolina não pode formar pontes de hidrogénio, esta é geralmente encontrada nas primeiras voltas da hélice- $\alpha$  (Murray *et al.*, 2002).

Apesar de, na maior parte das vezes as hélices- $\alpha$  se encontrarem à superfície das proteínas, elas podem estar totalmente localizadas no interior destas. A hélice- $\alpha$  anfifílica, um caso particular onde os resíduos são alternadamente hidrofóbicos e hidrofílicos, quer dizer que a cada três ou quatro monómeros, formam-se geralmente quando as hélices- $\alpha$  se encontram em ambientes polares e apolares. As hélices- $\alpha$  anfifílicas encontram-se nas lipoproteínas plasmáticas, assim como em algumas hormonas polipeptídicas como por exemplo, na glicoproteína do vírus da sida (Murray *et al.*, 2002).

A segunda configuração regular presente nas proteínas é a folha- $\beta$  pregueada. A folha- $\beta$  pregueada, vulgarmente denominada por “folha- $\beta$ ”, difere muito da hélice- $\alpha$ . Ela é constituída por uma “cadeia- $\beta$ ” quase totalmente distendida em vez de firmemente enrolada como acontecia na hélice- $\alpha$ . A distância entre aminoácidos adjacentes ao longo de uma cadeia- $\beta$  é de 3,5Å, em contraste com 1,5Å ao longo da hélice- $\alpha$ . As cadeias laterais de aminoácidos adjacentes apontam para sentidos opostos. Uma folha- $\beta$  é formada pela união de dois ou mais cadeias- $\beta$  por pontes de hidrogénio (Berg *et al.*, 2002). As cadeias- $\beta$  são retorcidos no sentido da mão direita. Estes formam o coração central de inúmeras proteínas globulares (Murray *et al.*, 2002). Algumas cadeias adjacentes podem estender-se em sentidos opostos, (antiparalelas), ou no mesmo sentido, (paralelas). No arranjo antiparalelo, os grupos (–NH) e (–CO) de cada aminoácido fazem respectivamente pontes de hidrogénio com os grupos (–CO) e (–NH) da cadeia adjacente. No arranjo paralelo, o esquema de pontes de hidrogénio é um pouco mais complicado. Para cada aminoácido, o grupo amina faz ponte de hidrogénio com o grupo carboxilo dum aminoácido na cadeia adjacente enquanto que o carboxilo faz ponte com a amina do aminoácido a dois monómeros de distância ao longo da cadeia. As folhas- $\beta$  podem ser constituídas por quatro, dez ou mais cadeias- $\beta$ . Elas podem ter arranjo exclusivamente paralelo, antiparalelo ou então misto. Em esquemas proteicos, as cadeias- $\beta$  são geralmente representadas por setas largas apontando no sentido da extremidade carboxi-terminal, indicando o tipo de folha- $\beta$ , paralela ou antiparalela. As folhas- $\beta$  podem ser relativamente planas mas a maioria adopta uma forma um pouco retorcida (Berg *et al.*, 2002).

A maioria das proteínas tem forma globular compacta. Isto leva necessariamente à inversão do sentido das suas cadeias peptídicas. Muitas destas inversões são feitas por um elemento estrutural chamado “volta reversa” ou “cotovelo”, (*Turn* em inglês). Em muitas destas voltas, o grupo carboxilo do aminoácido  $i$  de um péptido, forma uma ponte de hidrogénio com o grupo amina do aminoácido  $i+3$ . Esta interacção estabelece mudanças abruptas de sentido de cadeia peptídica (Berg *et al.*, 2002). Muitas vezes, as voltas contêm prolina e/ou glicina (Murray *et al.*, 2002).

Existem outras estruturas responsáveis pela inversão da cadeia peptídica. Essas estruturas denominam-se “alças” ou por vezes “novelo”. Ao contrário das hélice- $\alpha$  e das folhas- $\beta$ , as alças não têm estrutura periódicas regulares, mas são rígidas e bem definidas, o que não implica que sejam menos importantes do ponto de vista biológico (Berg *et al.*, 2002). Elas não devem ser confundidas com as alças estáticas, (Enrolamento ao acaso; “*Random coil*” em inglês), um termo que descreve configurações desordenadas e biologicamente pouco importantes das proteínas desnaturadas.

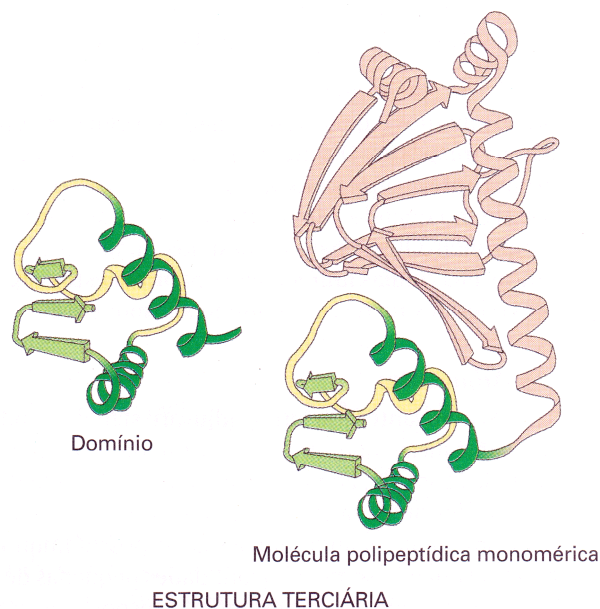
As alças de diferentes tamanhos e de diferentes formas, constituem as características principais da superfície das proteínas. Expostas aos solventes, ricas em resíduos polares ou carregadas mas não possuindo estrutura secundária regular, as alças podem ligar folhas- $\beta$  antiparalelas adjacentes (Murray *et al.*, 2002). Voltas e alças encontram-se sempre na

superfície das proteínas, participando muitas vezes em interações com outras proteínas ou outras moléculas como por exemplo os antígenos (Berg *et al.*, 2002).

### 1.5.3. Estrutura terciária

A trajectória global da cadeia peptídica de uma proteína é referida como sendo a estrutura terciária (Figura 6). O facto marcante é que o interior é constituído quase na totalidade por aminoácidos apolares, tais como a fenilalanina; leucina; metionina e a valina (Berg *et al.*, 2002). Em muitas proteínas, a parte externa tanto pode ser constituída de radicais polares como de radicais apolares, sendo os polares mais numerosos (Berg *et al.*, 2002; Voet *et al.*, 2004). Esta distribuição contrastante de radicais polares e apolares revela uma faceta importante da arquitectura das

proteínas. Num meio aquoso, o enrolamento das proteínas é impulsionado pela forte tendência dos aminoácidos hidrofóbicos para serem excluídos da água. Um sistema é termodinamicamente mais estável quando os grupos hidrofóbicos são aglomerados do que quando são projectados para o meio aquoso. A cadeia peptídica enrola-se de forma a que as suas cadeias laterais hidrofóbicas fiquem imersas e que as suas cadeias polares com carga eléctrica fiquem à superfície. Muitas hélices- $\alpha$  e folhas- $\beta$  têm uma face hidrofóbica orientada para o interior da proteína e uma face hidrofílica apontada para a solução. O papel da cadeia principal que acompanha as cadeias laterais hidrofóbicas também é importante. Um ( $-\text{CO}$ ) ou ( $-\text{NH}$ ) peptídico não emparelhado, prefere muito mais a água do que um meio apolar. Para emergir um segmento da cadeia principal num ambiente hidrofóbico basta emparelhar todos os seus grupos ( $-\text{CO}$ ) e ( $-\text{NH}$ ) por pontes de hidrogénio. Este emparelhamento também existe nas hélices- $\alpha$  e/ou folhas- $\beta$ . Interações de van der Waals entre cadeias laterais hidrocarbonadas firmemente acondicionadas também contribuem para a estabilidade das proteínas (Berg *et al.*, 2002).



**Figura 6:** Estrutura terciária, configuração das proteínas constituídas pelas diferentes estruturas: hélices- $\alpha$ ; folhas- $\beta$  e voltas. (Adaptado de Goldsby *et al.*, 2000)

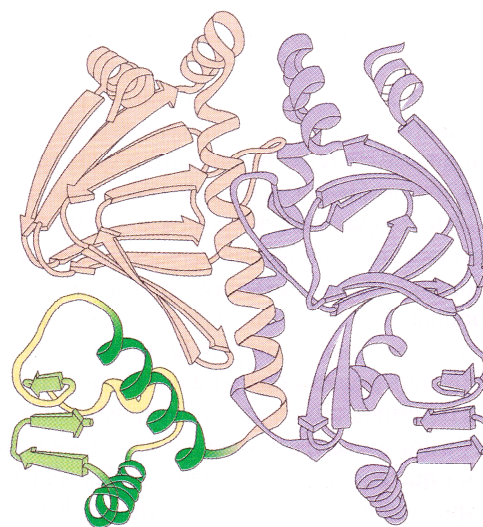
### 1.3.1. Estrutura quaternária

A maioria das proteínas tem mais do que uma cadeia peptídica (Figura 7). Essas proteínas exibem um quarto nível de organização estrutural. Cada cadeia peptídica nesse tipo de proteínas é chamada de subunidade. O tipo mais simples de estrutura quaternária é um dímero, constituído de duas subunidades idênticas (Berg *et al.*, 2002).

### 1.3.5. Localização dos aminoácidos na estrutura tridimensional

Os aminoácidos têm tendência a formar hélices- $\alpha$ ; folhas- $\beta$  e voltas ou alças. Como era de esperar, nem todos

os aminoácidos formam cada uma destas estruturas. Uns têm mais tendência para formar um tipo e/outros têm mais tendência para formar outro tipo de estrutura. Na tabela 4 pode ver-se a tendência de cada aminoácido em formar hélices- $\alpha$ ; folhas- $\beta$  e voltas. Pode observar-se por exemplo que a alanina; o glutamato e a leucina, tendem a estar presentes em hélice- $\alpha$  enquanto que a valina e a isoleucina tendem a formar folhas- $\beta$ . A glicina; a asparagina e a prolina têm tendência para ficar nas voltas. A valina; treonina e isoleucina são facilmente acomodados em folhas- $\beta$ , onde as suas cadeias laterais projectam-se para fora do plano que contem a cadeia principal. A serina; o aspartato e a asparagina, tendem a perturbar as hélices- $\alpha$  porque as suas cadeias laterais contêm doadores ou aceptores de pontes de hidrogénio próximos da cadeia principal, onde competem pelos grupos ( $-\text{NH}$ ) e ( $-\text{CO}$ ) desta última. A prolina tende a perturbar tanto as hélices- $\alpha$  como as folhas- $\beta$  porque não tem um grupo ( $-\text{NH}$ ) e porque a sua estrutura em anel restringe o seu valor de  $\varphi$  (Psi, ângulo de rotação em torno da ligação entre o carbono  $\alpha$  e o carboxilo) próximo de  $60^\circ$ . A glicina adapta-se a todas as estruturas e por essa razão não favorece particularmente a formação de hélices- $\alpha$  (Berg *et al.*, 2002).



ESTRUTURA QUATERNÁRIA

**Figura 7:** Estrutura quaternária. As proteínas exibem quatro níveis de organização estrutural, (subunidades). Neste caso trata-se de um dímero, duas subunidades. (Adaptado de Goldsby *et al.*, 2000)

**Tabela 4:** Os aminoácidos estão agrupados de acordo com a sua preferência (a negrito) em formar hélices- $\alpha$ , folhas- $\beta$  e voltas. A arginina não mostra preferência significativa por qualquer uma das estruturas. Adaptado de Berg et al., 2002

Aminoácido	Hélice $\alpha$	Folha $\beta$	Volta
Ala	<b>1,29</b>	0,90	0,78
Cis	<b>1,11</b>	0,74	0,80
Leu	<b>1,30</b>	1,02	0,59
Met	<b>1,47</b>	0,97	0,39
Glu	<b>1,44</b>	0,75	1,00
Gln	<b>1,27</b>	0,80	0,97
His	<b>1,22</b>	1,08	0,69
Lis	<b>1,23</b>	0,77	0,96
Val	0,91	<b>1,49</b>	0,47
Ile	0,97	<b>1,45</b>	0,51
Fen	1,07	<b>1,32</b>	0,58
Tir	0,72	<b>1,25</b>	1,05
Trp	0,99	<b>1,14</b>	0,75
Tre	0,82	<b>1,21</b>	1,03
Gli	0,56	0,92	<b>1,64</b>
Ser	0,82	0,95	<b>1,33</b>
Asp	1,04	0,72	<b>1,41</b>
Asn	0,90	0,76	<b>1,28</b>
Pro	0,52	0,64	<b>1,91</b>
Arg	0,96	0,99	0,88

As preferências de configuração de radicais de aminoácidos não são totalmente inclinadas para uma ou outra estrutura. Encontraram-se algumas sequências penta, ou hexapeptídicas adoptando uma estrutura numa proteína e uma forma totalmente diferente noutra proteína. Pode então concluir-se que algumas sequências de aminoácidos não determinam univocamente a estrutura secundária. Interações terciárias, podem ser decisivas na especificação da estrutura secundária de alguns segmentos. O contexto é com frequência crucial na determinação do resultado da configuração. A configuração de uma proteína evoluiu para funcionar num determinado ambiente/ou contexto (Berg *et al.*, 2002).

Em resumo, a estrutura primária é a sequência de aminoácidos de uma cadeia polipeptídica. A estrutura secundária é o arranjo espacial de radicais de aminoácidos que estejam perto de/outro na sequência, dando origem a estruturas periódicas como a hélice- $\alpha$  e a folha- $\beta$ . A estrutura terciária refere-se ao arranjo espacial de radicais que estão longe na sequência e ao padrão de pontes dissulfureto. Por último, a estrutura quaternária refere-se ao arranjo espacial das subunidades e à natureza das suas interações (Berg *et al.*, 2002).

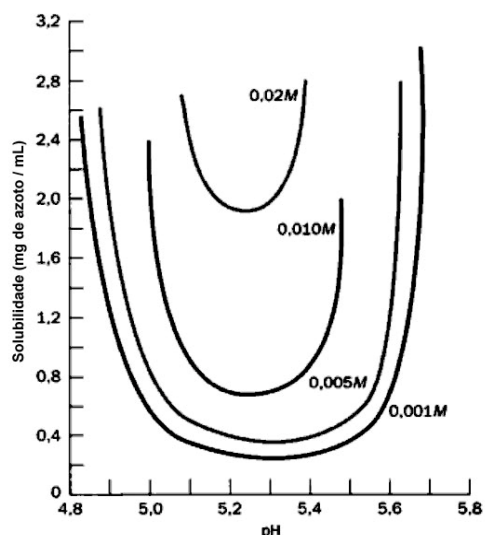
### 1.3.6. Solubilidade das proteínas

Os múltiplos grupos ácido-base numa proteína, tornam as suas propriedades de solubilidade dependentes da concentração em sal dissolvido, do carácter polar do solvente, do pH e da temperatura. As diferentes proteínas têm uma solubilidade que varia fortemente com as condições experimentais. Algumas proteínas precipitam em determinadas condições onde/outras permanecem perfeitamente solúveis. Este comportamento é sistematicamente usado para purificar proteínas (Voet *et al.*, 2004).

#### 1.3.6.1. Influência do pH

As proteínas possuem em geral, inúmeros grupos ionizáveis, dos quais os valores pK são diferentes. Existe um valor de pH característico de cada proteína, para a qual há tantas cargas positivas como cargas negativas na molécula. A este pH, que é o ponto isoeléctrico (pI) da proteína, ela tem uma carga formal igual a zero e consequentemente, permanece imóvel se for colocada num campo eléctrico (Voet *et al.*, 2004).

A figura 8 mostra que a solubilidade da  $\beta$ -lactoglobulina é mínima para um pH próximo do seu ponto isoeléctrico (pI), (5,2) em soluções de NaCl diluída e ela aumenta de forma mais ou menos simétrica de ambos os lados do pI. Esta variação de solubilidade é encontrada na maioria das proteínas e explica-se muito facilmente. Considerações físico-químicas sugerem que as propriedades de solubilidade de moléculas não carregadas, são insensíveis à concentração salina. Consequentemente, uma proteína no seu ponto isoeléctrico, não deveria dissolver-se em solução. No entanto, à medida que o pH se afasta do pI da proteína, o que provoca um aumento da carga formal da mesma, o “*Salting in*” encontra-se facilitado em razão do aumento das interacções electrostáticas entre moléculas vizinhas. Por consequente, nas soluções de concentrações salinas médias, a solubilidade de uma proteína em função do pH atingira um mínimo no ponto isoeléctrico da proteína e aumentara quando o pH se afastar do pI (Voet *et al.*, 2004; Garrett & Grisham, 2008).



**Figura ,8:** Solubilidade da  $\beta$ -lactoglobulina em função do pH, a concentrações de NaCl diferentes. Adaptado de Voet *et al.*, 2004

## 1.4. PROTEÍNAS EM ESTUDO

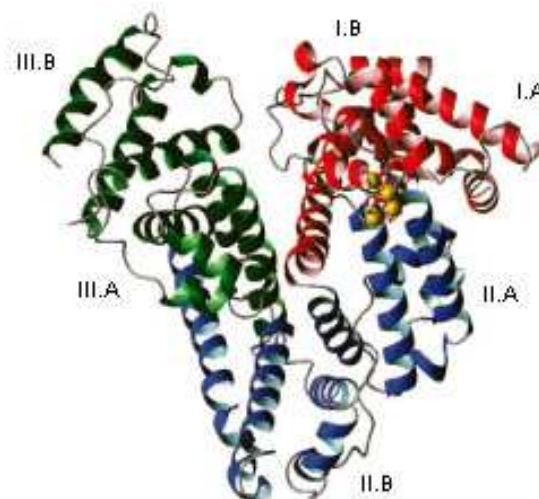
### 1.4.7. BSA

A albumina do soro bovino (BSA, em inglês), é nomeada (Bos d6) como sendo um alergénio (Jedrychowski & Wichers, 2009). Ela constitui 5% da totalidade das proteínas do soro do leite. A BSA é fisicamente e imunologicamente muito semelhante

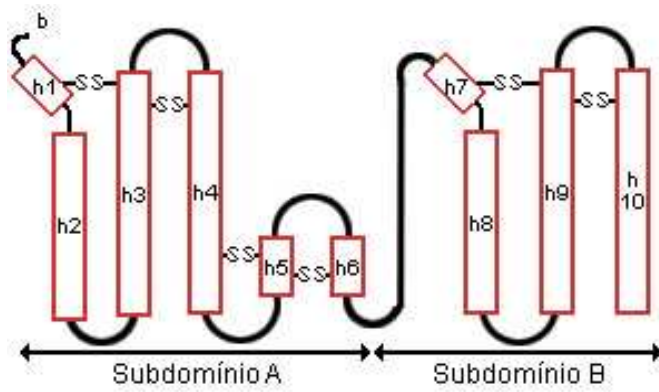
à albumina do soro humano. Ela regula em 80% a pressão osmótica do sangue, liga-se a vários cátions e é a principal transportadora dos ácidos gordos, hormonas e bilirrubina, sem a qual esta última não poderia dissolver-se no plasma. Para além de ser a proteína mais abundante no plasma, os seus níveis são inferiores no leite. Esta proteína foi observada como sendo um monómero e como sendo um dímero, com as respectivas massa moleculares de 66,474 e 133,029 kDa. A BSA é constituída por três domínios homólogos

(I, II, III), os quais estão divididos em nove voltas (V1-V9), estão protegidas no núcleo e estabilizadas por 17 pontes dissulfuretos (Jedrychowski & Wichers, 2009; Carter and Ho, 1994), tornando a estrutura terciária da BSA relativamente estável mesmo durante a desnaturação (Jedrychowski & Wichers, 2009). Cada domínio é constituído por dois subdomínios, (IA, IB; IIA, IIB e IIIA, IIIB) (Carter and Ho, 1994; Hamdani, 2009).

À luz de dados obtidos através da cristalografia por raio X, a estrutura da albumina é constituída em 67% de hélices- $\alpha$ . Os restantes resíduos de aminoácidos formam voltas e regiões flexíveis entre os subdomínios, mas não possui folhas-B (Figura 9). Cada domínio está dividido em dez hélices- $\alpha$ . Seis para o subdomínio A e quatro para o subdomínio B (Figura 10). Tanto o domínio I e II assim como o domínio II e III, estão interligados através de extensões helicoidais, da décima hélice do domínio I (10 I), à primeira hélice do domínio II (1 II) e da décima hélice do domínio II (10 II) à primeira hélice do domínio III (1 III), criando desta forma as duas hélices mais longas da albumina (Carter and Ho, 1994).



**Figura 9:** Estrutura tridimensional da albumina bovina serosa. Representação dos três domínios e seis subdomínios. Adaptado de Stewart et al., 2003

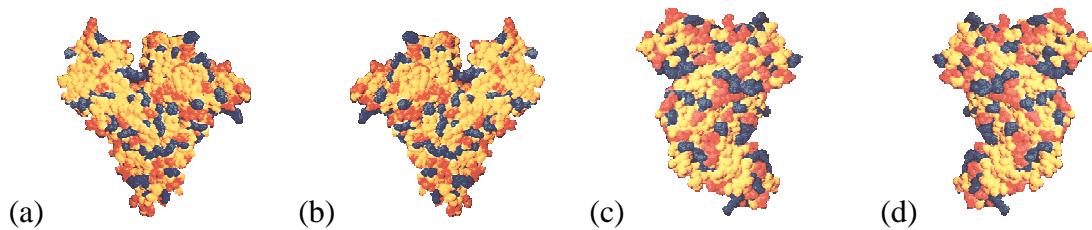


**Figura 10:** Localização das ligações de dissulfuro. (He and Carter, 1992)

Na BSA, as pontes dissulfuretos estão situadas nas seguintes posições: (1) 77-86; (2) 99-115; (3) 114-125; (4) 147-192; (5) 191-200; (6) 223-269; (7) 268-276; (8) 288-302; (9) 301-312; (10) 339-384; (11) 383-392; (12) 415-461; (13) 460-471; (14) 484-500; (15) 499-510; (16) 537-582; (17) 581-590. Os pares de pontes dissulfuretos localizam-se quase

exclusivamente entre as extensões helicoidais (Carter and Ho, 1994).

Ao Contrario da distribuição assimétrica das cargas na estrutura primária, a distribuição dessas mesmas cargas na estrutura terciária parece ser bastante uniforme (Figura 11) (Carter and Ho, 1994).

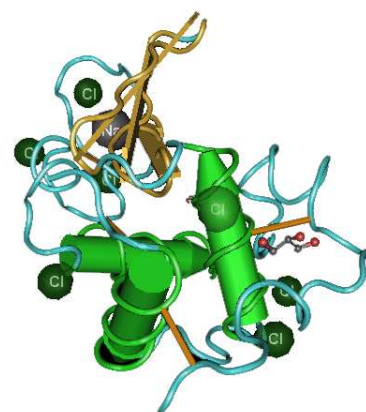


**Figura 11:** Representação da albumina sérica com os resíduos básicos coloridos a azul, os resíduos ácidos coloridos a vermelho e os resíduos neutros coloridos em amarelo. (a) vista frontal, (b) visão traseira, (c) vista lateral esquerda e (d) vista lateral direita. Adaptado de (Carter and Ho, 1994)

A albumina do soro é a proteína com maior solubilidade do sistema circulatório e tem muitas funções fisiológicas. A BSA tem sido uma das proteínas deste grupo mais estudadas, particularmente devido à sua similitude com a estrutura da albumina humana (Hamdani, 2009).

### 1.4.2. Lisozima C

A lisozima (Gal d4), é uma proteína ubíqua, presente em quase todos os organismos incluindo vírus e plantas (Figura 12). Representa apenas cerca de 3,5% das proteínas totais da clara de ovo de galinha, o que corresponde a cerca de 0,3 a 0,4g. Ela tem uma massa molecular relativa de 14 kDa e um ponto isoeléctrico (pI) de 10,7 (Jedrychowski & Wichers, 2009; Mine, 2008). Ela é vista como um alérgeno menor (Mine, 2008) e a sua alérgenicidade é rara quando comparada com a ovotransferrina (Gal d3). A lisozima desempenha um papel importante na defesa natural contra a infecção bacteriana do ovo.



**Figura 12:** Estrutura cristalina da lisozima C da clara de ovo de galinha. Adaptado de [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) (28/06/2010)

As suas propriedades anti-bacterianas são bem conhecidas mas também exibe actividades anti-virais, anti-tumorais e imunomodulatórias (Mine, 2008). Características bioquímicas como a forma da sequência da cadeia de aminoácidos ou o centro activo da lisozima, são bem conhecidos (Jedrychowski & Wichers, 2009). Com 4 ligações dissulfuretos a molécula de lisozima é constituída por dois domínios,  $\alpha$  e  $\beta$  ligados por uma longa hélice- $\alpha$  entre a qual está localizado o sítio activo. A forma da curva de estabilidade fisiológica da lisozima sugere a existência de uma única conformação dobrada devido a duas das suas quatro ligações dissulfuretos (Mine, 2008). A estabilidade da lisozima foi estudada através de vários métodos e condições diferentes e um aumento dessa estabilidade, por exemplo, contra a proteólise, pode diminuir a sua alérgenicidade (Jedrychowski & Wichers, 2009).

### 1.4.3. Ovalbumina

A ovalbumina (Gal d2), é a proteína em maior abundância na clara do ovo e a sua função permanece desconhecida. Ela é uma glicoproteína com uma massa molecular relativa de 45 kDa e constitui 54% das proteínas da clara do ovo (Jedrychowski & Wichers, 2009; Mine, 2008). Pertence à família das serpinas, se bem que ela tenha pouca actividade inibidora de proteases. A sequência de aminoácidos comporta 386 resíduos e o aminoácido terminal é acetilado. A ovalbumina não possui um N-terminal clássico que lidera a sequência de aminoácidos, apesar de ser uma proteína secretora. A sequência hidrofóbica entre os resíduos 21 e 47 podem actuar como uma sequência de sinal interno envolvida num local transmembranar. A estrutura cristalina da ovalbumina nativa de galinha, mostra um anel central reactivo intacto na forma de uma hélice- $\alpha$  exposta de três voltas que se destacam do corpo principal da molécula em duas hastes peptídicas (Jedrychowski & Wichers, 2009).

A ovalbumina pode possuir alguma actividade imunomoduladora, como foi encontrado para induzir a libertação do factor  $\alpha$  de necrose tumoral (TNF) e os péptidos imunogénicos da

ovalbumina foram usados para melhorar a resposta imunitária para a imunoterapia do cancro (Mine, 2008).

#### **1.4.4. Ovomucina**

A ovomucina representa aproximadamente 3,5% da totalidade das proteínas da clara do ovo e é uma macromolécula, altamente glicolisada, pesada (Jedrychowski & Wichers, 2009; Mine, 2008), com um teor em hidratos de carbono acima de 33%. Ela difere das restantes proteínas do ovo por ser extremamente grande. O seu peso molecular relativo apresentado na literatura (Alleoni, 2006) varia de 1800 a 8300 kDa. Contem um montante substancial de grupos dissulfuretos, éteres sulfónicos, um montante considerável de cistinas (interligadas através de ligações intermoleculares) e 50% do conteúdo de ácido ciálico total na clara do ovo (Jedrychowski & Wichers, 2009). A clara de ovo possui duas formas de ovomucina, uma subunidade solúvel e uma subunidade insolúvel (Jedrychowski & Wichers, 2009; Mine, 2008). A forma insolúvel caracteriza a grande viscosidade das finas camadas de ovomucina da clara do ovo, a qual tem actividade anti-viral. A alergenicidade da ovomucina é considerada minoritária quando comparada com outros alergénios como a ovomucóide e a ovalbumina (Jedrychowski & Wichers, 2009). Alias, vários estudos sobre alergia do ovo, apontam no sentido em que a ovomucina não é um alergénio dominante. Uma das razões pode ser que alguns processos de extracção para preparações de proteínas, inclui a diálise e a centrifugação para obter-se ovomucina livre de mucina em solução para evitar obstrução da coluna cromatográfica (Jedrychowski & Wichers, 2009).

#### **1.4.5. Ovomucóide**

A ovomucóide (Gal d1), é composta por 186 aminoácidos que constituem três domínios organizados em *tandem*, denominados (DI, DII e DIII) (Matsuda et al., 1981; Jedrychowski & Wichers, 2009; Huopalahti et al., 2007) e possuindo cada um deles 60 aminoácidos (Jedrychowski & Wichers, 2009). Ela foi identificada como um dos alergénios alimentares dominantes na clara do ovo e é a causa de muitas reacções alérgica nas crianças (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009). A ovomucóide exhibe graus variáveis de actividade inibitória relativamente a um número de serinas proteases como a tripsina por exemplo (Jedrychowski & Wichers, 2009; Matsuda et al., 1981; Feinstein, 1970). É uma proteína altamente glicolisada (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009) contendo 20%-25% de hidratos de carbono com um peso molecular de 28 kDa (Jedrychowski & Wichers, 2009 & Mine, 2008), representa 11% da proteína do clara do ovo, não coagula através do aquecimento e tem um ponto isoeléctrico (pI) de 4,1. Holen e Elsayed, 1990 e Mine, 2008 descreveram duas isoformas da ovomucóide numa gama de pI entre 4,4 e 4,6.

Tal como a avidina, a ovomucóide representa uma possível fonte de ligações peptídicas bioespecíficas, sendo possível estudar a sua incorporação nas micropartículas

poliméricas que superam a degradação de proteínas e enzimas proteolíticas. Na sua forma nativa, a ovomucóide demonstrou resistência à desnaturação a temperaturas extremas. O aquecimento da ovomucóide a 100°C tem apenas um pequeno efeito nas propriedades de ligação aos anticorpos. As propriedades antigénicas e alergénicas da ovomucóide mantêm-se activas após a digestão peptídica. Esse fenómeno, foi atribuído em parte, à presença das suas nove ligações de dissulfureto. De facto, a redução química da ovomucóide aperfeiçoa a sua digestibilidade e reduz a sua alergenicidade. Similarmente, a oxidação química da ovomucóide não altera a actividade de ligações à IgE específica de pacientes alérgicos ao ovo, enquanto que a sua forma reduzida mostra um decréscimo significativo. Em contraste, outros estudos referem que a forma reduzida da ovomucóide não afecta a sua alergenicidade ou conduz a um aumento da actividade de ligação IgE no soro de alguns pacientes (Mine, 2008).

#### 1.4.6. Caseínas

A caseína bovina é catalogada pela União Internacional das Sociedades imunológicas como sendo um único alergénio, (Bos d8). Na realidade, as caseínas são constituídas por quatro proteínas principais: a  $\alpha$ -caseína s1; a  $\alpha$ -caseína s2; a  $\beta$ -caseína e a  $\kappa$ -caseína em proporções aproximativas de 40%, 10%, 40%, 10% respectivamente (Jedrychowski & Wichers, 2009).

O centro das caseínas é hidrofóbico enquanto que a periferia é hidrofílica e contem locais de fosforilação. A  $\alpha$ -caseína s1; a  $\alpha$ -caseína s2 e a  $\beta$ -caseína, têm capacidade de quelatar o  $\text{Ca}^{2+}$ , o  $\text{Zn}^{2+}$  e o  $\text{Fe}^{3+}$  (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009). Este tipo de caseínas aumenta os níveis de cálcio no leite. A  $\alpha$ -caseína s1 e a  $\alpha$ -caseína s2, parecem ser os mais importantes alergénios, seguidos pela  $\beta$ -caseína. As caseínas mostram ter uma estrutura enrolada aleatória. De acordo com a literatura, demonstrou-se a reacção da IgE entre as caseínas do leite de vaca e as caseínas do leite de cabra e ovelha. No entanto, as caseínas do leite de égua parecem representar um baixo risco de provocar reacções mediadas pela IgE (Hoffmann-Sommergruber e Mills, 2009). Ficou provado que a desfosforilação reduz a ligação da IgE às caseínas (Jedrychowski & Wichers, 2009).

##### 1.4.6.1. $\alpha$ -caseína s1

A  $\alpha$ -caseína s1 tem um peso molecular de 27 a 32 kDa, dependendo das condições de análise. Ela é formada por dois polipéptidos de tamanhos diferentes que são constituídos pela mesma sequência de aminoácidos mas diferem no grau de fosforilação. Spuergin *at al.* (1997) identificaram um epítopo alergénico nas regiões 19-30, 93-98 e 141-150 da  $\alpha$ -caseína s1, como sendo um epítopo imunodominante. Foram reconhecidas algumas regiões de ligação às IgE nos aminoácidos 17-36, 39-48, 69-78, 93-102, 109-120, 123-132, 139-154, 159-174 e 173-194. Estudos recentes mostraram que os péptidos N-terminal e o C-terminal têm uma maior afinidade com as IgE humana (Jedrychowski & Wichers, 2009).

#### 1.4.6.2. $\alpha$ -caseína s2

A  $\alpha$ -caseína s2 é uma proteína com um peso molecular na ordem dos 27-32 kDa e é a mais hidrofílica de todas a caseína, porque possui um aglomerado de grupos aniônicos. Ela é constituída por dois componentes de maior dimensão e alguns de dimensão mais reduzida que exibem vários níveis de fosforilações pós-translacional. A  $\alpha$ -caseína s2 é fosforilada em múltiplos sítios. A sua estrutura é constituída aproximadamente de 25% hélices- $\alpha$ , contem três cisteínas e forma dímeros de ligação dissulfureto. Usando 99 decapeptídeos sintéticos, dez regiões apresentam ligações à IgE e foram identificadas como zonas alergénicas. Essas regiões situam-se nos aminoácidos: 31-44, 43-56, 83-100, 93-108, 105-114, 117-128, 143-158, 157-172, 165-188 e 191-200. Investigações sobre a presença de epítomos na  $\alpha$ -caseína s2, conduziram à identificação de alergias persistentes e temporárias. Os resíduos de aminoácidos: 33-42, 87-96 e 159-168 com ligações fracas em 145-154 e 171-180, originada de indivíduos com alergia ao leite de vaca persistente. Os pacientes com alergias temporárias mostraram apenas ligações fracas aos péptidos da  $\alpha$ -caseína s2 (Jedrychowski & Wichers, 2009).

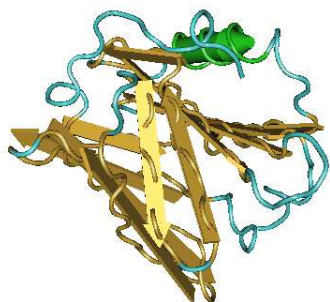
Foi sugerido que em algumas das regiões fosforiladas, os resíduos: 23-31, 73-76 e 144-146, formam parte do epítomo para o IgE e que os fosfatos estão envolvidos no reconhecimento dos anti-corpos (Bernard et al., 2000).

#### 1.4.6.3. $\beta$ -caseína

A  $\beta$ -caseína é uma proteína com um peso molecular de 26,6 kDa. Ela é o componente mais hidrofóbico da totalidade das caseínas. A  $\beta$ -caseína é menos fosforilada que a  $\alpha$ -caseína com cinco sítios potenciais de fosfoserina na parte da sequência N-terminal. Foram usados cem decapeptídeos sintéticos emparelhados, de maneira a poder averiguar quais as regiões da  $\beta$ -caseína susceptíveis de se ligarem às IgE. Os aminoácidos: 1-16, 45-54, 55-70, 83-92, 107-120, 135-144, 149-164, 167-184 e 185-208, foram descritos como característicos em pacientes com alergia persistente ao leite de vaca. Também foi observado que a resposta da IgE à  $\beta$ -caseína desfosforilada é inferior à  $\beta$ -caseína nativa. Sendo assim, a região do N-terminal desfosforilado, resíduos: 30-50, provavelmente inclui epítomos IgE importantes (Jedrychowski & Wichers, 2009).

A  $\beta$  caseína é uma das maiores proteínas do leite e uma das proteínas mais importantes envolvidas na formação e estabilização de emulsões dos alimentos. É uma molécula fortemente hidrofóbica com alta carga no N-terminal e C-terminal, contendo várias cadeias laterais hidrofóbicas e alguns grupos carregados. A pH neutro muitas proteínas com carga absoluta localizadas nos primeiros 21 resíduos da molécula, inclui quatro das cinco fosfoserinas (éster do ácido fosfórico e serina), enquanto a carga absoluta das restantes moléculas é perto de zero (Caessens *et al*, 1999).

### 1.4.7. *B*-lactoglobulina



**Figura 13:** Representação da *B*-lactoglobulina (Nativa). Adaptado de [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) (28/06/2010)

A *B*-lactoglobulina (Figura 13), é uma proteína do leite e do soro, com um peso molecular de 18 kDa e pertence à superfamília das lipocalinas. Ela é designada como sendo um alergénio nomeado (Bos d5) (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009). A *B*-lactoglobulina tem a capacidade de se ligar aos lípidos, incluindo: o retinol, a  $\beta$ -caroteno, os ácidos gordos saturados e insaturados e aos hidrocarbonetos alifáticos. Transporta moléculas hidrofóbicas as quais têm funções importantes. A *B*-lactoglobulina representa uma fracção importante da totalidade das proteínas

do soro do leite, mais de 50%. Em condições fisiológicas, a *B*-lactoglobulina existe numa mistura em equilíbrio sobre a forma de monómero e dímero. Ela possui três pontes dissulfetos e é sensível aos tratamentos térmicos. Se a *B*-lactoglobulina for aquecida entre 74°C e 90°C, reduz significativamente a ligação à IgE. Aquecendo a 90°C reduz a ligação à IgE de maneira mais significativa. No entanto, a inibição total é sempre possível com concentrações altas de alergénio aquecido, o que indica que a ligação à IgE não foi totalmente destruída (Ehn *et al.*, 2004).

Ao longo da cadeia polipeptídica da *B*-lactoglobulina, existem epítomos principais para a IgE humana. Esses epítomos são formados pelos resíduos de aminoácidos: 41-60, 102-124, os intermediários: 1-8, 25-40 e com menor influencia: 9-14, 84-91, 125-135 e 78-83 (Jedrychowski & Wichers, 2009).

### 1.4.8. *B*-2-microglobulina

A *B*-2-microglobulina ( $\beta$ 2m), é a cadeia lateral solúvel do MHC (Main Histocompatibility Complex) classe I. Ela desempenha um papel de factor de crescimento ósseo, que regula a actividade celular, tanto dos osteoblastos como dos osteoclastos (Quesada *et al.*, 1998).

### 1.4.9. Osteopontina

A osteopontina (OPN) é uma glicoproteína fosforilada (Higashikawa *et al.*, 2007; Wang e Denhardt, 2008) que é sintetizada em vários tecidos e é segregada nos fluidos corporais (Wang e Denhardt, 2008). Tem funções versáteis, incluindo um papel na reconstrução de tecidos, na fibrose, na mineralização, na imunomodulação e nas inflamações (Higashikawa *et al.*, 2007; Wang e Denhardt, 2008). A principal via utilizada pela osteopontina para influenciar o comportamento celular, é através da interacção com integrinas ou CD44 (Higashikawa *et al.*,

2007; Rodrigues *et al.*, 2007; Wang e Denhardt, 2008). As integrinas são glicoproteínas de superfície celular heterodímeras que intervêm na resposta celular para as proteínas da matriz extra-celular (Goldsby *et al.*, 2000; Higashikawa *et al.*, 2007; Rodrigues *et al.*, 2007). A osteopontina possui a sequência canónica de reconhecimento das integrinas, arginina (R) - glicina (G) - aspartato (D), (RGD), a qual é reconhecida por quatro integrinas:  $\alpha_v\beta_1$ ,  $\alpha_v\beta_3$ ,  $\alpha_v\beta_5$  e  $\alpha_5\beta_1$ . No entanto, a osteopontina pode ligar-se a outras integrinas sem ser através da sequência (RGD) (Higashikawa *et al.*, 2007; Rodrigues *et al.*, 2007; Wang e Denhardt, 2008). A osteopontina foi identificada como sendo um biomacador em vários tipos de cancros e doenças inflamatórias (Higashikawa *et al.*, 2007; Wang e Denhardt, 2008).

Geralmente, a osteopontina é extremamente hidrofílica (Rodrigues *et al.*, 2007; Wang e Denhardt, 2008) com um ponto isoeléctrico baixo (3,5) e mostra uma composição em aminoácidos inabitual com 42 serinas, 48 aspartatos e 27 glutamatos, o que constitui quase metade dos resíduos da osteopontina humana (298 resíduos) (Rodrigues *et al.*, 2007).

#### **1.4.10. Ubiquitina**

A ubiquitina (Ub) é uma pequena proteína de 76 resíduos de aminoácidos (Garzón *et al.*, 2007; Chew *et al.*, 2005) que pode ligar-se covalentemente a outras proteínas através de cascatas enzimáticas compostas por ubiquitina de activação e enzimas de conjugação (UBA ou E1 e UBC ou E2, respectivamente) e da ubiquitina ligase (E3) (Garzón *et al.*, 2007). Ela é uma proteína abundante nos eucariotas (Ent e Löwe, 2005; Chew *et al.*, 2005) e é essencialmente composta de hélices- $\alpha$  e várias voltas (Filby, 1998).

A ubiquitina é o membro mais estudado e mais entendido da família das proteínas ligadas à Ub, que inclui o tipo 1, ubiquitina como proteína (UBLs, Ubiquitin-Like) e o tipo 2, ubiquitina domínio de proteína (UDPs, Ubiquitin-Domain Proteins) (Chew *et al.*, 2005).

A ubiquitina pode ligar-se a uma proteína como um monómero ou como uma cadeia de poli-ubiquitinas, onde elas também estão ligadas através de pontes isopeptídicas. A ubiquitinização regula uma variedade de processos celulares, incluindo a endocitose; o tráfego vesicular; o controlo do ciclo celular; a resposta ao stress; a reparação do ADN (Ácido desoxirribonucleico); a sinalização e a transcrição de genes. Até agora, foram identificados dezasseis domínios de interacção com a ubiquitina e foram determinadas várias estruturas complexas (Hong *et al.*, 2009).

## 1.5. REACÇÕES ALÉRGICAS

Acredita-se que as reacções alérgicas aos alimentos resultam de uma resposta aumentada do sistema imunitário a antigénios da dieta, principalmente proteínas alimentares (Mine, 2008).

Os alimentos contêm uma grande variedade de proteínas, mas somente uma pequena fracção desses polipéptidos são alergénios (substância que provoca ou pode provocar alergia). A maioria das proteínas alimentares alergénicas tem massas moleculares entre os 10 e os 70 kDa. Contudo, foram já separadas várias proteínas de massa molecular inferior aos 10 kDa, originárias por exemplo do extracto de amendoim e de amostras de leite hidrolisado (Bayard e Lottspeich, 2000).

As alergias alimentares representam um crescente risco de saúde para as populações de países industrializados. A forma mais comum e mais estudada de alergia alimentar, também conhecida por Reacção de Hipersensibilidade Tipo I, é mediada por uma classe de anticorpos segregada pelos linfócitos B designada por Imunoglobulina E (IgE) (Mine, 2008). Pensa-se que cerca de 5-6% das crianças e mais de 2% dos adultos sofrem de algum tipo de alergia alimentar mediada pela IgE. Em geral, a reactividade da IgE e por consequência, o risco de desenvolver alergias contra proteínas alimentares, deve-se à estrutura 3D (tridimensional) dessas mesmas proteínas. Áreas específicas expostas à superfície das proteínas alergénicas, conduzem a um aumento da produção de anticorpos de IgE específicos (Mine, 2008, Hoffmann-Sommergruber e Mills, 2009). Muitos dos alergénios existem como isoalergénios, ou seja, proteínas constituídas por, pelo menos, 67% de sequências de aminoácidos idênticas e com diferentes graus de glicosilação. Os alergénios alimentares são geralmente glicoproteínas com um ponto isoeléctrico ácido (Bayard e Lottspeich, 2000). Foram identificados em várias proteínas, diversos resíduos de aminoácidos essenciais para o reconhecimento dos anticorpos (Hoffmann-Sommergruber e Mills, 2009).

### 1.5.1. Epítomos

As células imunológicas não interagem ou reconhecem uma molécula antigénica completa. Em vez disso, os linfócitos reconhecem sítios discretos na macromolécula denominados de epítomos ou determinantes antigénicos. Os epítomos são regiões imunologicamente activas de um antigénio onde se ligam os receptores de membrana específica do antigénio nos linfócitos ou nos anticorpos segregados (Goldsby *et al.*, 2000).

Os estudos com antigénios pequenos revelaram que as células T e B reconheceram epítomos diferentes numa mesma molécula antigénica. Os linfócitos com um antigénio complexo podem interagir a vários níveis com a estrutura do antigénio. Um epítomo num antigénio proteico pode envolver elementos da estrutura primária, secundária e terciária ou até mesmo quaternária da proteína. Nos polissacarídeos, uma ramificação extensa de cadeia lateral por ligações glicosídicas afecta a conformação dos epítomos (Goldsby *et al.*, 2000).

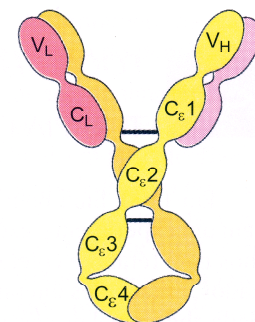
O reconhecimento dos antígenos pelas células B e T é fundamentalmente diferente. As células B reconhecem o antígeno solúvel quando este se liga ao seu anticorpo ligado à membrana. Como as células B se ligam ao antígeno livre na solução, os epítopos que elas reconhecem tendem a ser sítios altamente acessíveis na superfície exposta de um antígeno. A maioria das células T reconhece somente os péptidos combinados com as moléculas do MHC (Main Histocompatibility Complex) que estão na superfície das células de antígeno e nas células próprias alteradas. Como a maioria das células T reconhece o antígeno apenas quando ele está combinado com uma molécula do MHC, os epítopos da célula T, como regra, não podem ser considerados fora das suas moléculas do MHC associadas (Goldsby *et al.*, 2000).

### 1.5.2. *Imunoglobulinas (Igs)*

#### 1.5.2.1. IgE

A potente actividade biológica da IgE (Figura 14), permite-lhe ser identificada no soro, apesar da sua concentração média ser extremamente baixa (0,3µg/ml). Os anticorpos IgE medeiam as reacções de hipersensibilidade imediata que são responsáveis pelos sintomas da febre do feno; asma; urticária e choque anafilático. A presença do componente sérico responsável pelas reacções alérgicas foi pela primeira vez demonstrada em 1921 por K. Prausnitz e H. Kustner, que injectaram dermicamente o soro de uma pessoa alérgica num indivíduo não alérgico. Quando o antígeno apropriado foi injectado posteriormente no mesmo local, desenvolveu-se uma reacção de vermelhão e inchaço (análogo e urticária). Esta reacção, denominada de **reacção P-K**, foi a base para o primeiro ensaio biológico para a actividade da IgE (Goldsby *et al.*, 2000).

A IgE liga-se aos receptores Fc das membranas dos basófilos sanguíneos e dos mastócitos tissulares. A ligação cruzada das moléculas de IgE, ligadas ao receptor pelo antígeno (alergénio), induz a degranulação dos basófilos e dos mastócitos; como resultado, é libertada uma variedade de mediadores farmacologicamente activos dos grânulos, causando as manifestações alérgicas. A degranulação dos mastócitos induzidos pelas IgE pode também libertar mediadores que facilitam o recrutamento de várias células necessárias para a defesa antiparasitas (Goldsby *et al.*, 2000).

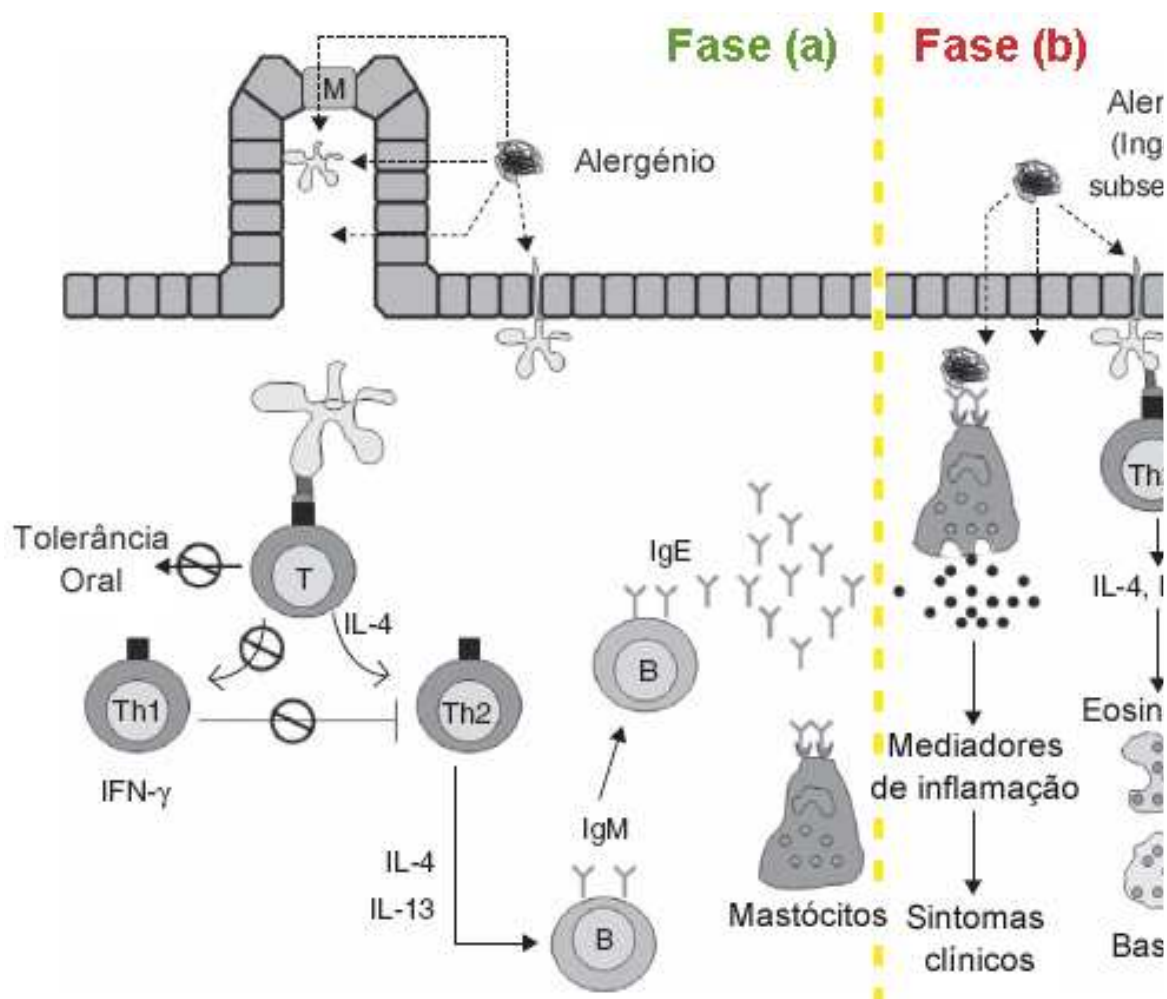


**Figura 14:** Estrutura geral da IgE. Representação das cadeias leves (rosa) e das cadeias pesadas (amarelo). As pontes dissulfuretos estão representadas por linhas pretas finas. (Adaptado de Goldsby *et al.*, 2000)

### 1.5.3. *Mecanismos celulares e moleculares da Alergia Alimentar*

A etiologia da Reacção de Hipersensibilidade Tipo I envolve duas fases: a) fase de sensibilização, sem sintomas, durante a qual o sistema imunitário está activado, resultando na

produção de anticorpos pertencentes à imunoglobulina de classe E e b) fase de estimulação, com respostas clínicas específicas como os sintomas alérgicos, as quais ocorrem após a subsequente exposição aos compostos alérgicos. Durante o início da reacção alérgica, demonstrou-se que os antígenos podem atravessar a barreira intestinal (Figura 15), serem capturados pelas células imunológicas basilares, serem processados e apresentados pelas células de apresentação de antígenos (células dendríticas) (Mine, 2008).



**Figura 15:** Mecanismo molecular e celular da alergia alimentar. As letras B e T representam respectivamente os linfócitos B e T. Th1 e Th2 representam respectivamente as células T helper 1 e 2. As diferentes interleucinas estão representadas através de (IL-4 e IL-13). IFN-γ representa o interferon gama (Interferon-γ). IgE e IgM representam respectivamente as imunoglobulinas E e M. A letra (M) representa uma célula ao acaso da membrana gastrointestinal. (Adaptado de Mine, 2008)

Usualmente é aceite que a resposta imunitária resulte de um balanço entre dois subtipos de linfócitos T, os CD4+ (cluster of differentiation 4) T helper 1 (Th1) e linfócitos T helper 2 (Th2) (Mosmann *et al.*, 1986; Mosmann e Coffman 1989). Em geral, as células Th1 desempenham um papel importante na imunidade mediada por células com a secreção de citocinas tais como a interferon γ (INF-γ), interleucina 2 (IL-2) e factor de necrose tumoral β (TNF-β), os quais promovem o desenvolvimento de linfócitos T citotóxicos e macrófagos

(Untersmayr e Jensen-Jarolim, 2006). Por outro lado, as células Th2 promovem respostas imunitárias humorais primárias, com a expressão de citocinas como a IL-4 e IL-13 (promovendo a síntese de IgE), IL-5 (proliferação eosinófila) e IL-9 (activação de mastócitos). O padrão dominante de citocinas que conduz à formação de células Th2, tem sido associado à resposta imunitária alérgica. A relação Th1/Th2 apresenta uma característica importante e curiosa, pois tanto um tipo celular como o outro têm capacidade de se regularem mutuamente. Os níveis e tipos de citocinas encontrados pelas células T helper “naive” têm um papel importante em determinar se ela se irá desenvolver numa célula Th1 ou numa Th2. Citocinas como a IL-4 promovem a diferenciação das células T helper “naive” em fenótipos Th2, enquanto que a IL-12 e a INF- $\gamma$  são importantes no desenvolvimento das células Th1 (Mine, 2008).

Em indivíduos saudáveis, a ingestão de antígenos inócuos leva a um estado de tolerância oral, enquanto que indivíduos alérgicos irão desenvolver, na mucosa gastrointestinal, um microambiente de citocinas onde a IL-4 é dominante. Quando os antígenos alimentares são apresentados às células T “naive”, através das células apresentadoras de antígeno, irão ser geradas células Th2 específicas desses antígenos. A activação das células Th2 irá conduzir à produção de IL-4 e IL-3, as quais promovem a produção de IgE pelas células B. A IgE específica dos alérgenos irá então ligar-se aos seus receptores de alta afinidade, presentes na superfície dos mastócitos. Numa ingestão subsequente do alérgeno alimentar, a apresentação do antígeno irá levar à activação rápida das células Th2, seguida pela activação dos eosinófilos e basófilos. Entretanto, os fragmentos alérgicos também podem interagir com IgE ligadas a receptores de mastócitos, desencadeando a agregação de receptores e a subsequente libertação de mediadores de inflamação, tais como as histaminas, as quais são directamente responsáveis pelos sintomas clínicos das alergias alimentares (Mine, 2008).

## 1.6. MÉTODOS DE ESTUDO

As principais características de qualquer proteína são: o peso molecular, a composição em aminoácidos e as cargas. Todas estas características podem ser calculadas com exactidão através da sequência de aminoácidos e todos estes cálculos podem ser efectuados, usando um programa de cálculo (Filby, 1998), tal como o Microsoft Excel.

Em quase todos os métodos descritos aqui, a proteína ou a sequência de aminoácidos a ser analisada, são inseridas na primeira coluna da primeira folha de cálculos, a qual é sempre etiquetada como “Cálculos” (Filby, 1998).

### 1.6.1. As escalas de hidrofobicidade

Uma característica que determina a estrutura e a estabilidade das proteínas é a interacção de aminoácidos hidrofóbicos. Estes aminoácidos tendem a associar-se no interior da proteína e em regiões da proteína que atravessam as membranas celulares. Muitas proteínas importantes, possuem um ou mais domínios que atravessam a membrana celular e a identificação de cada região tem implicações para a topologia e a função da proteína em causa. São necessários cerca de vinte aminoácidos para atravessar uma membrana celular típica. Os aminoácidos contidos nessa porção da sequência, são principalmente aminoácidos hidrofóbicos e são frequentemente regiões abrangentes da membrana. Tais regiões, são relativamente fáceis de delimitar com um programa de computador, porque elas são ricas em aminoácidos hidrofóbicos e porque as suas vizinhanças são ainda mais ricas em aminoácidos carregados e neutros.

Existe uma vasta variedade de valores para medir o grau de hidrofobicidade dos aminoácidos. Estes valores são baseados numa compilação de propriedades típicas dos aminoácidos incluindo a solubilidade em água em comparação à sua solubilidade em solventes orgânicos, e a frequência à qual eles são encontrados no interior das membranas ou em contacto com outros aminoácidos hidrofóbicos em proteínas conhecidas. Como não existe nenhum método padronizado de efectuar estes cálculos, vários autores propuseram diferentes valores de escala de hidrofobicidade (Tabela 5) (Filby, 1998). A escala mais utilizada é a de *Kyte & Doolittle* (Kyte e Doolittle, 1982), ou ainda a de *Sweet*, também conhecida como OMH (*Optimized Matching Hydrophobicity*) (Sweet e Eisenberg., 1983). A escala de *Kyte & Doolittle* é fundamentalmente parecida à escala de *Rao & Argos* (Rao e Argos, 1986) porque se baseia numa compilação de propriedades de solubilidade e estrutura dos diferentes aminoácidos e foi especificamente desenvolvida para identificar as regiões transmembranares das proteínas.

Os valores dos métodos de *Hopp & Woods* (Hopp e Woods, 1981) e *Welling & al.* (Welling *et al.*, 1985) são baseados em tentativas de prever regiões prováveis de se localizarem na superfície da proteína e conseqüentemente locais onde anticorpos (alergenicidade) e outras proteínas possam ligar-se. Em ambos os casos, os picos representam tais regiões. Como essas regiões são geralmente muito pequenas uma dimensão razoável de resíduos de aminoácidos será de cerca de 7 aminoácidos. Existem outras duas escalas comumente utilizadas, que são: a escala de *Eisenberg et al.*, (Eisenberg *et al.*, 1984) uma escala harmonizada e a de *Black* (Black e Mould, 1991), a mais recente. Ao escolher uma ou outra dessas escalas, cada valor é atribuído ao respectivo aminoácido. Os aminoácidos carregados têm um valor negativo, enquanto que os aminoácidos fortemente hidrofóbicos têm valores positivos (Filby, 1998). As escalas de hidrofobicidade reflectem a preferência das cadeias laterais dos aminoácidos por meios polares e/ou apolares e desta forma atribuir um parâmetro numérico a cada resíduo. Os valores de hidrofobicidade reflectem as determinações de energia livre de transferência de um aminoácido de um solvente apolar para um polar. Ao atribuir valores numéricos a cada

resíduo de uma sequência e seguidamente compara-los com uma amostragem de tamanho predefinido, estabelece-se o perfil hidrofóbico (Whitford, 2005).

**Tabela 5:** Representação de várias escalas de hidrofobicidade

Aminoácidos	Kyte & Doolittle	Hopp & Wood	Sweet	Welling	Eisenberg	Black
A	1,8	-0,5	-0,400	1,150	0,620	0,616
C	2,5	-1	0,170	-1,200	0,290	0,680
D	-3,5	3	-1,310	0,650	-0,900	0,028
E	-3,5	3	-1,220	-0,710	-0,740	0,043
F	2,8	-2,5	1,920	-1,410	1,190	1,000
G	-0,4	0	-0,670	-1,840	0,480	0,501
H	-3,2	-0,5	-0,640	3,120	-0,400	0,165
I	4,5	-1,8	1,250	-2,920	1,380	0,943
K	-3,9	3	-0,670	2,060	-1,500	0,283
L	3,8	-1,8	1,220	0,750	1,060	0,943
M	1,9	-1,3	1,020	-3,850	0,640	0,738
N	-3,5	0,2	-0,920	-0,770	-0,780	0,236
P	-1,6	0	-0,490	-0,530	0,120	0,711
Q	-3,5	0,2	-0,910	-0,110	-0,850	0,251
R	-4,5	3	-0,590	0,580	-2,530	0,000
S	-0,8	0,3	-0,550	-0,260	-0,180	0,359
T	-0,7	-0,4	-0,280	-0,450	-0,050	0,450
V	4,2	-1,5	0,910	-0,130	1,080	0,825
W	-0,9	-3,4	0,500	-1,140	0,810	0,878
Y	-1,3	-2,3	1,670	0,130	0,260	0,880

Quando estes valores estão distribuídos num intervalo de vinte aminoácidos, no gráfico aparece distintamente um pico, se houver muitos resíduos hidrofóbicos ou caso contrário, uma depressão (Filby, 1998).

A tabela 6, resume todos os elementos necessários para calcular o parâmetro  $H_7$  que corresponde à média dos valores de hidrofobicidade de 7 aminoácidos consecutivos; o  $H_\alpha$  (índice de hélices  $\alpha$ ) e o  $H_{Ag}$ , índice de antigenicidade. Todos estes cálculos baseiam-se em duas escalas de hidrofobicidade: a de *Kyte & Doolittle* serve para calcular  $H_7$  e  $H_\alpha$ , enquanto que a escala de *Hopp & Wood* serve para calcular  $H_{Ag}$  (Han e Tashjian, 2003). A mesma tabela representa apenas os vinte primeiros resíduos da BSA.

A coluna **A** representa a posição de cada aminoácido ao longo da cadeia. A coluna **B** representa o tipo de aminoácido de cada posição. As colunas **C** e **D**, representam o valor da hidrofobicidade de cada escala para cada tipo de aminoácido. A coluna **C** mostra a escala de *Kyte & Doolittle* e a coluna **D**, a escala de *Hopp & Wood* (Han e Tashjian, 2003).

**Tabela 6:** Representação da tabela de cálculos da hidrofobicidade para  $H_7$ ;  $H_\alpha$  e  $H_{Ag}$ , de uma amostra dos 20 primeiros aminoácidos da BSA. HYKD (escala hidrofóbica de *Kyte & Doolittle*); HYHW (escala hidrofóbica de *Hopp & Wood*).

A	B	C	D	E	F	G
Posição	Sequência de AA	HYKD	HYHW	$H_7$	$H_\alpha$	$H_{Ag}$
1	M	1,9	-1,3			-1,0
2	K	-3,9	3,0			-1,1
3	W	-0,9	-3,4			-1,6
4	V	4,2	-1,5	1,1		-1,3
5	T	-0,7	-0,4	0,7		-1,3
6	F	2,8	-2,5	1,8		-1,6
7	I	4,5	-1,8	2,5		-1,5
8	S	-0,8	0,3	2,5		-1,6
9	L	3,8	-1,8	3,1	1,2	-1,6
10	L	3,8	-1,8	3,1	0,8	-1,2
11	L	3,8	-1,8	2,3	0,9	-1,0
12	L	3,8	-1,8	2,3	1,0	-1,1
13	F	2,8	-2,5	2,1	1,7	-0,7
14	S	-0,8	0,3	1,3	1,9	0,2
15	S	-0,8	0,3	0,7	0,4	0,1
16	A	1,8	-0,5	-0,5	-0,1	-0,2
17	Y	-1,3	-2,3	-1,0	0,7	-0,5
18	S	-0,8	0,3	-0,3	0,4	0,4
19	R	-4,5	3,0	0,3	-0,5	0,8
20	G	-0,4	0,0	-0,6	-0,9	0,8
...	...	...	...	...	...	...

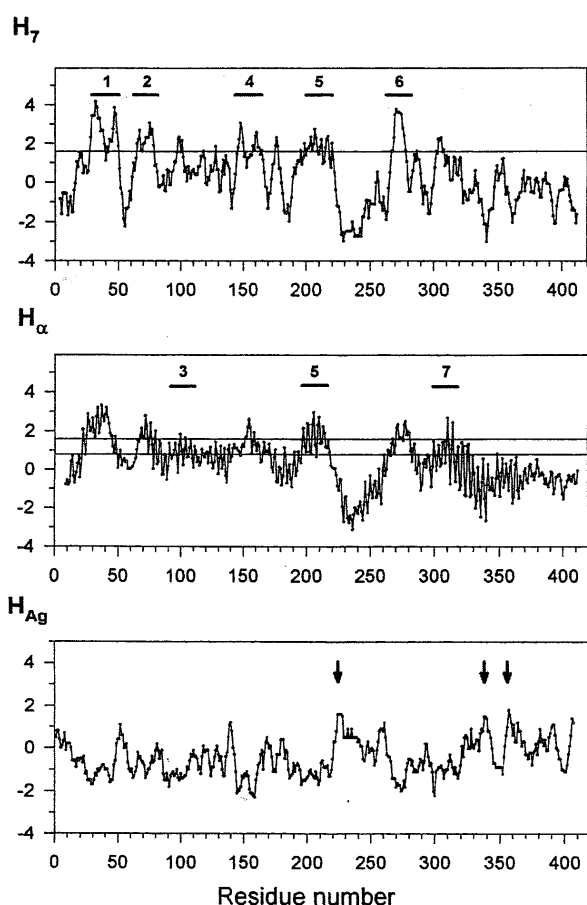
Os valores da coluna E, começam a partir do quarto resíduo e terminam quatro resíduos antes do fim da cadeia total, porque cada um destes valores, representa a média aritmética de sete aminoácidos consecutivos. Por exemplo: o primeiro valor da coluna E é a média dos valores que vão da posição 1 à posição 7 da coluna C. O segundo valor da coluna E é a média dos valores que vão da posição 2 à posição 8 da coluna C e assim sucessivamente até ao fim (Han e Tashjian, 2003).

Os valores da coluna F, começam a partir do nono resíduo e terminam nove resíduos antes do fim da cadeia total. A obtenção de cada um deles requer um cálculo particular (Equação 1.1), mas que pode traduzir-se por uma média ponderada que implica dezassete aminoácidos consecutivos (Han e Tashjian, 2003).

$$= \frac{C1+C17+C2+C16+\frac{(C4+C14)}{4}+C5+C13+\frac{(C6+C12)\times 3}{4}+\frac{(C8+C10)}{2}+C9}{10} \quad 1.1$$

A coluna **G** representa os valores da antigenicidade. Eles começam com o primeiro resíduo mas terminam seis resíduos antes do final da sequência de aminoácidos. O cálculo de cada valor consiste na média aritmética dos valores de hidrofobicidade de seis resíduos consecutivos, na escala de *Hopp & Wood*. O primeiro valor da coluna **G** é a média dos valores que vão da posição 1 à posição 6 da coluna **D**. O segundo valor da coluna **G** é a média dos valores que vão da posição 2 à posição 7 da coluna **D** e assim sucessivamente até ao fim (Han e Tashjian, 2003).

A obtenção de todos estes valores, (colunas **E**, **F** e **G**), permite elaborar gráficos (Figura 16) nos quais é possível prever o comportamento das proteínas em solventes como a água ( $H_7$ ), a existência de hélices  $\alpha$  ( $H_\alpha$ ) ou ainda o comportamento antigénico das proteínas ( $H_{Ag}$ ).



**Figura 16:** Análise do receptor TRH do rato para o  $H_7$ , o  $H_\alpha$  e o  $H_{Ag}$ . Adaptado de: Han e Tashjian, 2003

susceptível de gerar este tipo de gráficos a partir de uma matriz de dados como a tabela 6.

### 1.6.2. Composição da proteína

O uso das folhas de cálculo, permite visualizar a composição relativa em cada aminoácido e a contagem do número total de cada tipo de aminoácidos existentes na

Neste exemplo concreto, observa-se a análise dos três parâmetros ( $H_7$ ,  $H_\alpha$  e  $H_{Ag}$ ), para a TRH (*Thyrotropin-releasing hormone*, hormona tirotropina) do rato que tem um comprimento de 412 resíduos de aminoácidos. No  $H_7$  e  $H_\alpha$ , valores acima do valor da linha horizontal 1,6, são considerados hidrofóbicos. No  $H_\alpha$ , valores abaixo do valor da linha horizontal 0,8, são considerados hidrofílicos. No  $H_{Ag}$ , as regiões com valores elevados são zonas antigénicas plausíveis. Os traços nos gráficos de  $H_7$  e  $H_\alpha$  indicam uma previsão de hélices transmembranares. As setas no gráfico do  $H_{Ag}$  prevêm regiões de alta antigenicidade (Han e Tashjian, 2003).

Para as apresentações gráficas, pode utilizar-se qualquer programa (Microsoft<sup>®</sup> Excel, SPSS ou outros),

proteína. Esta informação é designada de “composição em aminoácidos” e é muito útil porque estes vinte valores (dos 20 aminoácidos que compõem as proteínas), são totalmente específicos a cada proteína (Filby, 1998).

A tabela 7 ajuda a visualizar e entender melhor este método. A coluna **A** representa a sequência (neste caso, incompleta) de aminoácidos da proteína em estudo. A coluna **B** representa o código dos vinte aminoácidos principais que compõem habitualmente as proteínas. A coluna **C** representa a quantidade real de cada aminoácido que compõem a proteína em estudo. A obtenção destes valores é feita através de uma expressão (Equação 1.2) matemática adequada ao programa de cálculo. Esta expressão significa que o programa vai contar todos os aminoácidos correspondentes ao “código”, ao logo da sequência de aminoácidos que no caso da BSA, são 607. Os números da expressão correspondem à linha da folha de cálculo onde se encontra o aminoácido e não à posição deste na cadeia polipeptídica. A segunda linha corresponde ao primeiro aminoácido, a terceira corresponde ao segundo e assim por ai adiante.

$$=CONTAR.SE(\$A\$2:\$A\$608;"Código") \quad 1.2$$

A coluna **D** corresponde à percentagem de cada aminoácido existente na proteína em estudo. Os valores são obtidos através da equação 1.3.

$$\frac{(Aminoácido\ em\ causa) \times 100}{Número\ total\ de\ AA\ da\ proteína} \quad 1.3$$

A coluna **E** corresponde à percentagem de cada aminoácido encontrado numa proteína de comparação. A coluna **F** representa a percentagem relativa de cada aminoácido presente na proteína em estudo. Esta relatividade refere-se à percentagem de cada aminoácido que a proteína (Coluna **D**) tem, relativamente à percentagem desse mesmo aminoácido na proteína de comparação (Coluna **E**), (Equação 1.4).

$$\frac{Valores\ da\ coluna\ D}{Valores\ da\ coluna\ E} \quad 1.4$$

A coluna **G** representa a média da percentagem relativa das proteínas em geral. Estipulou-se que essa média te um valor unitário (1) (Filby, 1998), seja qual for o aminoácido.

**Tabela 7:** Representação da composição real e relativa da BSA em aminoácidos, assim como a percentagem de cada um deles. Nesta tabela também está representada a média real e relativa dos aminoácidos das proteínas em geral.

A	B	C	D	E	F	G
Proteína em estudo				Outras proteínas	Relativa	
Seq. De AA	Código	Qd. De AA	% de AA	Média de AA de uma proteína de comparação	% de AA	Média de AA
M	A	48	7,91	7,26	1,089221	1
K	C	35	5,77	1,76	3,276172	1
W	D	40	6,59	5,18	1,272159	1
V	E	59	9,72	6,22	1,562690	1
T	F	30	4,94	4,06	1,217325	1
F	G	17	2,80	6,81	0,411257	1
I	H	17	2,80	2,23	1,255901	1
S	I	15	2,47	5,75	0,429769	1
L	K	60	9,88	5,88	1,681068	1
L	L	65	10,71	9,24	1,158918	1
L	M	5	0,82	2,33	0,353529	1
L	N	14	2,31	4,68	0,492826	1
F	P	28	4,61	4,98	0,926275	1
S	Q	20	3,29	4,04	0,815568	1
S	R	26	4,28	5,17	0,828503	1
A	S	32	5,27	7,39	0,713373	1
Y	T	34	5,60	5,94	0,942983	1
S	V	38	6,26	6,41	0,976645	1
R	W	3	0,49	1,32	0,374420	1
G	Y	21	3,46	3,26	1,061239	1
...						

A obtenção e recolha destes dados, são frequentemente usados na construção de gráficos em radar. Nesse tipo de gráficos, o tamanho dos picos está proporcionalmente ligado à percentagem e quantidade de cada aminoácido. Cada gráfico fornece uma visão global da composição em aminoácidos e são muito úteis para comparar o perfil de duas proteínas diferentes. Os aminoácidos são distribuídos no sentido horário e por ordem alfabética correspondente ao código de letra unitária de cada aminoácido. No entanto, pode ser utilizada qualquer outra ordem de distribuição (Filby, 1998).

O problema com os gráficos em radar, é que alguns aminoácidos são mais abundantes dentro da proteína que outros. Isto significa que a percentagem e/ou a quantidade dos aminoácidos presentes na proteína, irá desenhar um gráfico totalmente assimétrico. A

interpretação desses gráficos não é muito óbvia aos olhos de quem é inexperiente na matéria. É por essa razão que se compara os gráficos obtidos com gráficos de proteínas conhecidas (Filby, 1998).

A elaboração de gráficos tendo em vista as propriedades químicas dos aminoácidos como a solubilidade ou outro tipo de propriedades, é muito simples de fazer. Por exemplo: ao colocar os aminoácidos mais carregados à direita e os mais hidrofóbicos à esquerda, vai gerar um gráfico em radar no qual será mais fácil observar o comportamento global da proteína (Filby, 1998).

### **1.6.3. Determinação das cargas**

Em condições fisiológicas, alguns aminoácidos estão carregados positivamente ou negativamente. O montante e a localização dessas cargas, são características importantes da proteína e a concentração das cargas positivas ou negativas também estão associadas com tipos particulares de propriedades de ligação, por exemplo: são frequentemente encontradas regiões básicas na ligação de proteínas ao ADN (Filby, 1998).

Sob condições ligeiramente ácidas, apenas cinco aminoácidos contribuem para a carga total das proteínas. Os aminoácidos básicos, tais como a lisina e a arginina, têm carga perto de (+1). O aminoácido meio básico como a histidina, tem uma carga cerca de (+0,5). E os aminoácidos ácidos tais como o glutamato e o aspartato, têm uma carga perto de (-1). Sendo assim, coloca-se estes valores na folha de cálculo, à frente dos respectivos aminoácidos de maneira a poder calcular-se as cargas das proteínas. Frente aos restantes aminoácidos, coloca-se o valor de carga (0) (Filby, 1998).

A tabela 8, ilustra perfeitamente o método de cálculo das cargas das proteínas. A coluna A representa (na vertical) a sequência de aminoácidos da proteína em estudo. As colunas B, C, D, E e F, representam os valores da carga de cada um desses aminoácidos, quando ele está presente na cadeia polipeptídica, (coluna A). A coluna G, representa os valores do somatório de B a F. A coluna H, representa os mesmos valores que G, só que em absoluto (sem valores negativos). Cada célula da coluna I, representa a soma das cargas reais (G) de 28 resíduos de aminoácidos. Por exemplo: a primeira célula da coluna I, é a soma do resíduo 1 até ao resíduo 28 da coluna G. O mesmo acontece com a coluna J. A única diferença é que a soma é calculada na coluna H. Os cálculos continuam, deslocando-se cada vez de 1 aminoácido: de 2-29, de 3-30, etc...

Tabela 8: Representação dos cálculos de uma amostra de 28 aminoácidos da  $\alpha$ -caseína s1.

A	B	C	D	E	F	G	H	I	J
Seqüência de AA	Aspartato	Glutamato	Lisina	Arginina	Histidina	Somatório	Carga Abs.	Cargas num intervalo de 28 AA	
M	0	0	0	0	0	0	0	5	5
K	0	0	1	0	0	1	1	4	6
L	0	0	0	0	0	0	0	3	5
L	0	0	0	0	0	0	0	3	5
I	0	0	0	0	0	0	0	3	5
L	0	0	0	0	0	0	0	2	6
T	0	0	0	0	0	0	0	2	6
C	0	0	0	0	0	0	0	2	6
L	0	0	0	0	0	0	0	2	6
V	0	0	0	0	0	0	0	3	7
A	0	0	0	0	0	0	0	3	7
V	0	0	0	0	0	0	0	3	7
A	0	0	0	0	0	0	0	3	7
L	0	0	0	0	0	0	0	3	7
A	0	0	0	0	0	0	0	3	7
R	0	0	0	1	0	1	1	3	7
P	0	0	0	0	0	0	0	2	6
K	0	0	1	0	0	1	1	1	7
H	0	0	0	0	0,5	0,5	0,5	0	6
P	0	0	0	0	0	0	0	-0,5	5,5
I	0	0	0	0	0	0	0	-0,5	5,5
K	0	0	1	0	0	1	1	0,5	6,5
H	0	0	0	0	0,5	0,5	0,5	-1,5	6,5
Q	0	0	0	0	0	0	0	-1	7
G	0	0	0	0	0	0	0	-1	7
L	0	0	0	0	0	0	0	-1	7
P	0	0	0	0	0	0	0	-2	8
Q	0	0	0	0	0	0	0	-2	8

### 1.6.4. Método de Chou & Fasman

O maior problema em biologia molecular, é a previsão da estrutura tridimensional proteica com base na sua sequência de aminoácidos. Actualmente é muito mais fácil obter essa mesma sequência de uma proteína do que a sua estrutura. As proteínas contêm apenas alguns tipos de estrutura fundamentais; estas são: a hélice- $\alpha$ ; a folha- $\beta$  e a volta. Cada uma tem um conteúdo e arranjo característico de aminoácidos, e existem vários métodos computacionais destinados a identificar este tipo de sequências. Um método de prever uma estrutura proteica foi estabelecido por *Chou e Fasman* (Filby, 1998) e uma versão funcional do mesmo pode ser implantado no Excel. Esta versão é uma das grandes variedades que podem ser construídas na base do algoritmo de *Chou e Fasman*, que não está claramente definido na literatura (Filby, 1998).

Existem dois factores de probabilidades das voltas, chamados PT (Probability Turn) e  $P_t$  (probability turn). Apesar de terem o mesmo nome, estes factores são calculados de maneira distinta. PT é a média dos valores de quatro aminoácidos consecutivos calculados a partir de J4, (Tabela 10). O factor ( $P_t$ ) é o produto das probabilidades calculadas nas colunas X a AA, (Tabela 9), através dos mesmos quatro aminoácidos (Filby, 1998).

**Tabela 9:** Matriz de probabilidades através da qual se aplica o algoritmo de Chou & Fasman para a elaboração da tabela 9. Esta matriz é idêntica para todas as proteínas.

S	T	U	V	W	X	Y	Z	AA
Aminoácidos	Posição	Hélice	Folha	Turn	f i	f i+1	f i+2	f i+3
A	0	1,42	0,83	0,66	0,06	0,076	0,035	0,058
C	1	0,7	1,19	1,19	0,159	0,053	0,117	0,128
D	2	1,01	0,54	1,46	0,147	0,11	0,179	0,081
E	3	1,51	0,37	0,74	0,056	0,06	0,077	0,064
F	4	1,13	1,38	0,6	0,059	0,041	0,065	0,065
G	5	0,57	0,75	1,56	0,102	0,085	0,19	0,152
H	6	1	0,87	0,95	0,14	0,047	0,093	0,054
I	7	1,08	1,6	0,47	0,043	0,034	0,013	0,056
K	8	1,16	0,74	1,01	0,055	0,115	0,072	0,095
L	9	1,21	1,3	0,59	0,061	0,025	0,036	0,07
M	10	1,45	1,05	0,6	0,068	0,082	0,014	0,055
N	11	0,67	0,89	0,98	0,161	0,083	0,191	0,091
P	12	0,57	0,55	1,52	0,102	0,301	0,034	0,068
Q	13	1,11	1,1	0,98	0,074	0,098	0,037	0,098
R	14	0,98	0,93	0,95	0,7	0,106	0,099	0,085
S	15	0,77	0,75	1,43	0,12	0,139	0,125	0,106
T	16	0,83	1,19	0,96	0,086	0,108	0,065	0,079
V	17	0,7	1,7	0,5	0,062	0,048	0,028	0,053
W	18	1,08	1,37	0,96	0,077	0,013	0,064	0,167

Y	19	0,69	1,47	1,14	0,082	0,065	0,114	0,125
---	----	------	------	------	-------	-------	-------	-------

A existência de uma hélice- $\alpha$  numa determinada região da proteína só é possível se o valor da probabilidade num intervalo de sete aminoácidos for superior a 1,05. O mesmo acontece com as folhas- $\beta$ . Esta determinação é efectuada pela equação (1.5) em O6 (Tabela 10).

$$= SE(E(K6 > L6; K6 > 1,05) = VERDADEIRO; 2; 1) \quad 1.5$$

A função (E) (Equação 1.5) dá uma resposta de (VERDADEIRO) apenas se o valor de hélice- $\alpha$  for superior ao valor de folha- $\beta$  ( $K6 > L6$ ) e o valor de hélice- $\alpha$  for superior a 1.05 ( $K6 > 1.05$ ) (Tabela 10). Se (E) for (VERDADEIRO), então a célula da folha de cálculo toma o valor 2, caso contrário ela toma o valor 1. Os números 1 e 2 são escolhidos para produzir um gráfico claro e sem significados intrínsecos. A determinação de probabilidade das folhas- $\beta$  é parecida à das hélices- $\alpha$ , em que a probabilidade das folhas- $\beta$  deve ser superior à probabilidade das hélices- $\alpha$  e também deve ser superior a 1.05. A equação (1.6) na célula P6 (Tabela 10) devolve um valor igual a 4 se a região completar estes critérios e 3 se tal não acontecer. Mais uma vez os números 4 e 3 são escolhidos por razões gráficas.

$$= SE(E(L6 > 1,05; L6 > K6) = VERDADEIRO; 4; 3) \quad 1.6$$

A previsão das voltas é mais complexa. As voltas são previstas se o valor de  $PT > 1$ , se for superior a ambas as probabilidades hélice- $\alpha$  e folha- $\beta$  e se o valor de  $P_t > 7.5 \times 10^{-5}$ . Os valores  $PT$  (Tabela 10, coluna M), são calculados da mesma forma que as hélices e folhas. Os valores de  $P_t$  (Tabela 10, coluna N) são também utilizados e baseados na tabela que enumera o quanto frequente determinados aminoácidos são encontrados em voltas em proteínas conhecidas. Estes valores são incluídos nas colunas de X a AA da matriz de probabilidades (Tabela 9). Estes valores referem-se à frequência com que um aminoácido ocorre na primeira; segunda; terceira e quarta posição numa volta ( $f_i$ ,  $f_{i+1}$ ,  $f_{i+2}$  e  $f_{i+3}$  respectivamente) (Filby, 1998).

Para se perceber bem este método, é preciso ter em conta que a elaboração da tabela 10 é feita com base na tabela 9, onde consta as probabilidades de encontrar os diversos aminoácidos nas diferentes estruturas, (hélices- $\alpha$ , folhas- $\beta$  e voltas). Por essa razão, todas as equações necessárias para o cálculo dos diferentes valores das diversas colunas, utilizam valores da tabela 9. É por isso muito importante ter-se em conta a letra que designa cada coluna em ambas as tabelas.

A coluna F representa a sequência de aminoácidos da proteína. A coluna G representa a posição do respectivo aminoácido na matriz de probabilidades (Tabela 9, coluna T).

**Tabela 10:** Representação da folha de cálculos a partir da qual se constrói o gráfico de Chou & Fasman para as diferentes estruturas: hélice- $\alpha$ , folha- $\beta$  e volta. Amostra dos primeiros 9 aminoácido da BSA.

F	G	H	I	J	K	L	M	N	O	P	Q
Seq AA	No AA	Hélice	Folha	Volta	Hélice	Folha	Volta (PT)	Volta (pt)	Hélice	Folha	Volta
W	18	1,08	1,37	0,96							
V	17	0,7	1,7	0,5							
T	16	0,83	1,19	0,96	0,97142 9	1,32714 3	0,755	1,56E -05	1	4	5
F	4	1,13	1,38	0,6	0,99	1,31714 3	0,632 5	2,44E -05	1	4	5
I	7	1,08	1,6	0,47	1,06285 7	1,26	0,865	4,86E -06	1	4	5
S	15	0,77	0,75	1,43	1,11714 3	1,27571 4	0,772 5	1,76E -05	1	4	5
L	9	1,21	1,3	0,59	1,11714 3	1,27571 4	0,77	1,51E -05	1	4	5
L	9	1,21	1,3	0,59	1,07285 7	1,15428 6	0,8	7,56E -06	1	4	5
L	9	1,21	1,3	0,59	1,07285 7	1,15428 6	0,59	3,84E -06	1	4	5
...	...	...	...	...	...	...	...	...	...	...	...

De maneira sucinta, a tabela 11 representa as diferentes funções de programação dos cálculos para as restantes colunas da tabela 10.

**Tabela 11:** Resumo das diferentes funções que permitem calcular os valores das diferentes colunas da tabela 9.

Coluna	Função
H	=DESLOCAMENTO(\$U\$4;\$G4;0;1;1)
I	=DESLOCAMENTO(\$V\$4;\$G4;0;1;1)
J	=DESLOCAMENTO(\$W\$4;\$G4;0;1;1)
K	=MÉDIA(H4:H10)
L	=MÉDIA(I4:I10)
M	=MÉDIA(J4:J7)
N	=PRODUTO(DESLOCAMENTO(\$X\$4;G4;0;1;1);DESLOCAMENTO(\$X\$4;G5;1;1;1);DESLOCAMENTO(\$X\$4;G6;2;1;1);DESLOCAMENTO(\$X\$4;G7;3;1;1))
O	=SE(E(K6>L6;K6>1,05)=VERDADEIRO;2;1)
P	=SE(E(L6>1,05;L6>K6)=VERDADEIRO;4;3)
Q	=SE(E(M6>K6;M6>L6;M6>1;N6>0,000075)=VERDADEIRO;6;5)

Os dados são apresentados graficamente. Os picos representam uma previsão para a estrutura relevante, enquanto que a falta dos mesmos, prevê a ausência dessa mesma estrutura (Filby, 1998).

### 1.6.5. Estruturas super-enroladas “Coiled-coil”

Um exemplo ainda mais complexo, é o uso do Excel para prever regiões prováveis de formar “coiled coils” (super-hélices), (Tabelas 13; 14). O algoritmo foi descrito por Lupas e os seus colaboradores e é um pouco complexo, (Tabela 15) (Filby, 1998). A super-hélice- $\alpha$  “coiled coil”, é um arranjo periódico de aminoácidos hidrofóbicos e carregados, em “hepteno” (Filby, 1998; Bujnicki, 2008). Os aminoácidos num segmento de estrutura super-hélice são rotulados como a, b, c, d, e, f e g (Filby, 1998). As posições a e d geralmente contêm aminoácidos hidrofóbicos de grandes dimensões, frequentemente leucina (Filby, 1998; Bujnicki, 2008), enquanto que os restantes aminoácidos geralmente possuem carga. As posições a e d de duas hélice- $\alpha$  super-enrolada, interagem e encontram-se geralmente no interior da estrutura. Lupas e os seus colaboradores determinaram a frequência com que cada aminoácido é encontrado nas sete posições possíveis em proteínas conhecidas. Estes dados, que estão reproduzidos na Tabela 12, mostram a frequência relativa em que cada aminoácido é encontrado numa super-hélice comparando-a com a esperada numa outra proteína qualquer. Existem sete diferentes valores para cada aminoácido dependendo em que posição pode ser encontrada numa super-hélice, (Tabela 13; colunas I, J, K, L, M, N, O). O próximo aminoácido também obtém sete valores, mas tal valor deve ser alterado por uma posição relativa ao anterior aminoácido, visto que o próximo aminoácido é mais posterior na possível estrutura super-enrolada (Tabela 13; coluna H). O valor para a posição “a” para um aminoácido N deve estar ligada ao da posição “b” para o aminoácido N+1 e por assim adiante. É usado um intervalo de vinte e oito aminoácidos visto que esse parece ser o comprimento mínimo para as super-hélices na natureza. Os sete conjuntos de valores para cada possível péptido de 28 aminoácidos são multiplicados em conjunto para determinar a raiz vigésima oitava (Equação 1.7).

$$\sqrt[28]{(\text{Produto dos 28 aminoácidos em cada posição})} \quad 1.7$$

Os valores mais elevadas de entre os sete possíveis para cada um dos péptidos de vinte e oito aminoácidos é convertido numa probabilidade baseada em valores obtidas com moléculas da super-hélices conhecida e esta probabilidade é representada graficamente (Tabela 14; coluna Y). Este programa tem sido muito eficaz e prevê super-hélices- $\alpha$  em várias moléculas. Em vários casos estas previsões têm sido comprovadas (Filby, 1998).

**Tabela 12:** Frequência com a qual, cada aminoácido é encontrado em cada uma das sete posições possíveis em proteínas super-enroladas conhecidas. Adaptado de Filby, 1998.

<b>AA</b>	<b>AB</b>	<b>AC</b>	<b>AD</b>	<b>AE</b>	<b>AF</b>	<b>AG</b>	<b>AH</b>	<b>AI</b>
<b>Aminoácidos</b>	<b>Posição</b>	<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>e</b>	<b>f</b>	<b>g</b>
A	0	1,297	1,551	1,084	2,612	0,377	1,284	0,877
C	1	0,824	0,022	0,308	0,152	0,18	0,156	0,044
D	2	0,03	2,352	2,268	0,237	0,663	1,62	1,448
E	3	0,262	3,496	3,108	0,998	5,685	2,494	3,048
F	4	0,531	0,076	0,403	0,662	0,189	0,106	0,013
G	5	0,045	0,275	0,578	0,216	0,211	0,426	0,156
H	6	0,347	0,275	0,679	0,395	0,294	0,579	0,213
I	7	2,597	0,098	0,345	0,894	0,514	0,471	0,431
K	8	1,375	2,639	1,763	0,191	1,815	1,961	2,795
L	9	3,167	0,297	0,398	3,902	0,585	0,501	0,483
M	10	2,24	0,37	0,48	1,409	0,541	0,772	0,663
N	11	0,835	1,475	1,534	0,039	1,722	2,456	2,28
P	12	0	0,008	0	0,013	0	0	0
Q	13	0,179	2,114	1,778	0,631	2,55	1,578	2,526
R	14	0,659	1,163	1,21	0,031	1,358	1,937	1,798
S	15	0,382	0,583	1,052	0,419	0,525	0,916	0,628
T	16	0,169	0,702	0,955	0,654	0,791	0,843	0,647
V	17	1,665	0,403	0,386	0,949	0,211	0,342	0,36
W	18	0,24	0	0	0,456	0,019	0	0
Y	19	1,417	0,09	0,122	1,659	0,19	0,13	0,155

De um ponto de vista mais claro, os cálculos efectuados para cada posição (a, b, c, d, e, f e g) e para cada aminoácido da proteína, são elaborados na tabela 13 (Referências 1, 2, 3, 4, 5, 6 e 7) tendo por base a matriz de frequências relativas, (Tabela 12) e aplicando as funções do Microsoft® Excel, (Tabela 15). Por sua vez, os valores obtidos servem de base para novos cálculos, (Tabela 14). O valor de cada célula da tabela 14, da coluna P à coluna V, é a multiplicação de 28 valores de aminoácidos consecutivos, elevado a  $(1/28)$  ou então  $(^{28}\sqrt{\quad})$ , (Equação 1.7).

**Tabela 13:** Representação da folha de cálculos a partir da qual se calcula os valores das 7 referências relativas às 7 posições (a, b, c, d, e, f e g) e para cada aminoácido da cadeia polipeptídica. Amostra dos primeiros 31 aminoácidos.

<b>F</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>J</b>	<b>K</b>	<b>L</b>	<b>M</b>	<b>N</b>	<b>O</b>
Seq. AA	Qtd. AA	Índice	Ref. 1	Ref. 2	Ref. 3	Ref. 4	Ref. 5	Ref. 6	Ref. 7
M									
K	18	1	0	0	0,456	0,019	0	0	0,24
W	17	2	0,386	0,949	0,211	0,342	0,36	1,665	0,403
V	16	3	0,654	0,791	0,843	0,647	0,169	0,702	0,955
T	4	4	0,189	0,106	0,013	0,531	0,076	0,403	0,662
F	7	5	0,471	0,431	2,597	0,098	0,345	0,894	0,514
I	15	6	0,628	0,382	0,583	1,052	0,419	0,525	0,916
S	9	7	3,167	0,297	0,398	3,902	0,585	0,501	0,483
L	9	8	0,297	0,398	3,902	0,585	0,501	0,483	3,167
L	9	9	0,398	3,902	0,585	0,501	0,483	3,167	0,297
L	9	10	3,902	0,585	0,501	0,483	3,167	0,297	0,398
L	4	11	0,189	0,106	0,013	0,531	0,076	0,403	0,662
F	15	12	0,916	0,628	0,382	0,583	1,052	0,419	0,525
S	15	13	0,628	0,382	0,583	1,052	0,419	0,525	0,916
S	0	14	1,297	1,551	1,084	2,612	0,377	1,284	0,877
A	19	15	0,09	0,122	1,659	0,19	0,13	0,155	1,417
Y	15	16	1,052	0,419	0,525	0,916	0,628	0,382	0,583
S	14	17	0,031	1,358	1,937	1,798	0,659	1,163	1,21
R	5	18	0,211	0,426	0,156	0,045	0,275	0,578	0,216
G	17	19	0,342	0,36	1,665	0,403	0,386	0,949	0,211
V	4	20	0,013	0,531	0,076	0,403	0,662	0,189	0,106
F	14	21	0,659	1,163	1,21	0,031	1,358	1,937	1,798
R	14	22	1,163	1,21	0,031	1,358	1,937	1,798	0,659
R	2	23	2,268	0,237	0,663	1,62	1,448	0,03	2,352
D	16	24	0,654	0,791	0,843	0,647	0,169	0,702	0,955
T	6	25	0,294	0,579	0,213	0,347	0,275	0,679	0,395
H	8	26	1,961	2,795	1,375	2,639	1,763	0,191	1,815
K	15	27	0,628	0,382	0,583	1,052	0,419	0,525	0,916
S	3	28	0,262	3,496	3,108	0,998	5,685	2,494	3,048
E	7	29	0,098	0,345	0,894	0,514	0,471	0,431	2,597
I	0	30	1,084	2,612	0,377	1,284	0,877	1,297	1,551

Por exemplo: o primeiro valor da coluna **P** da tabela 14, é a multiplicação de 28 valores da coluna **I** da tabela 13, que vão desde o índice 1 ao índice 28 (coluna **H**). O produto dessa multiplicação é de seguida elevado a  $(1/28)$ . O mesmo acontece com os valores que vão desde o índice 2 ao índice 29 e assim de seguida para todos os aminoácidos e todos as colunas de **P** a **V** da tabela 14.

**Tabela 14:** Representação da folha de cálculos a partir da qual se calcula os valores que servirão à elaboração do gráfico das estruturas super-enroladas “Coiled coil”. Amostra dos primeiros 30 aminoácidos da BSA.

<b>P</b>	<b>Q</b>	<b>R</b>	<b>S</b>	<b>T</b>	<b>U</b>	<b>V</b>	<b>W</b>	<b>X</b>	<b>Y</b>
Reslt. 1	Reslt. 2	Reslt. 3	Reslt. 4	Reslt. 5	Reslt. 6	Reslt. 7	Máx	Prob · S.E.	Gráf ·
0	0	0,47419 7	0,51883 7	0	0	0,69483 9	0,69483 9		0
0,43677 8	0,57121 5	0,48573 6	0,58368 8	0,50531 8	0,57835 4	0,75652 2	0,75652 2		0
0,45318 6	0,59224 8	0,49590 9	0,61192 8	0,52164 5	0,57321 8	0,79382 6	0,79382 6		0
0,44509 8	0,57168	0,4893	0,58812 2	0,53522 2	0,55435	0,78421 5	0,78421 5		0
0,47757 7	0,63418 8	0,58349 1	0,59267 5	0,58999 3	0,57655	0,70299 6	0,70299 6		0
0,45280 4	0,55964 6	0,55133 3	0,58731 8	0,59327 7	0,57039 7	0,67832 1	0,67832 1		0
0,47760 4	0,58584	0,58188 1	0,59824 9	0,57686 2	0,59623 4	0,69701 4	0,69701 4		0
0,40438 9	0,63077 5	0,61919 2	0,54129 5	0,57944 7	0,62175 5	0,72488 8	0,72488 8		0
0,40438 9	0,63077 5	0,61919 2	0,54129 5	0,57944 7	0,62175 5	0,72488 8	0,72488 8		0
0,40981 4	0,56883 7	0,59704 7	0,53816 9	0,55652 4	0,53412 1	0,72289 9	0,72289 9		0
0,39033 6	0,61696 2	0,63227 2	0,57476 8	0,50913	0,58328 8	0,77795 7	0,77795 7		0
0,44079 2	0,69062 7	0,76833 2	0,56044 9	0,58373 2	0,62743 4	0,78944 6	0,78944 6		0
0,43363	0,66446 6	0,7657	0,54560 8	0,57467 5	0,62611 3	0,77326 7	0,77326 7		0
0,37755 1	0,67232 8	0,71196 2	0,52722 8	0,58414	0,60368	0,71594 5	0,71594 5		0
0,37833 9	0,68521 2	0,72443 6	0,48020 8	0,61786 4	0,61287 9	0,74620 4	0,74620 4		0
0,39373 7	0,72435 6	0,67356 4	0,48200 9	0,64461 8	0,61302	0,65970 9	0,72435 6		0
0,38030	0,78444	0,67617	0,47173	0,63860	0,66112	0,64400	0,78444		0

3	4	3	3	2	2	8	4		
0,42973 2	0,73397 8	0,63556 7	0,4454	0,66009 5	0,63656 5	0,61825 8	0,73397 8		0
0,44567 1	0,73824 1	0,66174 5	0,51847 7	0,66191 2	0,62813 9	0,68557 7	0,73824 1		0
0,45079 5	0,74300 2	0,67233 4	0,49294 5	0,65926 3	0,62680 1	0,70772 8	0,74300 2		0
0,52396 5	0,76708 2	0,74879 7	0,51067 6	0,69238 6	0,64245	0,77366 8	0,77366 8		0
0,51993 9	0,69587 1	0,71996 4	0,56967 6	0,64529 9	0,57912 8	0,64877 7	0,71996 4		0
0,50727 3	0,69240 2	0,79013	0,55066 4	0,62826 9	0,55777 5	0,63626 4	0,79013		0
0,50288 2	0,71704 6	0,82907 2	0,55014 8	0,64088	0,59451 6	0,63384 5	0,82907 2		0
0,51988 1	0,68143 5	0,77552 6	0,52277 7	0,69144 6	0,55246 3	0,58893 5	0,77552 6		0
0,53281 4	0,67792 3	0,79853 8	0,56573 5	0,69334 9	0,54202 3	0,63913 4	0,79853 8		0
0,52869 5	0,67547 7	0,74245 8	0,56127 1	0,69355 9	0,56565 8	0,64694 2	0,74245 8		0
0,55564	0,65743 6	0,77741 3	0,57189	0,70377 5	0,59850 4	0,65963 2	0,77741 3	0	0
0,57885	0,54858 1	0,71580 9	0,53471 7	0,62213 1	0,54209 5	0,56697 9	0,71580 9	0	0
0,52930 3	0	0,61542 6	0	0	0	0	0,61542 6	0	0

A coluna **W**, representa os valores máximos de cada linha de **P** a **V**. A coluna **X** representa a probabilidade de existir estruturas super-enroladas. Os seus valores iniciam a partir do vigésimo oitavo aminoácido e estão sujeitos a oito condições.

1. Se o valor da coluna **W** for  $>1,55$  então assume o valor 1
2. Se o valor da coluna **W** for  $>1,48$  então assume o valor 0,9
3. Se o valor da coluna **W** for  $>1,4$  então assume o valor 0,8
4. Se o valor da coluna **W** for  $>1,36666$  então assume o valor 0,6
5. Se o valor da coluna **W** for  $>1,33333$  então assume o valor 0,4
6. Se o valor da coluna **W** for  $>1,22$  então assume o valor 0,2
7. Se o valor da coluna **W** for  $>1$  então assume o valor 0,1
8. Se o valor da coluna **W** não obedecer a estas condições então assume o valor 0

Cada valor da coluna X tem de obedecer a pelo menos uma das sete primeiras condições. Caso contrário, eles assumem sempre o valor (0). A coluna Y representa os dados que irão servir para a elaboração do gráfico. Cada valor desta coluna, indica o valor máximo da coluna X num intervalo de 28 aminoácidos visto que é o comprimento mínimo para as super-hélices (Filby, 1998).

De maneira sucinta, a tabela 15 representa as diferentes funções de programação dos cálculos para as colunas da tabela 13 e 14.

**Tabela 15:** Resumo das diferentes funções que permitem calcular os valores das diferentes colunas da tabela 12 e 13.

Coluna	Função
G	=PROC(\$F4;\$AA\$4:\$AA\$23;\$AB\$4:\$AB\$23)
I	=DESLOCAMENTO(\$AC\$4;\$G3;RESTO(\$H3;7);1;1)
J	=DESLOCAMENTO(\$AC\$4;\$G3;RESTO(\$H3+1;7);1;1)
K	=DESLOCAMENTO(\$AC\$4;\$G3;RESTO(\$H3+2;7);1;1)
L	=DESLOCAMENTO(\$AC\$4;\$G3;RESTO(\$H3+3;7);1;1)
M	=DESLOCAMENTO(\$AC\$4;\$G3;RESTO(\$H3+4;7);1;1)
N	=DESLOCAMENTO(\$AC\$4;\$G3;RESTO(\$H3+5;7);1;1)
O	=DESLOCAMENTO(\$AC\$4;\$G3;RESTO(\$H3+6;7);1;1)
P	=POTÊNCIA(PRODUTO(I3:I30);1/28)
Q	=POTÊNCIA(PRODUTO(J3:J30);1/28)
R	=POTÊNCIA(PRODUTO(K3:K30);1/28)
S	=POTÊNCIA(PRODUTO(L3:L30);1/28)
T	=POTÊNCIA(PRODUTO(M3:M30);1/28)
U	=POTÊNCIA(PRODUTO(N3:N30);1/28)
V	=POTÊNCIA(PRODUTO(O3:O30);1/28)
W	=MÁXIMO(P3:V3)
X	=SE(W3>1,55;1;SE(W3>1,48;0,9;SE(W3>1,4;0,8;SE(W3>1,36666;0,6;SE(W3>1,33333;0,4;SE(W3>1,22;0,2;SE(W3>1;0,1;0)))))))))
Y	=MÁXIMO(X3:X30)

## 2. OBJECTIVO

---

---

Este estudo tem como objectivo verificar se a partir das propriedades físico-químicas dos aminoácidos da estrutura primária das proteínas é possível prever propriedades das proteínas como a hidrofobicidade e alergenicidade.

Por outro lado, pretende-se também apresentar previsões sobre a estrutura secundária das mesmas proteínas utilizando diferentes métodos.

Dada a importância da hidrofobicidade para as propriedades das proteínas, serão também determinadas as zonas hidrofóbicas tendo em conta diferentes escalas de hidrofobicidade para os aminoácidos.

## 3. MATERIAIS E MÉTODOS

### 3.1. MATERIAIS

#### 3.1.1. Material informático

Para este estudo, foi utilizado um computador PC com as seguintes características e programas:

Tabela 16: Características do equipamento e softwares utilizados.

<b>Computador</b>	
<b>Marca</b>	Asus
<b>Modelo</b>	M6Ne
<b>Processador</b>	Intel Pentium M, 1.70GHz
<b>Memória RAM</b>	2.00Gb
<b>Programa de base</b>	
<b>Sistema operativo</b>	Microsoft Windows XP
<b>Edição</b>	Home Edition
<b>Versão</b>	2002, Service Pack 3
<b>Programas de cálculos e tratamento de dados</b>	
<b>Cálculos</b>	Microsoft Office Excel, Edição 2007
<b>Tratamento gráfico</b>	
<b>Tratamento de dados</b>	Microsoft Office Word, Edição 2007

#### 3.1.2. Sequências das proteínas

Para a realização deste estudo foram utilizadas oito proteínas alimentares, sendo quatro do ovo de galinha (lisozima C; ovalbumina; ovomucina e ovomucóide) e outras quatro do leite de vaca ( $\alpha$ -caseína s1;  $\alpha$ -caseína s2;  $\beta$ -caseína e  $\beta$ -lactoglobulina); três presentes no organismo humano ( $\beta$ -2-microglobulina; osteopontina e ubiquitina) e ainda a albumina do soro de vaca (BSA) (Tabela 17). O tratamento de dados das mesmas só foi possível graças às sequências de aminoácidos de cada proteína disponíveis em banco de dados na internet ([www.uniprot.org](http://www.uniprot.org)).

Tabela 17: Proteínas em estudo, da clara do ovo, do leite, do organismo humano e BSA.

<b>Proteínas do ovo</b>		<b>Código de referência (UniProt)</b> [www.uniprot.org]
<b>Proteínas da clara de ovo</b>	Lisozima C	P00698
	Ovalbumina	P01012
	Ovomucina	Q98U19
	Ovomucóide	P01005
<b>Proteínas do leite</b>		
<b>Proteínas presentes no leite</b>	A-caseína s1	P02662
	A-caseína s2	P02663
	B-caseína	P02666
	B-lactoglobulin	P02754
<b>Proteína de “referência”</b>		
<b>Albumina do Soro Bovino</b>	ASB (BSA)	P02769
<b>Proteínas presentes no organismo humano</b>		
<b>Proteínas do organismo</b>	Osteopontina	P31096
	B-2-microglobulina	P01888
	Ubiquitina	P62988

### 3.2. PROTOCOLO

1. Utilizam-se folhas de cálculo do “Microsoft Excel” descritas na literatura (Filby, 1998; Han e Tashjian, 2003) para cada proteína em estudo. As diferentes folhas de cálculo servem para determinar a hidrofobicidade; a alergenicidade; os pontos isoeléctricos; as cargas; a detecção de estruturas super-enroladas da hélice  $\alpha$  e o prognóstico de Chou e Fasman na detecção das hélices  $\alpha$  e das folhas  $\beta$  para cada proteína em estudo.
2. Procura da sequência de aminoácidos de cada proteína, na base de dados “UniProt” (www.uniprot.org/).
3. No espaço “Query”, colocar o nome da proteína desejada e de seguida carregar em “search”.

4. Na lista de proteínas que aparecem com o mesmo nome, seleccionar aquela que se deseja para o estudo.
5. Copiar a sequência de aminoácidos da proteína escolhida para o “Microsoft Word”.
6. Em *Localizar e substituir* escolher *substituir* e de seguida: *Mais* → *Especial* → *Espaço em Branco* → *Especial* → *Marca de parágrafo* → *Substituir tudo*.
7. Deve obter-se a sequência de aminoácidos sob a forma de uma listagem.
8. Eliminar os espaços em branco restantes.
9. Copiar a sequência de aminoácidos sem os espaços em branco e colá-la numa célula da folha de cálculo previamente programada.
10. Elaborar gráficos.
11. Repetir os passos 2 a 10 para cada proteína em estudo.

## 4. REGISTO E DISCUSSÃO DOS RESULTADOS

### 4.1. HIDROFOBICIDADE E ALERGENICIDADE

Nas folhas de cálculos para cada proteína, foram usadas cinco escalas de hidrofobicidade para averiguar sobre a existência de diferentes resultados em virtude das diferentes escalas (Tabela 18). Isto significa que na elaboração dos dois primeiros gráficos de cada proteína (Anexo I), estão representadas essas mesmas cinco escalas, a saber: a escala de Kyte & Doolittle (Kyte e Doolittle, 1982); de Sweet *et al.* (Sweet e Eisenberg., 1983), de Welling *et al.* (Welling *et al.*, 1985), de Eisenberg *et al.* (Eisenberg *et al.*, 1984) e a escala de Black (Black e Mould, 1991). Contudo para se escolher a escala mais adequada procedeu-se à aplicação das diferentes escalas a um conjunto de 150 aminoácidos da BSA (Gráficos 1 a 8), no estudo de  $H_7$  e  $H_\alpha$ . Deste modo, em cada ponto da sequência primária o valor de  $H_7$  corresponde à média do valor de hidrofobicidade de sete aminoácidos vizinhos usando qualquer uma das cinco escalas de hidrofobicidade, obtendo-se valores representados por  $H_7$ .  $H_\alpha$  representa o índice de hélice- $\alpha$  com base nas mesmas escalas de hidrofobicidade e através da equação 4.1, sendo  $C1$  o primeiro resíduo de aminoácido da proteína. Esta equação permite obter os gráficos de  $H_\alpha$ , onde  $C$  representa a célula onde está registado o valor da hidrofobicidade do aminoácido correspondente.

As bases do método estão apresentadas em 4.1.

$$H_\alpha = \frac{C1+C17+C2+C16+\frac{(C4+C14)}{4}+C5+C13+\frac{(C6+C12)\times 3}{4}+\frac{(C8+C10)}{2}+C9}{10} \quad 4.1$$

Nos gráficos de hidrofobicidade  $H_7$  de cada proteína (Gráficos: 9; 16; 25; 34; 43; 46; 47; 58; 67; 76; 85; 94; 103 e 112), está representado o comportamento hidrofóbico da média de sete aminoácidos consecutivos. Segundo Bomie Han *et al.*, (2003), são consideradas hidrofóbicas todas as zonas em que o valor de  $H_7$  é superior a 1,6 (Recta vermelha). Por outro lado, são consideradas hidrofílicas todas as zonas em que o valor de  $H_\alpha$  é inferior a 0,8 (Recta verde), como se pode observar no segundo gráfico das diferentes proteínas (Gráficos: 10; 17; 26; 35; 44; 48; 49; 59; 68; 77; 86; 95; 104 e 113).

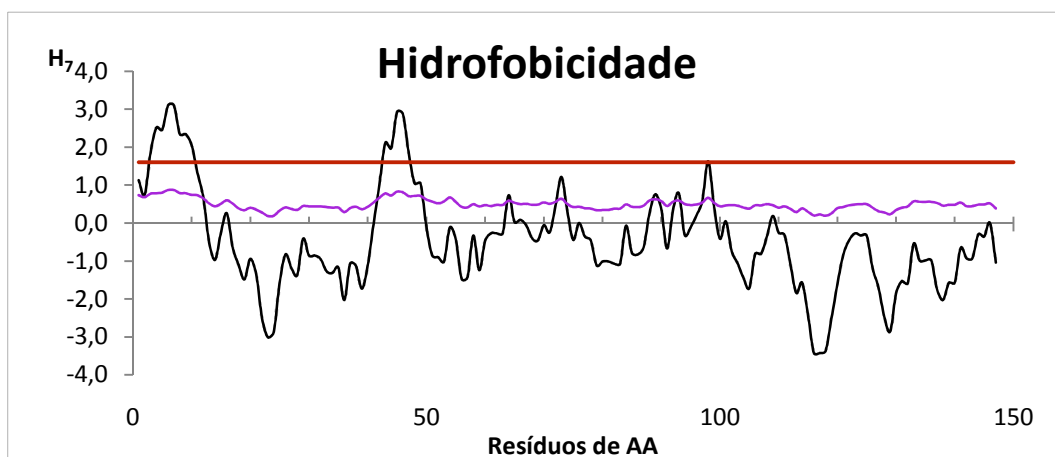
No terceiro gráfico de cada proteína (Gráficos: 11; 18; 27; 36; 45; 50; 51; 60; 69; 78; 87; 96; 105 e 114), está representado a antigenicidade das mesmas. Os cálculos foram realizados a partir da escala de hidrofiliicidade de Hopp & Woods (Tabela 18) (Han e Tashjian, 2003), correspondendo cada pico na escala positiva a uma previsão de região de elevada antigenicidade.

Tabela 18: Escalas de hidrofobicidade utilizadas no trabalho.

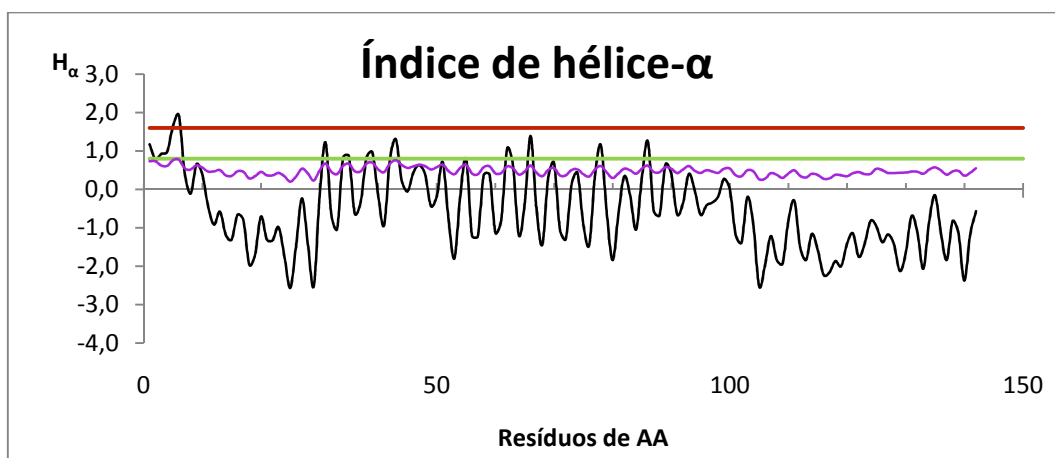
Aminoácidos	Código dos aminoácidos	Kyte & Doolittle	Hopp & Woods	Sweet	Welling	Eisenberg	Black
Arginina	R	-4,5	3,0	-0,590	0,580	-2,530	0,000
Aspartato	D	-3,5	3,0	-1,310	0,650	-0,900	0,028
Glutamato	E	-3,5	3,0	-1,220	-0,710	-0,740	0,043
Lisina	K	-3,9	3,0	-0,670	2,060	-1,500	0,283
Serina	S	-0,8	0,3	-0,550	-0,260	-0,180	0,359
Asparagina	N	-3,5	0,2	-0,920	-0,770	-0,780	0,236
Glutamina	Q	-3,5	0,2	-0,910	-0,110	-0,850	0,251
Glicina	G	-0,4	0,0	-0,670	-1,840	0,480	0,501
Prolina	P	-1,6	0,0	-0,490	-0,530	0,120	0,711
Treonina	T	-0,7	-0,4	-0,280	-0,450	-0,050	0,450
Alanina	A	1,8	-0,5	-0,400	1,150	0,620	0,616
Histidina	H	-3,2	-0,5	-0,640	3,120	-0,400	0,165
Cisteína	C	2,5	-1,0	0,170	-1,200	0,290	0,680
Metionina	M	1,9	-1,3	1,020	-3,850	0,640	0,738
Valina	V	4,2	-1,5	0,910	-0,130	1,080	0,825
Isoleucina	I	4,5	-1,8	1,250	-2,920	1,380	0,943
Leucina	L	3,8	-1,8	1,220	0,750	1,060	0,943
Tirosina	Y	-1,3	-2,3	1,670	0,130	0,260	0,880
Fenilalanina	F	2,8	-2,5	1,920	-1,410	1,190	1,000
Triptofano	W	-0,9	-3,4	0,500	-1,140	0,810	0,878

## 4.2. ESCOLHA DA ESCALAS DE HIDROFOBICIDADE

Ao observar-se as curvas obtidas com as diferentes escalas de hidrofobicidade, tanto para  $H_7$  como para  $H_\alpha$ , rapidamente se distingue que seja qual for a proteína, a escala que apresenta uma maior amplitude nas suas variações é a escala de Kyte & Doolittle (Han e Tashjian, 2003). Num sentido totalmente oposto, a escala de Black não revela grandes diferenças entre regiões com diferentes aminoácidos (Gráficos: 1 e 2).

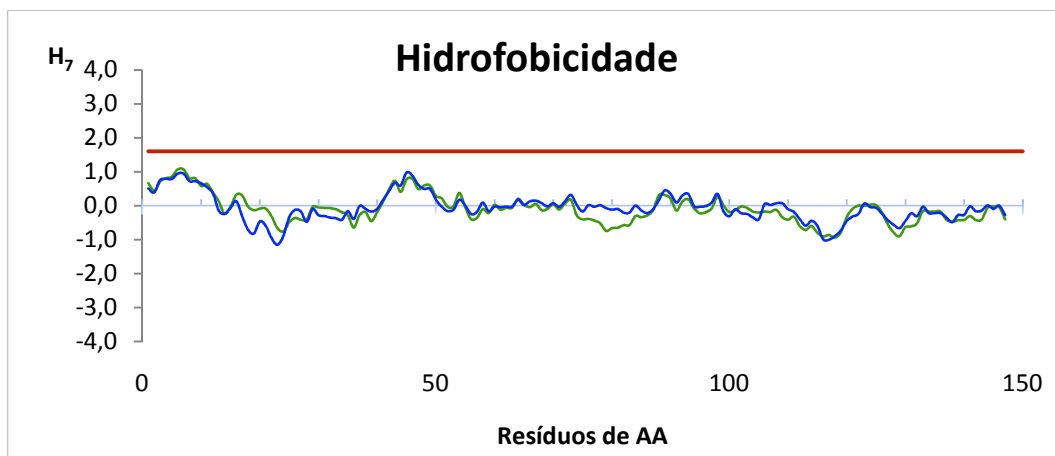


**Gráfico 1:** Representação gráfica das curvas de hidrofobicidade  $H_7$  para a escala de Kyte & Doolittle (preto) e a escala de Black (roxo) na BSA, (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).

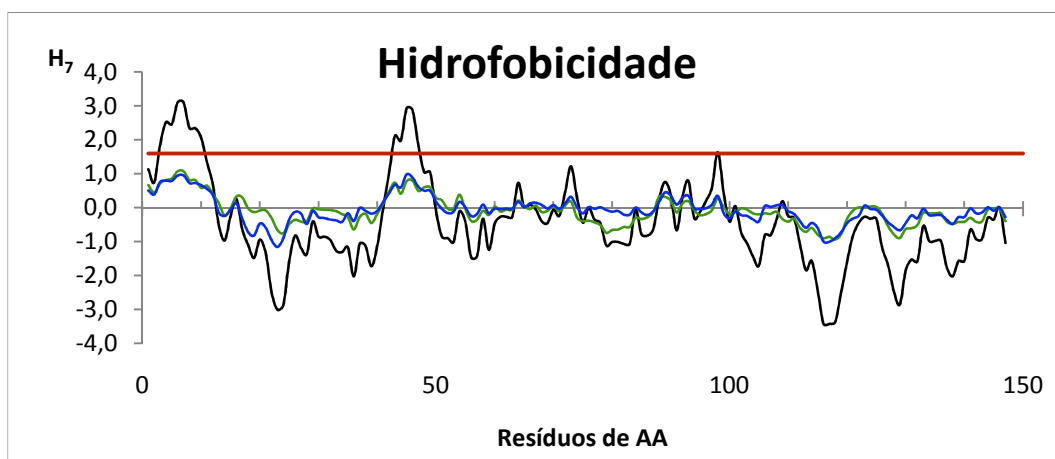


**Gráfico 2:** Representação gráfica das curvas de índice de hélice- $\alpha$  ( $H_\alpha$ ) para a escala de Kyte & Doolittle (preto) e a escala de Black (roxo) na BSA, (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

As escalas de Sweet *et al.* e de Eisenberg *et al.* são muito semelhantes tanto em termos de variação como no comportamento ao longo da sequência de aminoácidos (Gráfico 3). Para além disso, apesar de não terem grande amplitude de variação, estas duas escalas têm um comportamento muito idêntico à escala de Kyte & Doolittle (Gráfico 4).

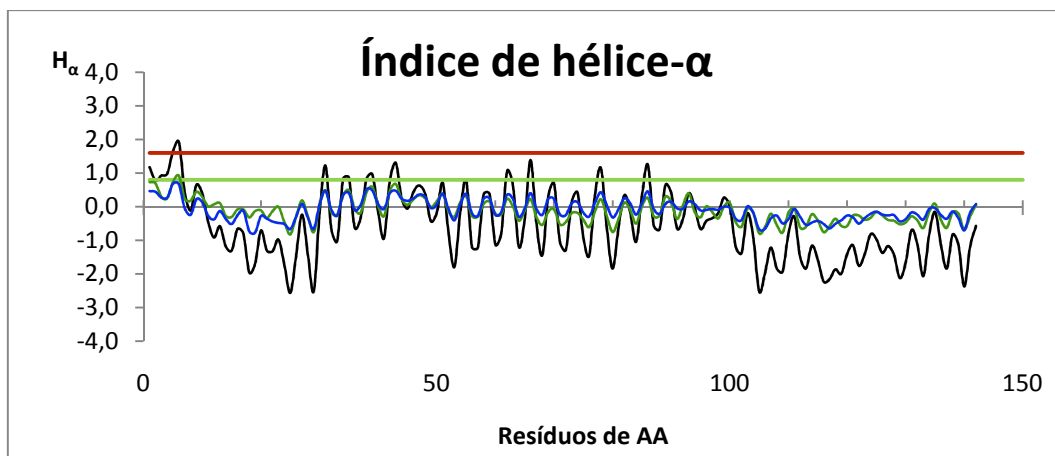


**Gráfico 3:** Representação gráfica das curvas de hidrofobicidade  $H_7$  para a escala de Sweet *et al.* (verde) e a escala de Eisenberg *et al.* (azul) na BSA, (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).



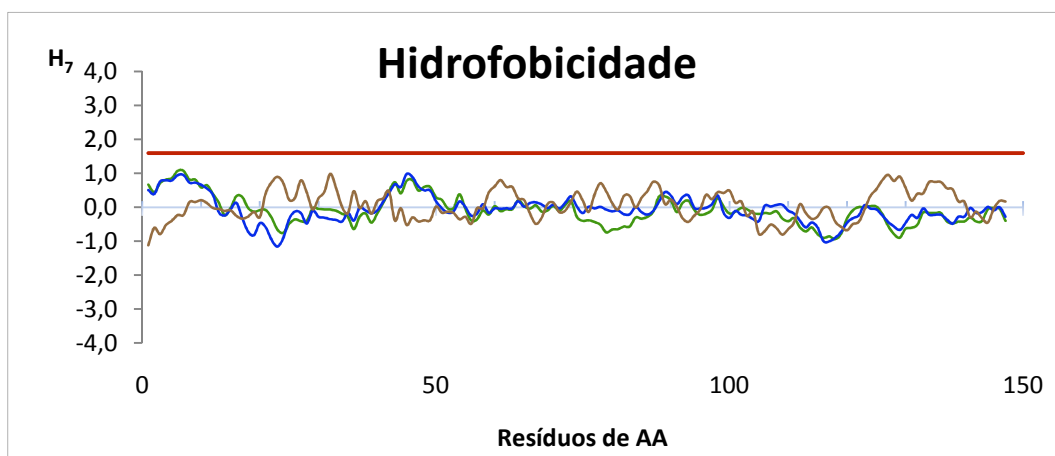
**Gráfico 4:** Representação gráfica das curvas de hidrofobicidade  $H_7$  para as escalas de Kyte & Doolittle (preto), de Sweet *et al.* (verde) e de Eisenberg *et al.* (azul) na BSA, (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).

As escalas de Sweet *et al.* e de Eisenberg *et al.* têm o mesmo comportamento quando aplicadas ao cálculo do índice de hélices- $\alpha$  (Gráfico 5), tal como acontece com o cálculo de hidrofobicidade  $H_7$  (Gráfico 4).

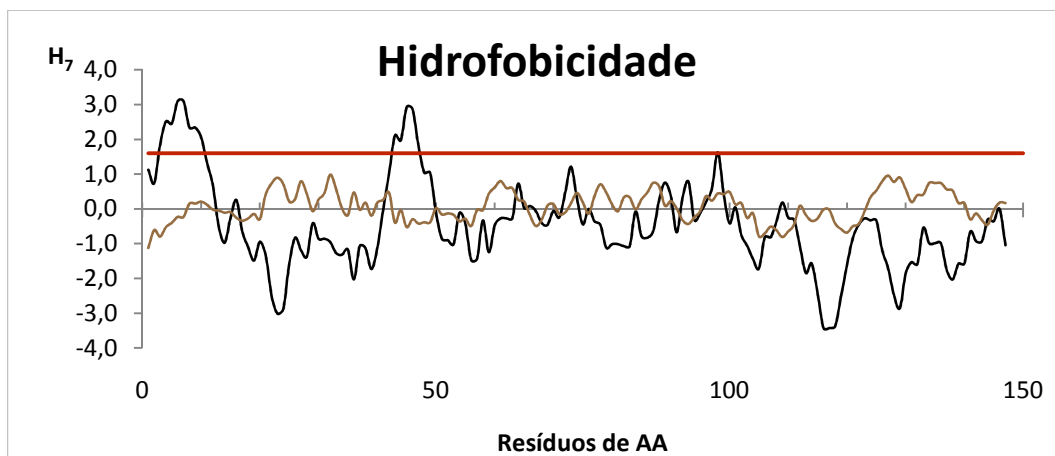


**Gráfico 5:** Representação gráfica das curvas de índice de hélice- $\alpha$  ( $H_\alpha$ ) para as escalas de Kyte & Doolittle (preto), de Sweet *et al.* (verde) e de Eisenberg *et al.* (azul) na BSA, (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

Os gráficos obtidos com a escala de Welling *et al.* apresentaram variações ligeiramente superiores às obtidas com as escalas de Sweet *et al.* e de Eisenberg *et al.* (Gráfico 6). No entanto, continuam a ser inferiores às variações obtidas com a escala de Kyte & Doolittle (Gráfico 7).

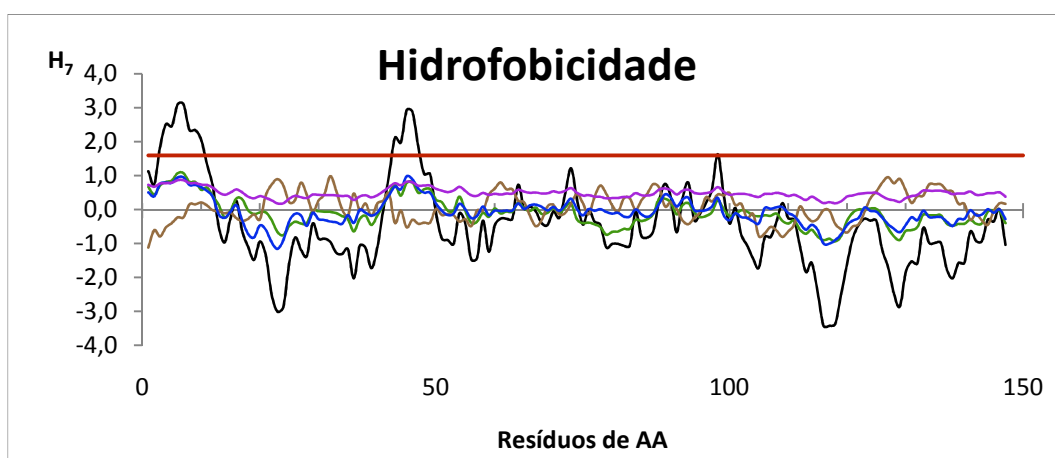


**Gráfico 6:** Representação gráfica das curvas de hidrofobicidade  $H_7$  para as escalas de Welling *et al.* (castanho), de Sweet *et al.* (verde) e de Eisenberg *et al.* (azul) na BSA, (numa amostra de 150 aminoácidos). Linha de referência de 1,6 (vermelho).



**Gráfico 7:** Representação gráfica das curvas de hidrofobicidade  $H_7$  para as escalas de Kyte & Doolittle (preto) e de Welling *et al.* (castanho) na BSA, (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).

Para além disso, dependendo dos aminoácidos aos quais se refere, ela pode ter comportamentos totalmente opostos à de qualquer uma das outras escalas (Gráfico 8).



**Gráfico 8:** Representação gráfica das curvas de hidrofobicidade  $H_7$  para as escalas de Kyte & Doolittle (preto), de Sweet *et al.* (verde), de Willing *et al.* (castanho), de Eisenberg *et al.* (azul) e de Black (roxo), na BSA, (numa amostra de 150 aminoácidos de um total de 607). Linha de referência de 1,6 (vermelho).

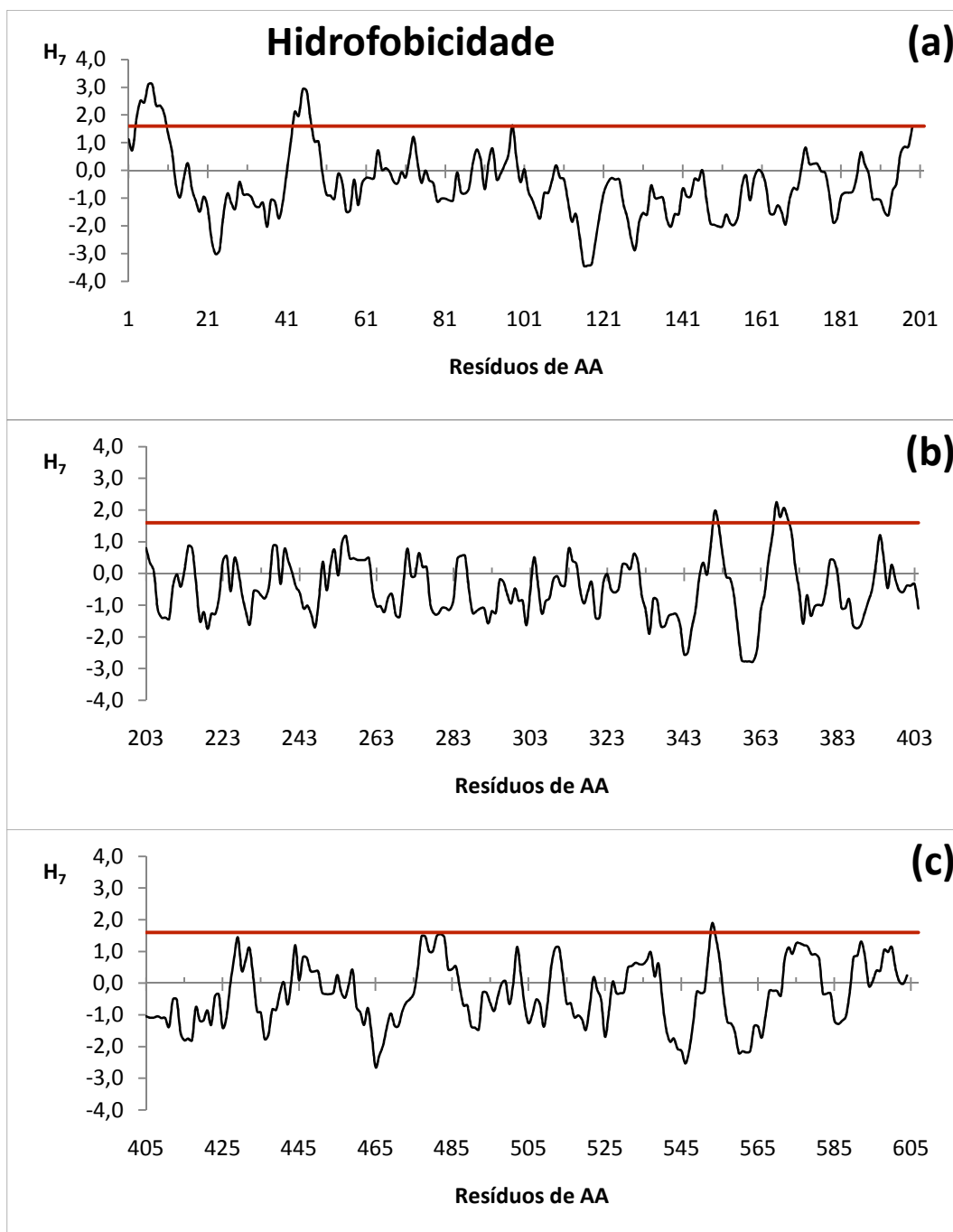
Apesar de não estar representada graficamente, confirma-se que a escala de Welling *et al.* tem o mesmo tipo de comportamento relativamente às outras escalas, quando é aplicada ao cálculo do índice de hélice- $\alpha$  ( $H_\alpha$ ).

Tendo em conta os resultados obtidos para a BSA (Gráficos 1 a 8) que são semelhantes aos obtidos para as restantes proteínas (Anexo I), optou-se pela utilização da escala de Kyte & Doolittle. Sendo assim, a apresentação e a discussão dos restantes gráficos referentes à hidrofobicidade ( $H_7$ ) e índice de hélice- $\alpha$  ( $H_\alpha$ ), serão elaborados a partir dessa mesma escala.

#### 4.2.1. BSA

A BSA é uma proteína de transporte cujas propriedades e características estão bem definidas. O estudo com a BSA tem como objectivo verificar se os cálculos conduzem a conclusões concordantes como as referidas na bibliografia (Jedrychowski & Wichers, 2009; Carter and Ho, 1994).

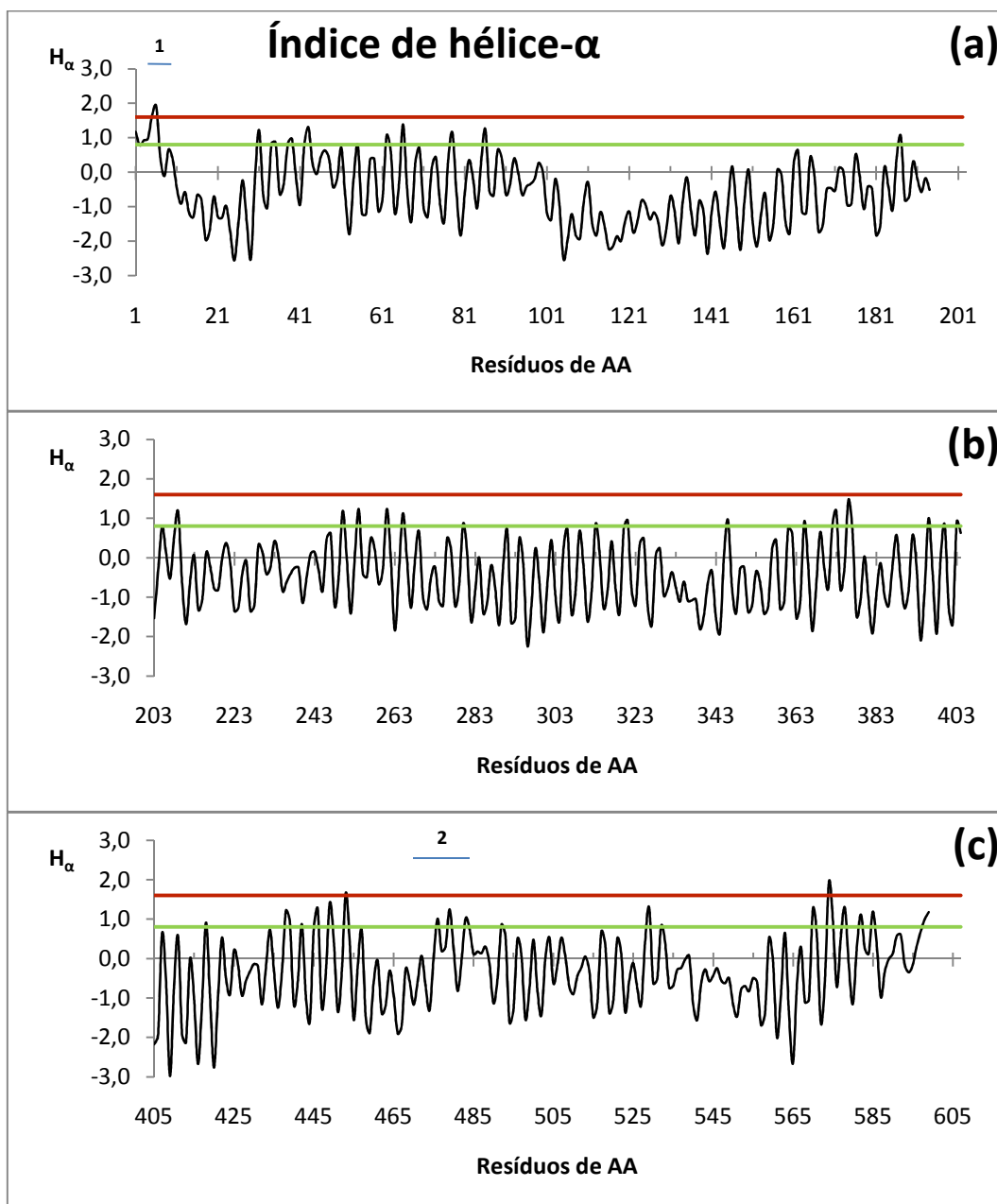
No gráfico 9 encontra-se o resultado do estudo de hidrofobicidade ( $H_7$ ) para a BSA.



**Gráfico 9:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a BSA (607 resíduos de aminoácidos) usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho). (a) resíduos 1-201; (b) resíduos 203-404 e (c) resíduos 405-605.

Apesar de poder observar-se seis picos que ultrapassam o valor de hidrofobicidade de 1,6 não se pode afirmar que a BSA possui zonas hidrofóbicas bem definidas, porque as diferentes secções de aminoácidos necessários para calcular a média de 7 aminoácidos, para cada valor que define os diferentes picos, é inferior a 19 resíduos (Gráfico 9) (Han e Tashjian, 2003). Também, nada indica que a vizinhança de cada um desses picos tenha um comportamento muito hidrofóbico. Esta ideia completa-se pelo facto de se saber que a albumina é hidrossolúvel (Jedrychowski & Wichers, 2009; Carter and Ho, 1994).

Graças à mesma escala é possível calcular o índice de hélice- $\alpha$  (Gráfico 10).

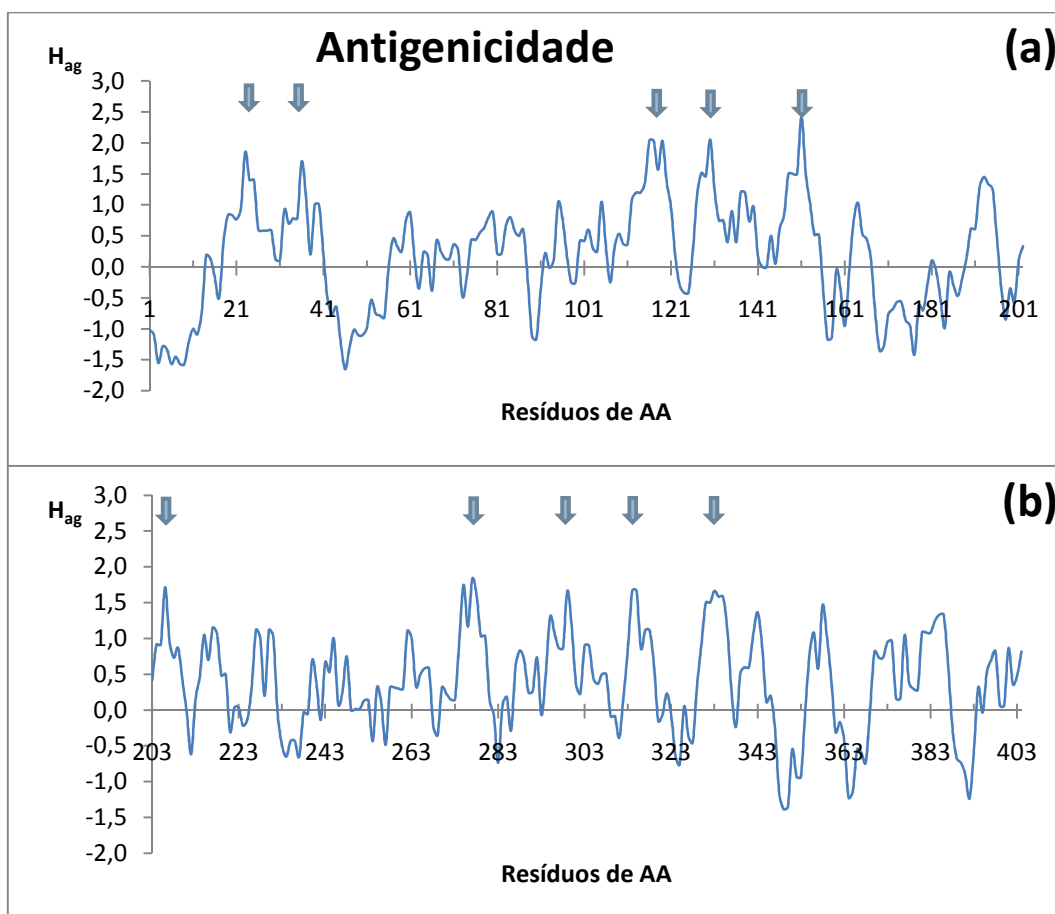


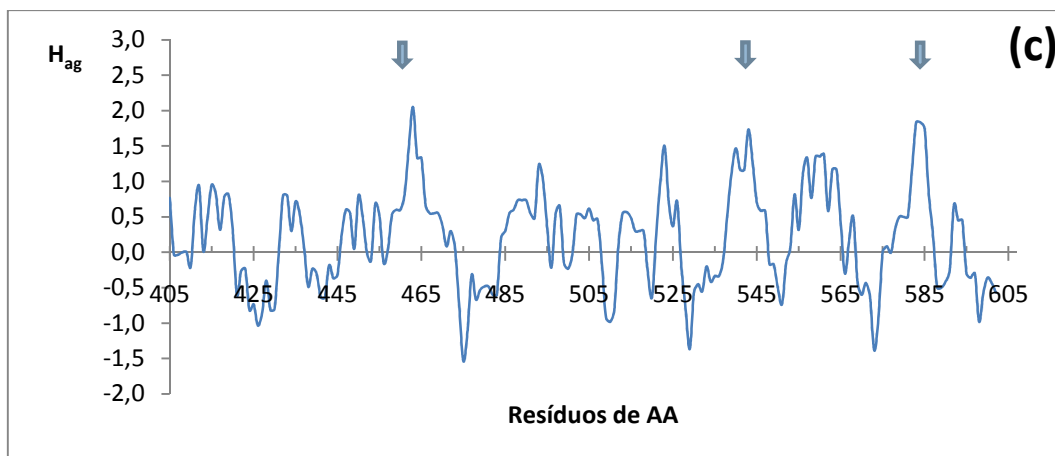
**Gráfico 10:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a BSA, usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro). (a) resíduos 1-202; (b) resíduos 203-404 e (c) resíduos 405-607.

Tendo em conta que a BSA é constituída a 67% de hélices- $\alpha$  (Carter and Ho, 1994), é perfeitamente normal que existam poucos picos hidrofóbicos superiores a 1,6. A razão para o sucedido é que as hélices- $\alpha$  são na sua maioria hidrofílicas. Excepto, quando elas são constituídas por resíduos alternadamente hidrofóbicos e hidrofílicos, onde elas têm um comportamento anfifílico, dependendo do meio em que se encontram, (polar ou apolar) (Murray *et al.*, 2002)

É de notar a presença de uma zona hidrofóbica (Gráfico 10, pico 1) e uma zona anfifílica (Gráfico 10, pico 2). Este último pico poderá indicar a presença de uma hélice- $\alpha$  anfifílica situada no interior da BSA.

A antigenicidade da BSA pode ser observada no gráfico 11.





**Gráfico 11:** Representação gráfica das zonas antigénicas, para a BSA, usando a escala de Hopp & Wood (azul). (a) resíduos 1-202; (b) resíduos 203-404 e (c) resíduos 405-607. Zonas antigénicas possíveis (setas azuis).

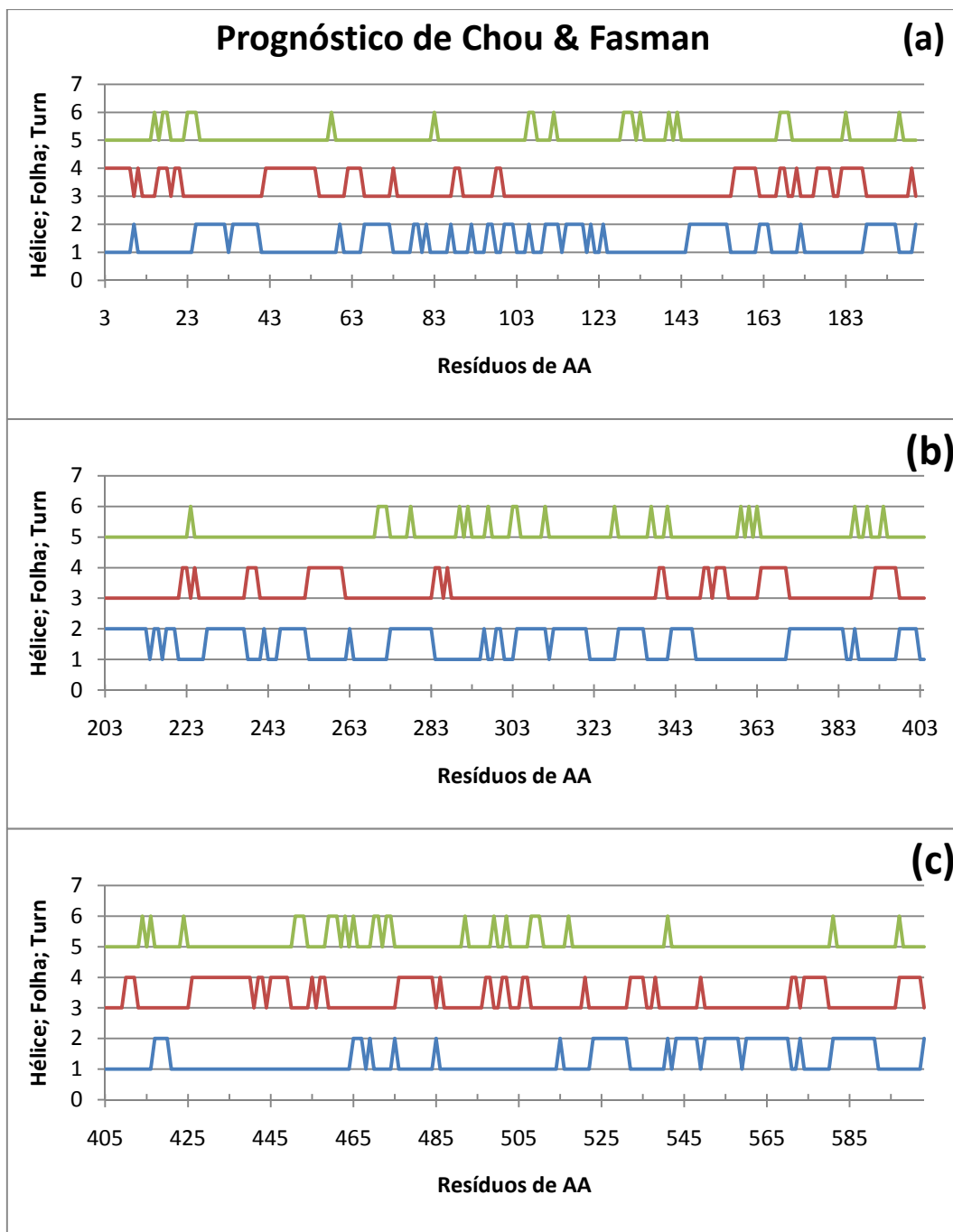
Seria de esperar que nas zonas hidrofílicas da BSA existissem regiões antigénicas. De facto, ao observar-se o gráfico 11, constata-se a presença de treze picos com valores mais elevados que correspondem às possíveis zonas antigénicas da BSA. Como a BSA é fisicamente e imunologicamente muito semelhante à albumina do soro humano (Jedrychowski & Wichers, 2009; Carter and Ho, 1994), seria de esperar que a BSA não manifestasse qualquer tipo de antigenicidade. No entanto, ela é considerada como sendo um alérgico (Jedrychowski & Wichers, 2009).

A presença ou a ausência dos principais elementos da estrutura secundária das proteínas: as hélices- $\alpha$ ; folhas- $\beta$  e voltas, (Gráficos: 12; 19; 28; 37; 52; 61; 70; 79; 88; 97; 106 e 115), são definidas através do algoritmo de Chou & Fasman. Cada pico (triângulo) corresponde à probabilidade de encontrar uma das três estruturas, dependendo da linha que se está a ler. Quanto mais larga for a forma do pico (trapézio isósceles) maior probabilidade de encontrar a estrutura naquela região da cadeia polipeptídica.

O método baseia-se na atribuição de um valor 1 para uma distribuição ao acaso de um aminoácido em hélice- $\alpha$ , folhas- $\beta$  ou voltas. Um valor superior a 1 sugere uma tendência para encontrar o resíduo num elemento particular (Whitford, 2005). Por exemplo, a alanina tem  $P_{\alpha}=1,41$ ;  $P_{\beta}=0,72$  e  $P_t=0,82$  o que significa que a alanina ocorre 41% mais frequentemente do que seria de esperar se a sua ocorrência numa hélice fosse ao acaso (Chou & Fasman, 1978). Seguindo o mesmo raciocínio, a probabilidade de ocorrência de alanina em folhas- $\beta$  é 28% inferior à probabilidade ao acaso. A presença de alanina em voltas é 18% inferior à probabilidade ao acaso.

Resumidamente, segundo o algoritmo de Chou & Fasman, regiões com pelo menos 4 resíduos consecutivos em 6, apresentam  $P_{\alpha} > 1,0$  correspondem a uma hélice- $\alpha$ . As folhas- $\beta$  são identificadas em regiões onde 3 de 5 resíduos têm um valor de  $P_{\beta} > 1,0$  (Whitford, 2005).

O gráfico 12 demonstra a aplicação do método de Chou & Fasman para a BSA.



**Gráfico 12:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na BSA. Cada pico corresponde à presença de um possível elemento. (a) resíduos 3-202; (b) resíduos 203-404 e (c) resíduos 405-603.

Para uma melhor leitura e visto que a BSA é uma proteína de grande dimensão, o gráfico 12 de Chou & Fasman foi dividido em três partes (a, b e c).

O método de Chou & Fasman, mostra que existe uma grande probabilidade de encontrar hélices- $\alpha$ ; folhas- $\beta$  e voltas ao longo da cadeia polipeptídica da BSA (Gráfico 12). Sabe-se que a BSA possui 30 hélices- $\alpha$  (Carter and Ho, 1994) e 9 voltas (Jedrychowski & Wichers, 2009; Carter and Ho, 1994). Se for tido em conta apenas os picos largos (trapézio isósceles), é possível contar 9 voltas e 29 hélices- $\alpha$  (Gráfico 12). Isto revela uma certa

concordância entre o que se conhece sobre a BSA e os resultados obtidos através deste método. No entanto, sabe-se que a BSA não possui folhas- $\beta$  (Carter and Ho, 1994). Apesar disso, o gráfico 12 indica a presença de inúmeras folhas- $\beta$ , o que revela uma possível inadequação deste método relativamente à detecção das folhas- $\beta$ .

O algoritmo de Chou & Fasman pode apresentar uma taxa de sucesso até cerca de 70%. O principal problema do algoritmo é que ele só considera interações locais e negligencia as de longo alcance que influenciam a estabilidade da estrutura secundária. Por outro lado também não distingue entre tipos de hélices, tipos de voltas ou a orientação das folhas- $\beta$  (Whitford, 2005).

É de notar que a probabilidade de encontrar hélices- $\alpha$  e folhas- $\beta$  nunca se sobrepõem. No entanto, é possível encontrar sobreposições entre as voltas e qualquer uma das duas outras estruturas.

Num princípio quase idêntico mas usando outro tipo de cálculos, as super-hélices- $\alpha$ , (Gráficos: 13; 20; 29; 38; 53; 62; 71; 80; 89; 98; 107 e 116), são definidas através do algoritmo de Lupas *et al.* Tal como para o método de Chou & Fasman, o método de Lupas *et al.* (Filby, 1998) também está fundamentado em estudos estatísticos e probabilísticos. Uma das diferenças é que este método está orientado para a detecção de estruturas super-enroladas (super-hélices- $\alpha$ ) das quais as hélices- $\alpha$  fazem parte. Estes investigadores propuseram a existência de uma probabilidade para que cada aminoácido se encontre numa das sete posições previstas por eles, posições (a; b; c; d; e; f e g) (Filby, 1998). Estas posições correspondem às posições referenciadas na tabela 12.

A aplicação do algoritmo de Lupas pode ser observada através do gráfico 13 para a BSA

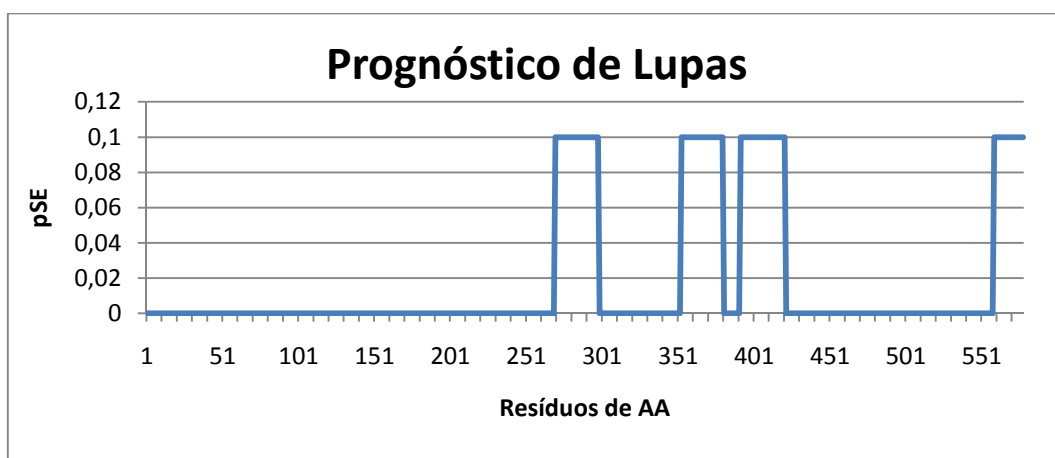


Gráfico 13: Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na BSA. A cada pico corresponder a possível presença dessa estrutura.

Observa-se que existe alguma probabilidade de encontrar quatro super-hélices- $\alpha$  na BSA (Gráfico 13). A primeira estrutura super enrolada encontra-se entre os resíduos 270-298,

a segunda entre os resíduos 353-380 e a terceira entre os resíduos 392-421. Há fortes indícios para que haja uma quarta estrutura a partir dos resíduos 559.

Quando se faz um estudo deste género, é importante ter em conta várias características das proteínas. Se até aqui foi observado como se comportam em média os aminoácidos constituintes das proteínas, deve-se também ter em conta a quantidade real de cada aminoácido que compõem cada uma das proteínas (Gráficos: 14; 21; 30; 39; 54; 63; 72; 81; 90; 99; 108 e 117); a percentagem relativa desses mesmos aminoácidos (Gráficos: 22; 31; 40; 55; 64; 73; 82; 91; 100; 109 e 118); as características químicas (Gráficos: 23; 32; 41; 56; 65; 74; 83; 92; 101; 110 e 119) e por fim as cargas das proteínas (Gráficos: 15; 24; 33; 42; 57; 66; 75; 84; 93; 102; 111 e 120).

A visão global da composição das proteínas pode ajudar a prever se esta ou aquela proteína possui na sua estrutura secundária a formação de elementos como: as hélices- $\alpha$ ; folhas-B e/ou voltas. De facto, foi visto anteriormente (Tabela 4), que determinados aminoácidos são constituintes que favorecem, ou perturbam, a formação de cada um desses elementos. Por exemplo; alanina, glutamato e leucina, tendem a estar presentes em hélice- $\alpha$  enquanto que valina, treonina e isoleucina tendem a formar folhas-B. Glicina, asparagina e prolina têm tendência para formar voltas (Berg et al., 2002).

O gráfico 14 proporciona uma visão global da composição em aminoácidos da BSA.

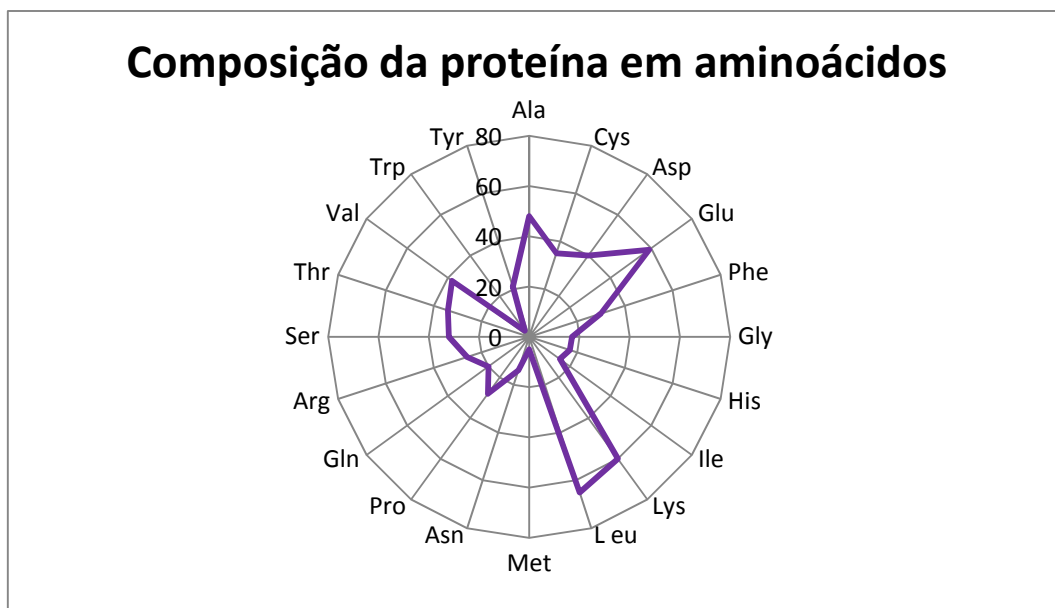


Gráfico 14: Representação gráfica da quantidade real de cada aminoácido na BSA (607 resíduos de aminoácido).

Observa-se que a BSA possui grandes quantidades (cerca de 60 unidades), de leucina; lisina e glutamato (Gráfico 14). Possui igualmente uma quantidade substancial, de todos os outros aminoácidos, que varia entre 14 e 48 unidades. No entanto, a BSA possui muito pouca quantidade de metionina e triptofano. Se for referida a tendência que alguns aminoácidos têm em formar hélices- $\alpha$ ; folhas-B e voltas, pode dizer-se que a BSA é constituída

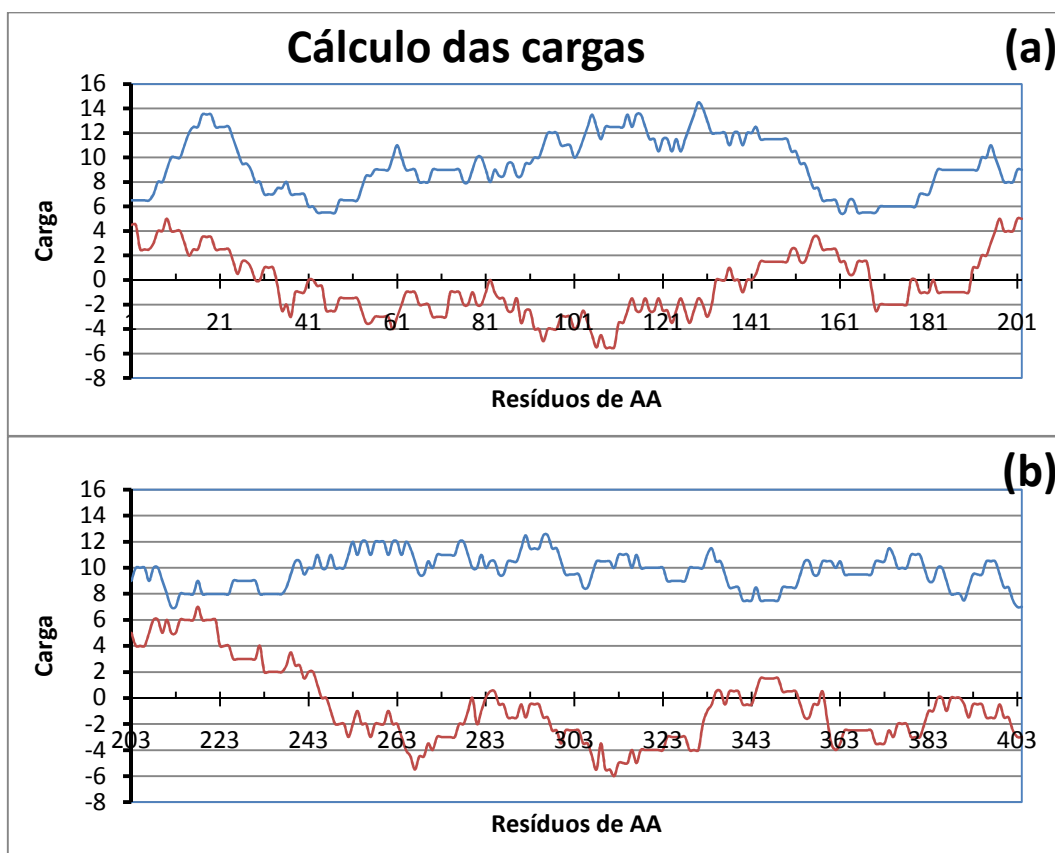
maioritariamente de hélices- $\alpha$ , devido à grande quantidade de leucina e glutamato (Berg et al., 2002). Isto confirma o facto de a BSA ser constituída em mais de 60% por hélices- $\alpha$  (Carter and Ho, 1994).

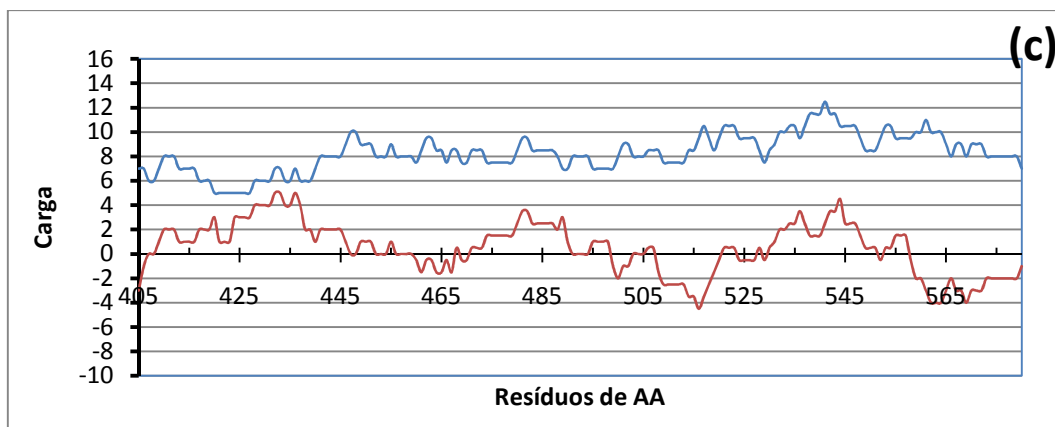
Quando se fala da carga global de uma proteína, apenas são responsáveis por essa carga, cinco dos vinte aminoácidos principais. Esses aminoácidos são: a lisina; arginina; histidina; glutamato e aspartato. Nos cálculos que serviram para a elaboração do gráfico das cargas, foram considerados os seguintes valores: 1; 0,5 e -1. Os aminoácidos básicos (lisina e arginina) assumem o valor 1 e os aminoácidos ácidos (glutamato e aspartato) assumem o valor -1. A histidina, como é meia básica, assume o valor 0,5 (Filby, 1998). Assim, de cada vez que aparece um destes aminoácidos ao longo da cadeia polipeptídica da proteína em estudo, o aminoácido assume o valor que lhe é devido. Cada carga é calculada através da soma dos valores obtidos de 28 aminoácidos consecutivos ao longo de toda a cadeia polipeptídica. No cálculo das cargas absolutas, procedeu-se da mesma maneira só que todos os valores negativos passaram a positivos.

As cargas reais, como o nome indica, são cargas que reflectem o comportamento eléctrico dos referidos aminoácidos ao longo da cadeia peptídica da proteína.

As cargas absolutas, reflectem o comportamento eléctrico global de determinadas regiões da proteína, visto que todas as cargas negativas são convertidas em positivas.

As cargas reais e as cargas absolutas da BSA são apresentadas no gráfico 15.





**Gráfico 15:** Representação gráfica das cargas da BSA. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos. (a) resíduos 1-202; (b) resíduos 203-404 e (c) resíduos 405-580.

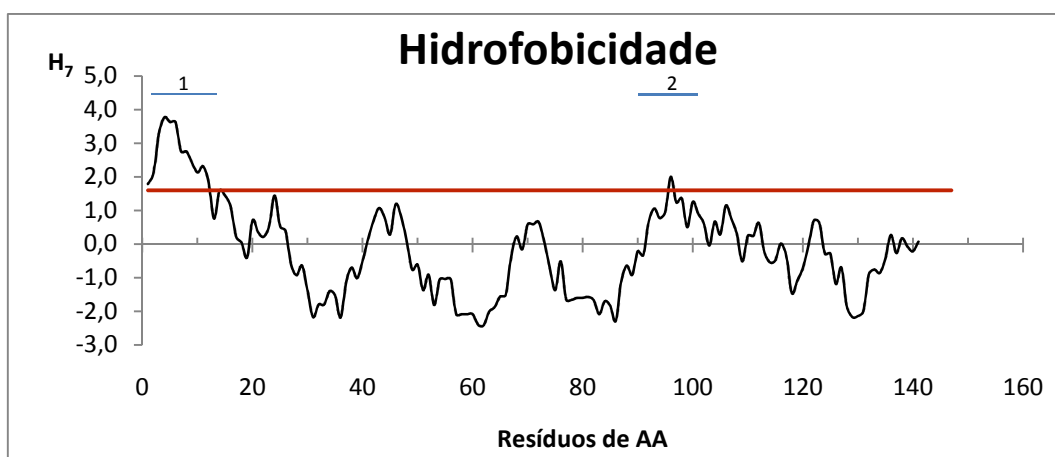
Observa-se que a BSA possui cargas reais positivas entre os intervalos 1-28, 142-167 (Gráfico 15a); 191-249 (Gráfico 15a, b); 413-450, 475-494 e 534-555 (Gráfico 15c). Contudo, existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores (Gráfico 15).

Ao observar-se as cargas absolutas da BSA, verifica-se que ela tem um comportamento polar positivo ao longo de praticamente toda a cadeia polipeptídica. Isto significa que a BSA colocada numa solução sob a acção de uma corrente eléctrica, como na electroforese desloca-se para o cátodo (pólo negativo).

#### 4.2.2. Proteínas do ovo

##### 4.2.2.1. Lisozima C

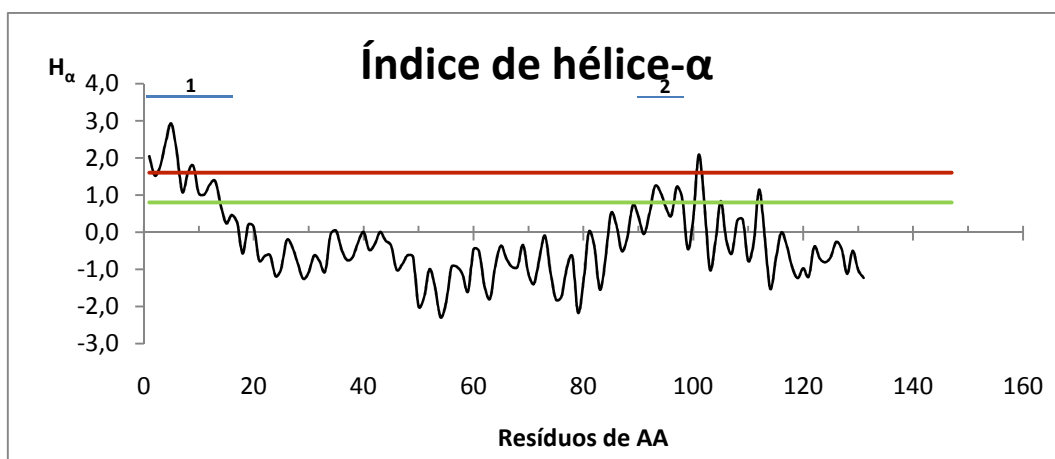
A hidrofobicidade  $H_7$  para a lisozima C pode ser representada do seguinte modo.



**Gráfico 16:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  com base na escala de Kyte & Doolittle (preto) para a lisozima C (147 resíduos de aminoácido). Linha de referência de 1,6 (vermelho).

Pela análise do gráfico 16, excepto no ponto 1 e no ponto 2, que representam possíveis regiões internas da proteína (Voet *et al.*, 2004), verifica-se que a lisozima C tem um comportamento hidrofílico, uma vez que os valores de hidrofobicidade encontram-se abaixo de 1,6. Tendo em conta que a lisozima C está presente, entre outros lados, na clara de ovo (Jedrychowski & Wichers, 2009; Mine, 2008) e que esta última é constituída essencialmente de água (Mine, 2008; Belitz *et al.*, 2009), é perfeitamente normal que a lisozima C tenha um comportamento hidrofílico.

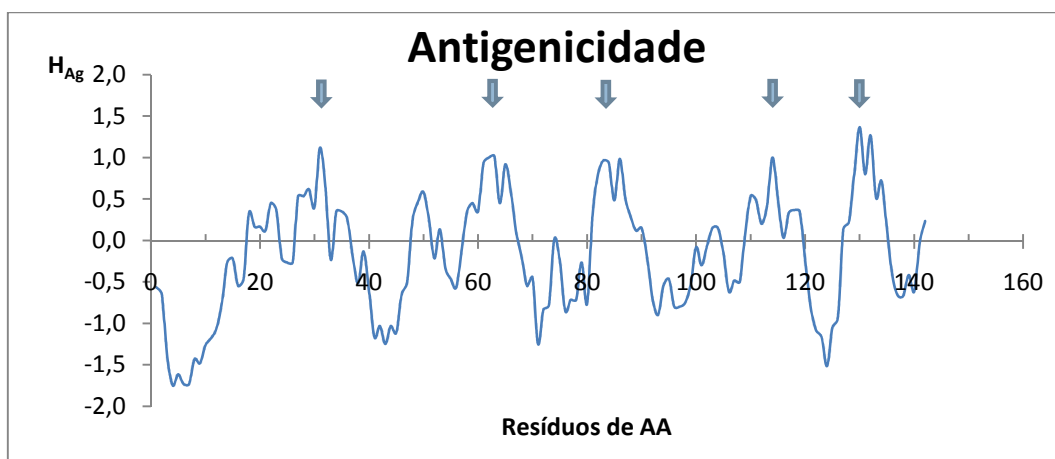
No gráfico 17, observa-se o resultado dos cálculos para o índice de hélice- $\alpha$  para a lisozima C.



**Gráfico 17:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a lisozima C, usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

Através do gráfico 17, observa-se uma zona inicial hidrofóbica (1) e uma pequena zona anfifílica (2). Este último ponto pode indicar a presença de uma hélice- $\alpha$  anfifílica. Este tipo de hélice- $\alpha$  é um caso particular onde os resíduos são alternadamente hidrofóbicos e hidrofílicos (Murray *et al.*, 2002).

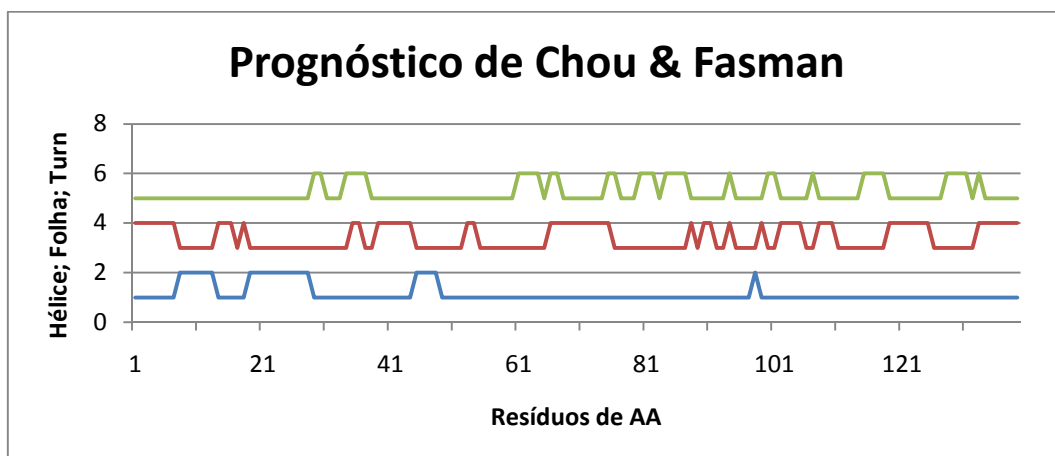
A antigenicidade da lisozima C pode ser representada através do gráfico 18.



**Gráfico 18:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a lisozima C, usando a escala de Hopp & Wood (azul). Zonas antigénicas possíveis (setas azuis).

No que diz respeito aos antígenos, o gráfico 18 revela a previsão de cinco picos antigénicos. Este comportamento era pouco previsível, visto que a lisozima C é um alergénio menor (Mine, 2008) e por essa razão deveria haver poucos picos antigénicos.

O algoritmo de Chou & Fasman, para a lisozima C pode ser representado no gráfico 19.



**Gráfico 19:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na lisozima C. Cada pico corresponde à presença de um possível elemento.

Observa-se que existe a probabilidade de encontrar três hélices- $\alpha$  ao longo da cadeia polipeptídica da lisozima C (Gráfico 19). Este facto confirma-se visto que ela é constituída por dois domínios,  $\alpha$  e  $\beta$  ligados por uma longa hélice- $\alpha$  entre a qual está localizado o sítio activo (Mine, 2008). O gráfico 19 revela igualmente a existência hipotética de dez voltas e onze folhas- $\beta$ .

Tendo em conta a incerteza associada ao método de Chou & Fasman, é de prever que as onze folhas- $\beta$  detectadas na lisozima C, não reflectam a realidade.

Depois de efectuados todos os cálculos, o algoritmo de Lupas apresenta-se no gráfico 20.

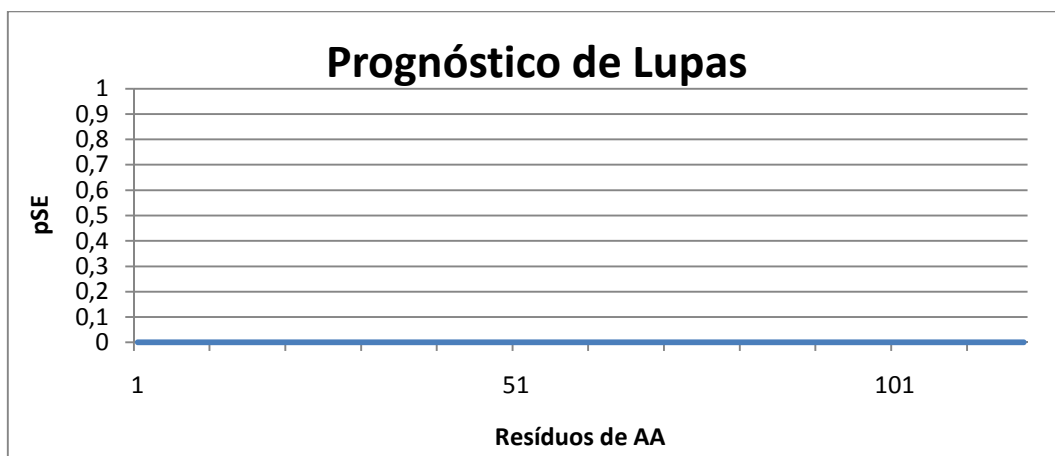


Gráfico 20: Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na lisozima C.

Observa-se que a probabilidade de haver super-hélices- $\alpha$  na lisozima C é nula (Gráfico 20). Apesar de se ter observado anteriormente que a lisozima C pode possuir várias hélices- $\alpha$  ao longo da sua cadeia polipeptídica, nada garante que estas possam formar super-hélices- $\alpha$ .

Os gráficos em radar ajudam a visualizar a composição das proteínas em aminoácidos.

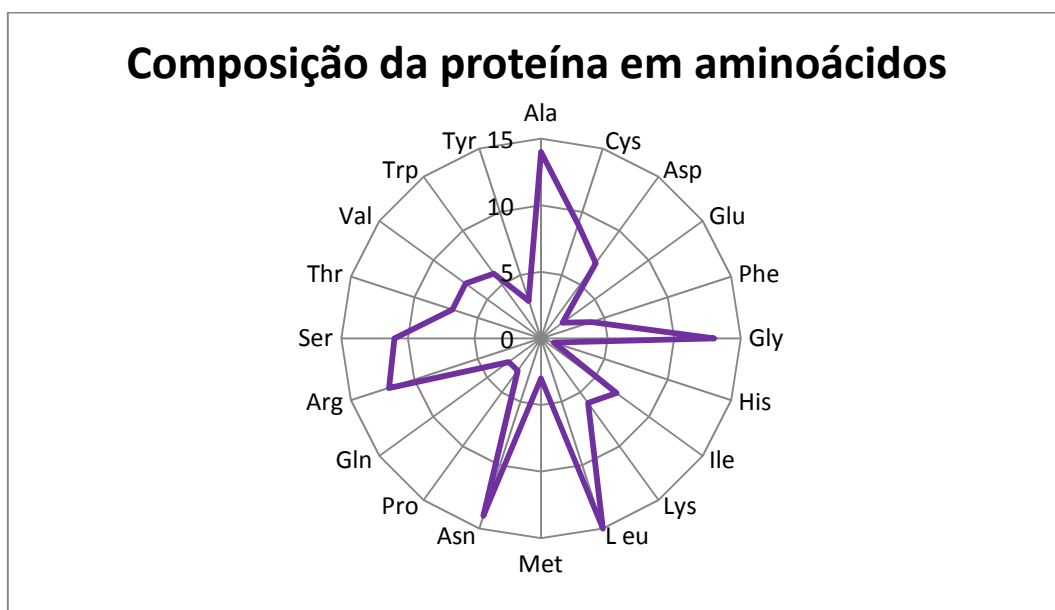


Gráfico 21: Representação gráfica da quantidade real de cada aminoácido na lisozima C (147 resíduos de aminoácido).

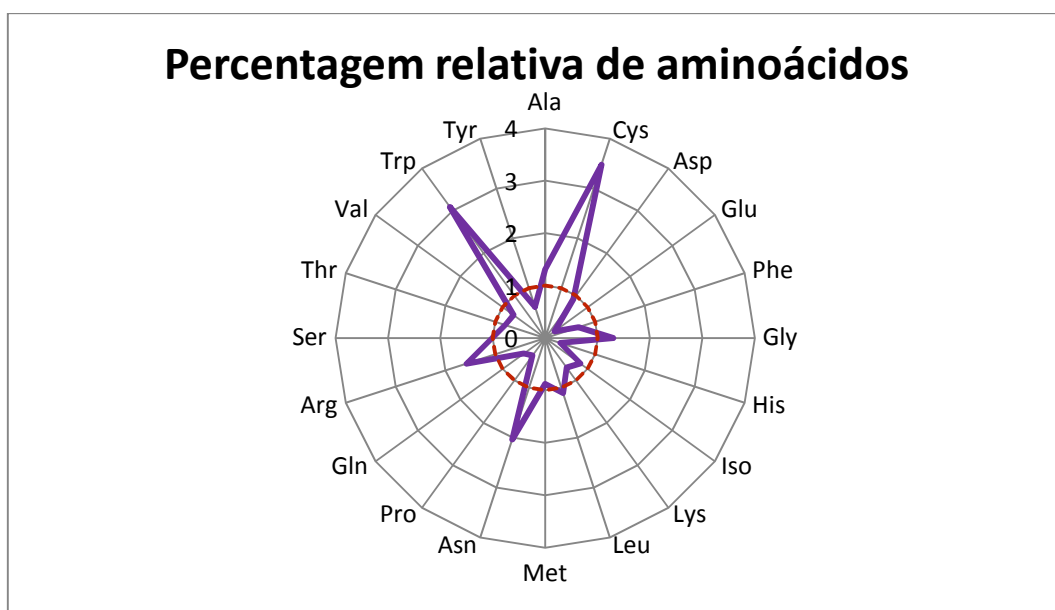
A lisozima C é constituída pelos vinte principais aminoácidos (Gráfico 21). Observa-se que a lisozima C é maioritariamente composta por: leucina; alanina; asparagina; glicina; arginina e serina. Tem uma quantidade substancial de: cisteína; aspartato; isoleucina; treonina; valina; lisina; triptofano e fenilalanina. Possui um valor residual de: metionina; prolina; glutamina; tirosina; glutamato e histidina. Visto que a lisozima C é rica em leucina e alanina, existe uma forte suspeita de que ela tenha pelo menos uma hélice- $\alpha$  na sua estrutura

(Berg et al., 2002). No entanto, a grande quantidade de asparagina e serina pode perturbar a formação dessa mesma hélice- $\alpha$  (Berg et al., 2002). Os valores médios de isoleucina, treonina e valina também revelam a possível presença de folhas- $\beta$  (Berg et al., 2002). A presença de voltas é igualmente possível devido à existência de valores residuais de prolina (Murray *et al.*, 2002).

Para se obter a percentagem relativa de cada aminoácido numa proteína relativamente as outras, é necessário dividir a percentagem de cada aminoácido da proteína em estudo, pela percentagem desse mesmo aminoácido na proteína comparativa (equação 4.2). Como proteína comparativa, escolheu-se a GenPep#97, que representa na realidade, a média das percentagens de cada aminoácido em 10000 proteínas.

$$\% \text{ relativa de cada aminoácido} = \frac{\% \text{ de aminoácidos na proteína em estudo}}{\% \text{ de aminoácidos na GenPep\#97}} \quad 4.2$$

O gráfico 22 mostra a composição relativa da lisozima C em relação à GenPep#97.

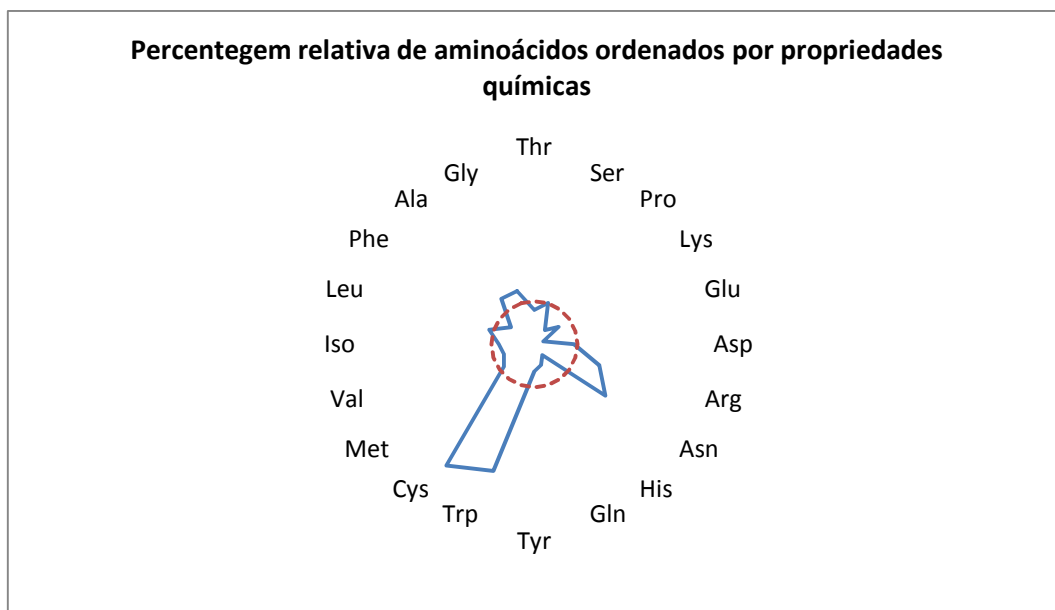


**Gráfico 22:** Representação gráfica da percentagem de cada aminoácido na lisozima C, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína (147 resíduos de aminoácido).

Observa-se que a lisozima C possui oito aminoácidos com uma percentagem relativa acima da média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína), representada a tracejado vermelho (Gráfico 22). Desses oito, os mais evidentes são a cisteína, seguida do triptofano; asparagina; arginina; alanina; glicina; leucina e serina. Os restantes aminoácidos têm um valor abaixo da média.

Os gráficos de radar podem ser construídos de maneira a distribuir os aminoácidos mediante as suas propriedades químicas e físicas (Gráfico 23). Neste caso concreto, os aminoácidos foram distribuídos segundo duas propriedades. Os aminoácidos mais carregados encontram-se à direita do gráfico e os mais hidrofóbicos à esquerda (Filby, 1998).

A representação gráfica das propriedades químicas dos aminoácidos da lisozima C pode ser visualizada no gráfico 23.



**Gráfico 23:** Representação gráfica da percentagem relativa dos aminoácidos da lisozima C, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

Observa-se que a lisozima C possui uma grande percentagem relativa de cisteína e triptofano, relativamente à média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína), representada a tracejado vermelho (Gráfico 23). Como estes dois aminoácidos são hidrofóbicos, seria de esperar que esta proteína tivesse o mesmo tipo de comportamento. No entanto, a lisozima C é hidrossolúvel. Este facto deve-se certamente à quantidade de aminoácidos carregados, arginina e asparagina, que também eles são superiores à média. A lisozima C possui ainda mais três aminoácidos hidrofóbicos, com uma percentagem relativa superior à média: alanina, glicina e leucina.

As cargas reais e absolutas da lisozima C podem ser representadas sob a forma de gráfico.

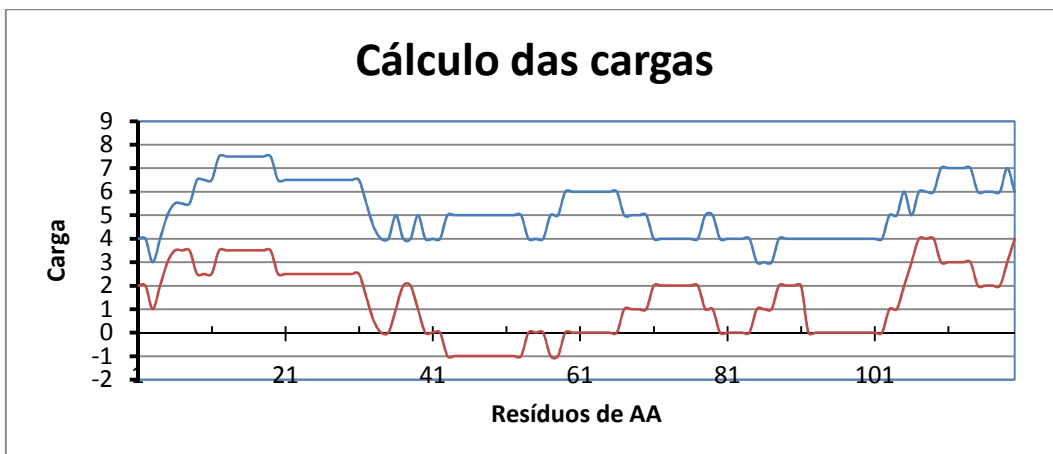


Gráfico 24: Representação gráfica das cargas da lisozima C. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

Observa-se que a lisozima C possui cargas reais positivas entre os intervalos 1-33 e 103-120 (Gráfico 24). Existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores. Esses intervalos são: 36-39, 67-79 e 85-91.

Observa-se igualmente que as cargas absolutas da lisozima C, mostram que ela tem um comportamento polar positivo ao longo de praticamente toda a cadeia polipeptídica, mas essencialmente no início e no final. Isto significa que se a lisozima C for colocada numa solução com uma corrente eléctrica, ela deslocar-se-á para o cátodo (pólo negativo).

## 4.2.2.2. Ovalbumina

Devido ao facto da ovalbumina ter um tamanho relativamente grande, (mais de 270 resíduos de aminoácidos), foi necessário dividir em dois os diferentes gráficos ( $H_7$ ,  $H_\alpha$  e  $H_{Ag}$ ).

A hidrofobicidade da ovalbumina pode representar-se graficamente para uma melhor visualização (Gráfico 25).

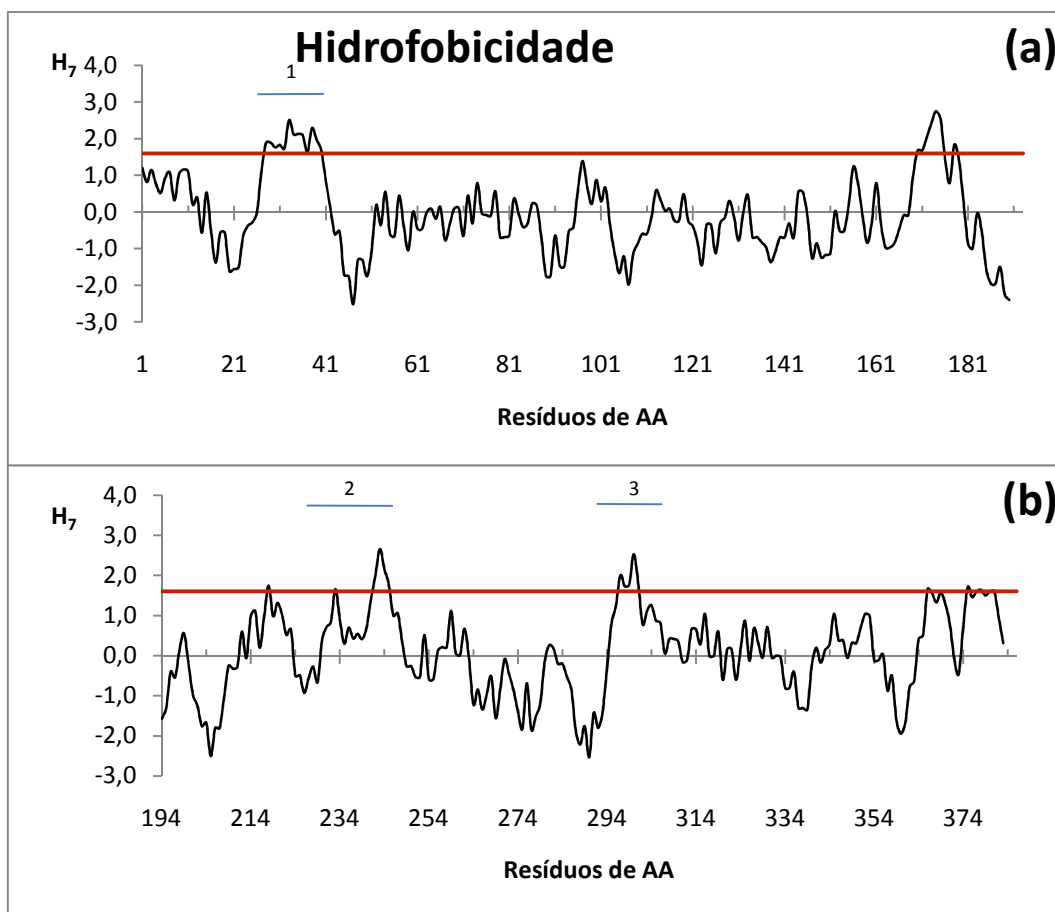


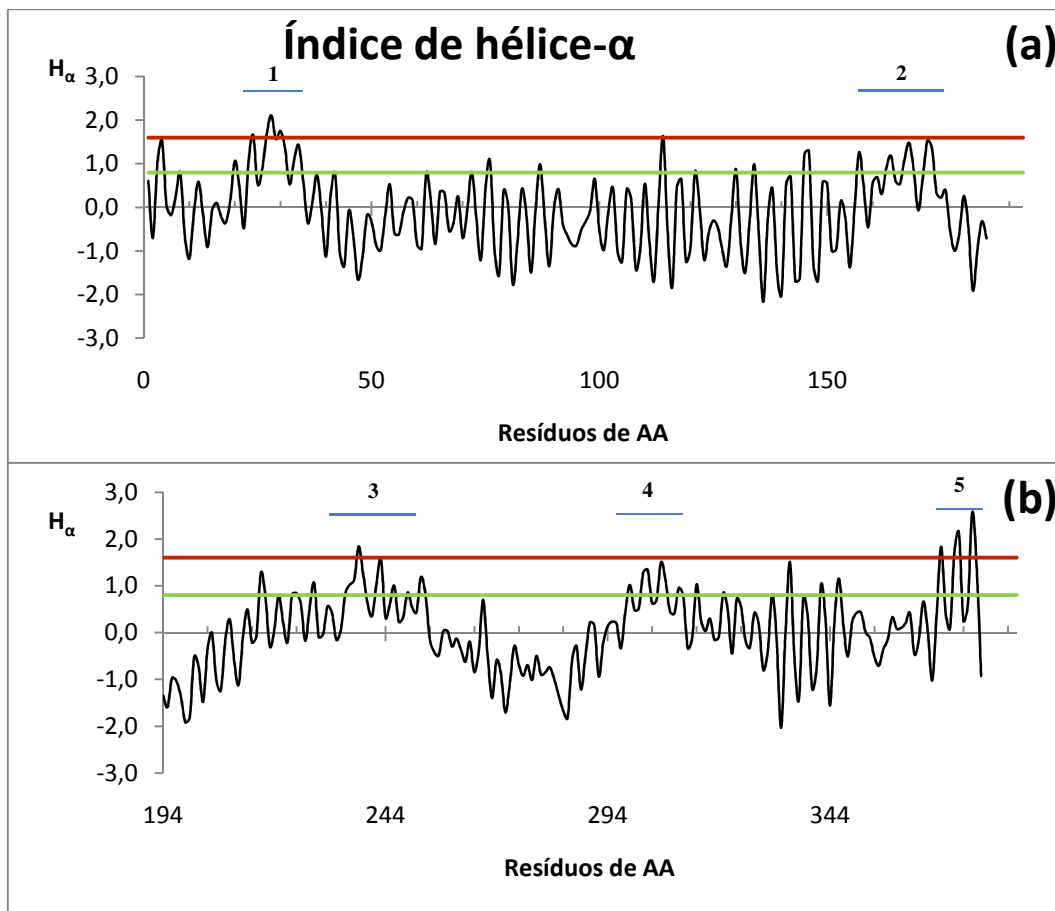
Gráfico 25: Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a ovalbumina (386 resíduos de aminoácido) usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho). (a) resíduos 1-193 e (b) resíduos 194-386.

Observa-se um pico de hidrofobicidade entre os resíduos 30 e 45 (Gráfico 25a), valores incluídos no intervalo estipulado na teórica (Jedrychowski & Wichers, 2009).

Apesar de haver apenas 13 médias de resíduos com um grau de hidrofobicidade superior ou igual a 1,6 neste intervalo, são necessários 21 para efectuar os cálculos que permitem desenhar este pico. Como a vizinhança próxima também tem um comportamento hidrofóbico, apesar de ser inferior a 1,6 considera-se uma zona hidrofóbica, porque possui mais de 19 resíduos de aminoácidos. O mesmo se sucede com os picos 2 e 3 (Gráfico 25b), que apesar de ter poucos resíduos com uma hidrofobicidade acima de 1,6 têm ambos uma vizinhança bastante hidrofóbica. Estes três picos representam possíveis regiões internas da ovalbumina (Voet *et al.*, 2004),

Exceptuando estes três pontos, pode dizer-se, através deste gráfico, que a ovalbumina mostra ter um comportamento bastante hidrofílico.

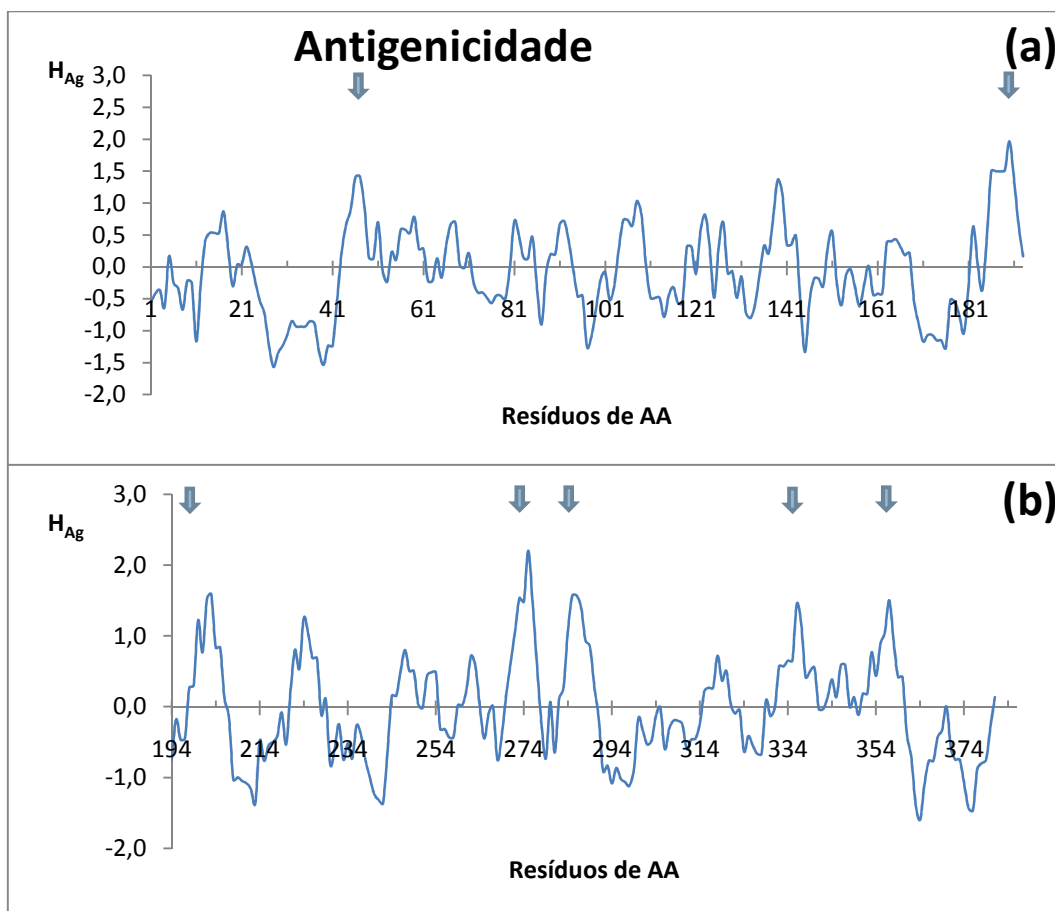
As hélices- $\alpha$  podem ser detectadas através do índice de hélices- $\alpha$  (Gráfico 26).



**Gráfico 26:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a ovalbumina, usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro). (a) resíduos 1-193 e (b) resíduos 194-386.

No que diz respeito ao índice de hélice- $\alpha$  (Gráfico 26), nota-se que apenas o pico 1 tem um comportamento hidrofóbico. Os restantes picos (2; 3 e 4), correspondem a zonas anfifílicas, podendo indicar cada um deles a presença de uma hélice- $\alpha$  anfifílica. A zona 5, apesar de ter um valor superior a 1,6 é demasiado irregular para se ter em conta.

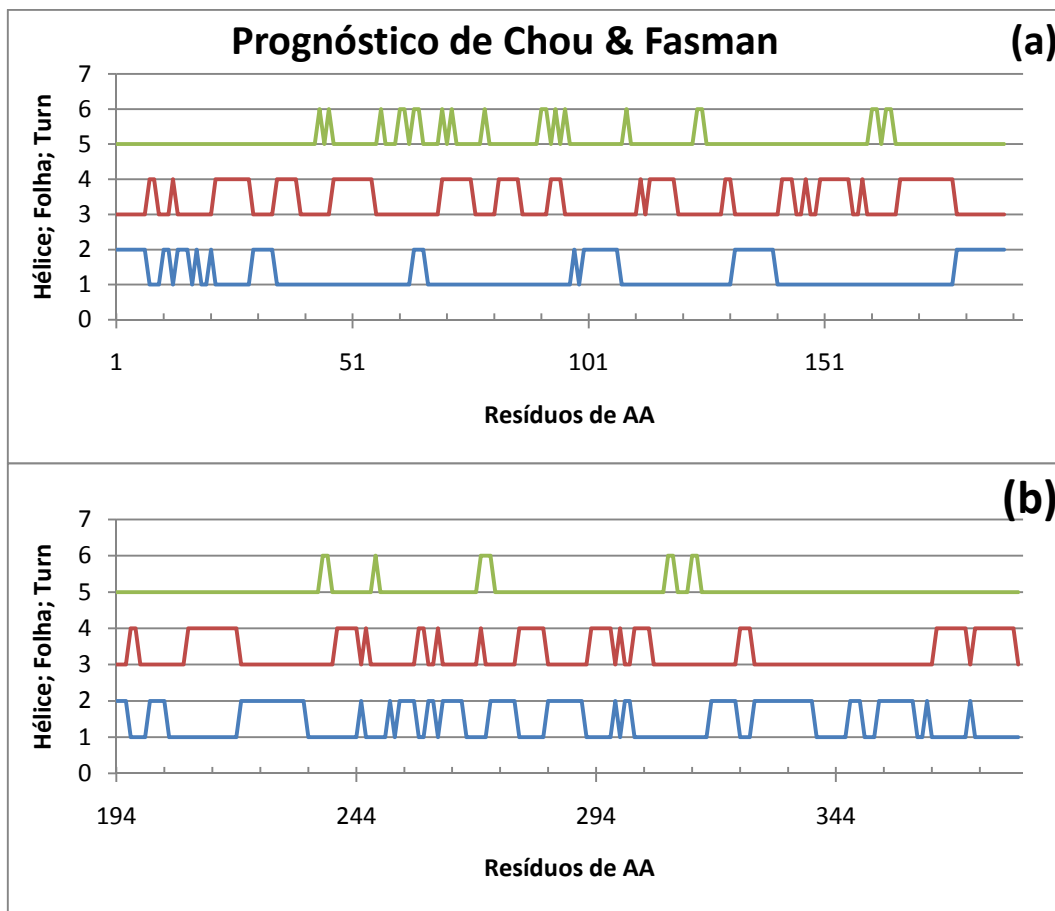
A antigenicidade da ovalbumina pode ser representada graficamente do seguinte modo:



**Gráfico 27:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a ovalbumina, usando a escala de Hopp & Wood (azul). (a) resíduos 1-193 e (b) resíduos 194-386. Zonas antigénicas possíveis (setas azuis).

Na previsão de epítomos (ligações antigénicas) (Gráfico 27), a ovalbumina parece possuir sete picos susceptíveis de serem regiões antigénicas. Pode afirmar-se que a ovalbumina é bastante reactiva no que diz respeito à sua antigenicidade, o que está de acordo com a bibliografia (Mine, 2008). Devido à dimensão da proteína, foi necessário dividir o gráfico em duas partes iguais (a e b), para se poder ver com mais detalhe todos os picos. Em a, estão representados os aminoácidos de 1 a 193 e em b de 194 a 386.

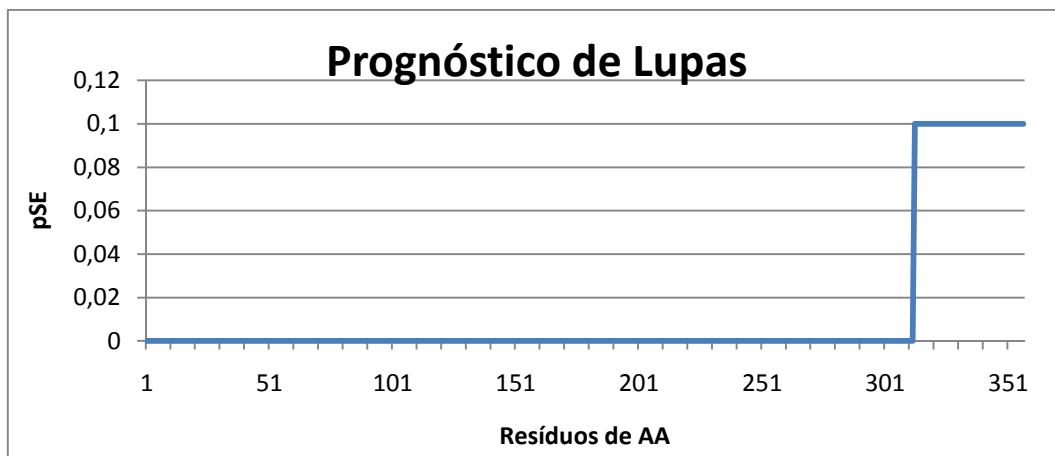
A previsão das hélices- $\alpha$ ; folhas- $\beta$  e voltas podem ser observadas através do gráfico 28



**Gráfico 28:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na ovalbumina. Cada pico corresponde à presença de um possível elemento estrutural. (a) resíduos 1-193 e (b) resíduos 194-383.

Observa-se que existe uma probabilidade de encontrar 19 hélices- $\alpha$  ao longo da cadeia polipeptídica da ovalbumina (Gráfico 28). Este facto confirma o pressuposto teórico em que a estrutura desta proteína mostra um anel central reactivo intacto na forma de uma hélice- $\alpha$  (Jedrychowski & Wichers, 2009). No que diz respeito às folhas- $\beta$ , observam-se, no mínimo, 21. No entanto, este facto não pode ser confirmado por dados teóricos. Em relação às voltas, observam-se pelo menos 10, contudo, com base na informação teórica, a ovalbumina possui três voltas (Jedrychowski & Wichers, 2009).

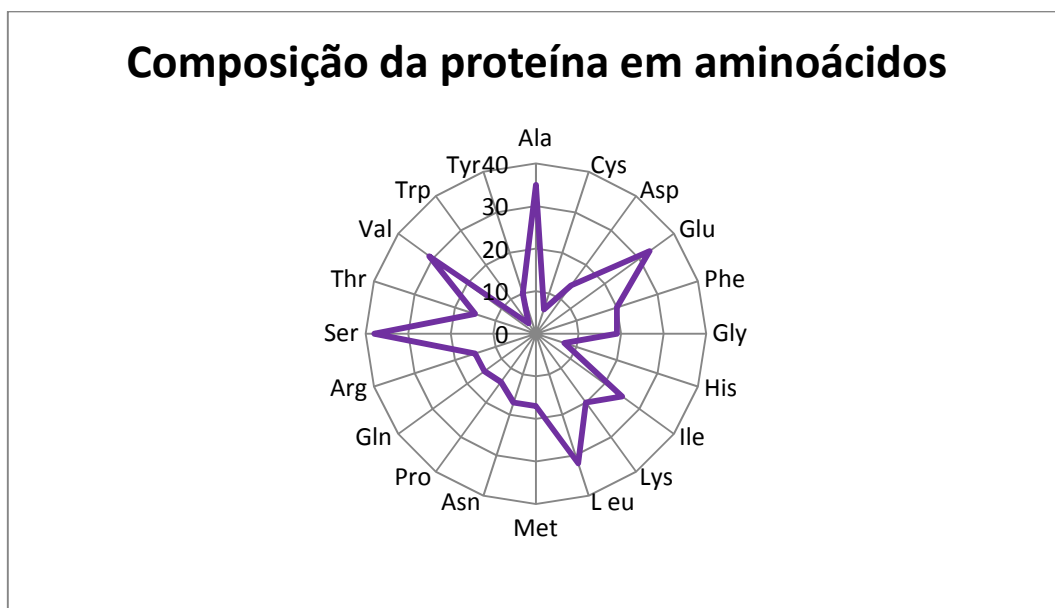
O prognóstico de Lupas para a ovalbumina é apresentado no gráfico 29.



**Gráfico 29:** Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na ovalbumina. A cada pico corresponder a possível presença dessa estrutura.

Observa-se que existe uma probabilidade de encontrar uma super-hélice- $\alpha$  na ovalbumina (Gráfico 29). A estrutura encontra-se a partir do resíduo 313 e possivelmente prolonga-se até ao final da cadeia.

A composição da ovalbumina pode ser representada por um gráfico em radar (Gráfico 30).

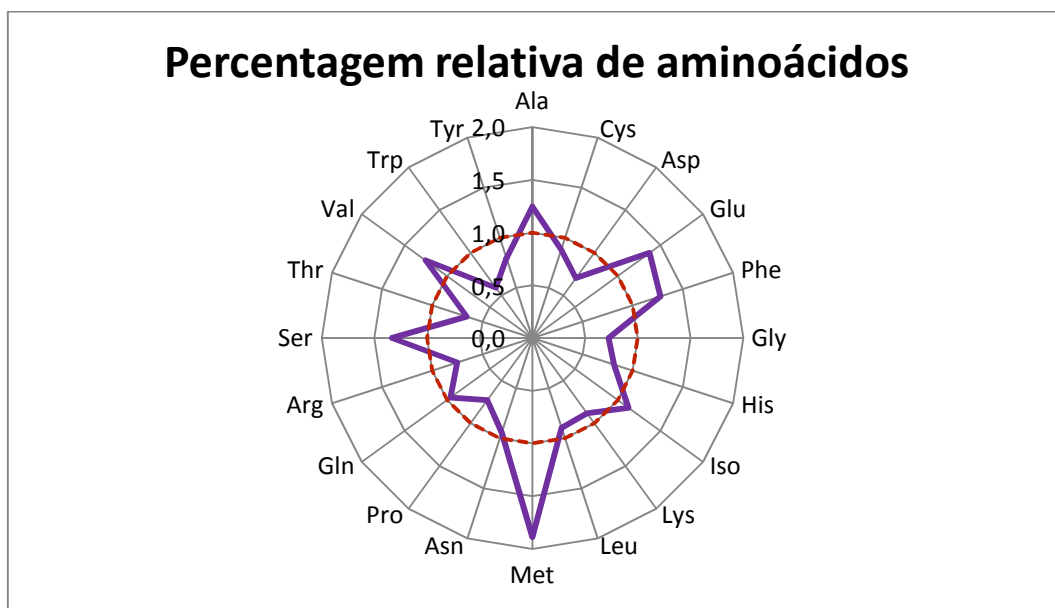


**Gráfico 30:** Representação gráfica da quantidade real de cada aminoácido na ovalbumina (386 resíduos de aminoácido).

A ovalbumina é constituída pelos vinte principais aminoácidos (Gráfico 30). Sendo assim, a ovalbumina é constituída maioritariamente por (ordem decrescente): serina; alanina; glutamato; leucina; valina; isoleucina; fenilalanina e lisina. Tem uma quantidade relativa de: glicina; metionina; asparagina; glutamina; arginina; treonina; aspartato; prolina e tirosina. A ovalbumina possui igualmente alguns aminoácidos residuais tais como: histidina; cisteína e

triptofano. Visto que a ovalbumina é rica em alanina; glutamato e leucina, isto confirma a existência de uma hélice- $\alpha$  na sua estrutura (Berg et al., 2002). No entanto, a grande quantidade de serina e valina pode perturbar a estabilidade dessa mesma hélice- $\alpha$ . O valor médio de isoleucina e o valor elevado de valina revelam também uma possível presença de folhas-B. A presença de voltas é igualmente confirmada devido à existência de valores mínimos de glicina, asparagina e prolina (Berg et al., 2002).

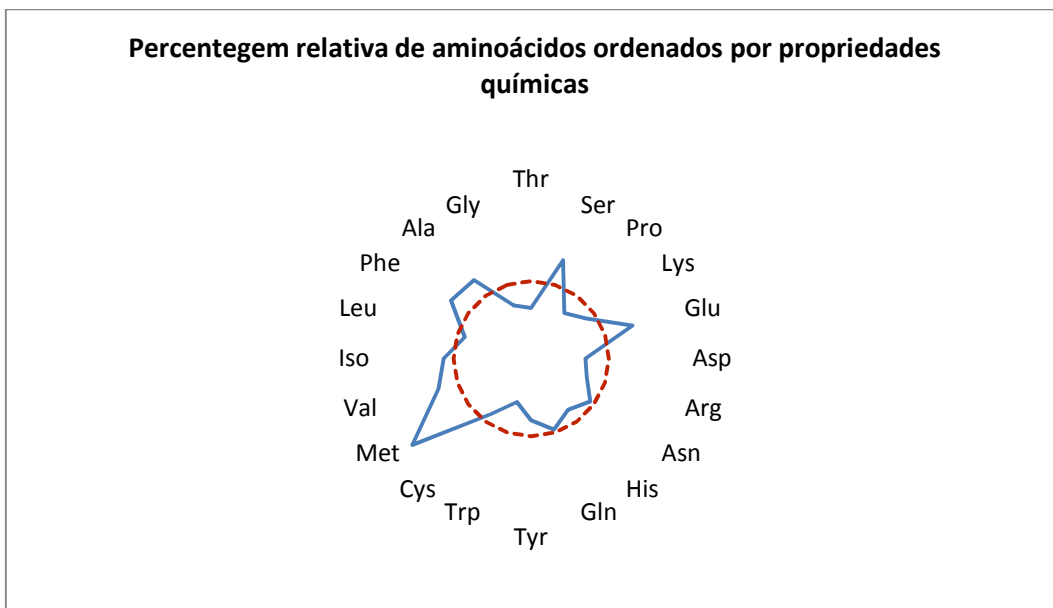
A comparação da ovalbumina e da GenPep#97 pode ser feita através das percentagens de aminoácidos que compõem cada uma delas (Gráfico 31).



**Gráfico 31:** Representação gráfica da percentagem de cada aminoácido na ovalbumina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína (386 resíduos de aminoácido).

Observa-se que a ovalbumina possui sete aminoácidos numa percentagem acima da média, relativamente à GenPep#97, sendo essa média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína e está representada a tracejado vermelho (Gráfico 31). Desses sete, os mais evidentes é a metionina, seguido de glutamato; serina; fenilalanina; valina; alanina e isoleucina. Os restantes aminoácidos têm um valor abaixo da média.

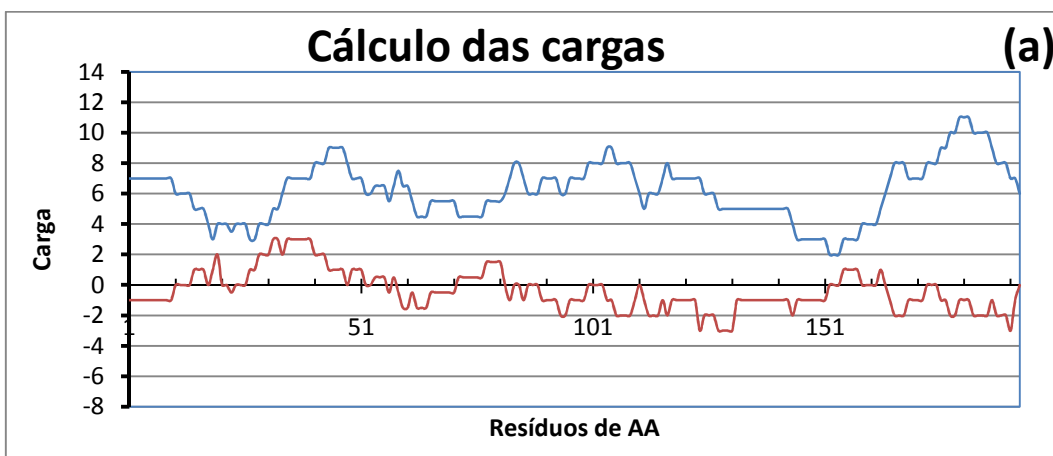
A distribuição dos aminoácidos segundo as suas propriedades químicas, revela algumas características da ovalbumina (Gráfico 32).

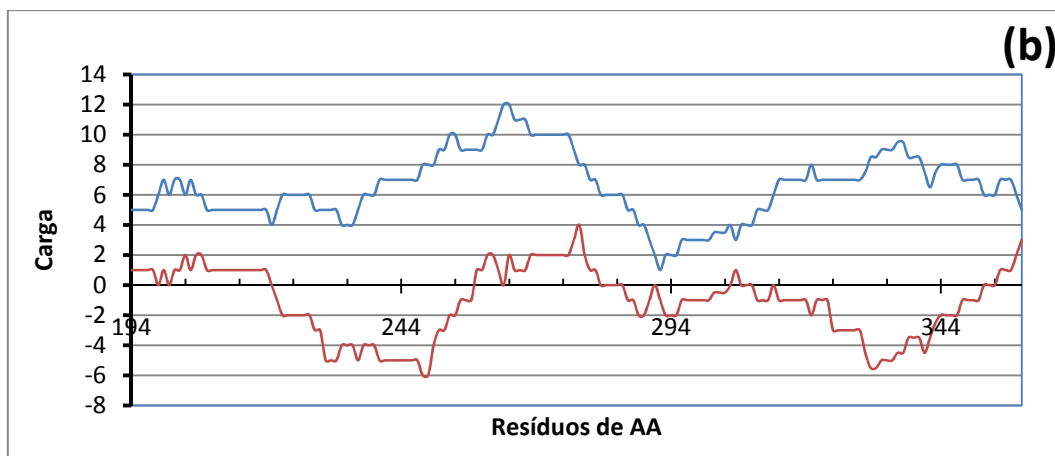


**Gráfico 32:** Representação gráfica da percentagem relativa dos aminoácidos da ovalbumina, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão à direita e os mais hidrofóbicos estão à esquerda. O círculo a tracejado vermelho representa a média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

Observa-se que a ovalbumina possui sete aminoácidos numa percentagem relativa superiores à média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína), representada a tracejado vermelho (Gráfico 32). Desses sete aminoácidos, cinco são hidrofóbicos: metionina; valina; isoleucina; fenilalanina e alanina. Dois são carregados: serina e glutamato. Esta composição sugere que a ovalbumina tem um comportamento hidrofóbico, ou pelo menos uma parte comporta-se de maneira hidrofóbica.

As cargas da ovalbumina podem ser observadas através do gráfico 33.





**Gráfico 33:** Representação gráfica das cargas da ovalbumina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos. (a) resíduos 1-193 e (b) resíduos 194-359.

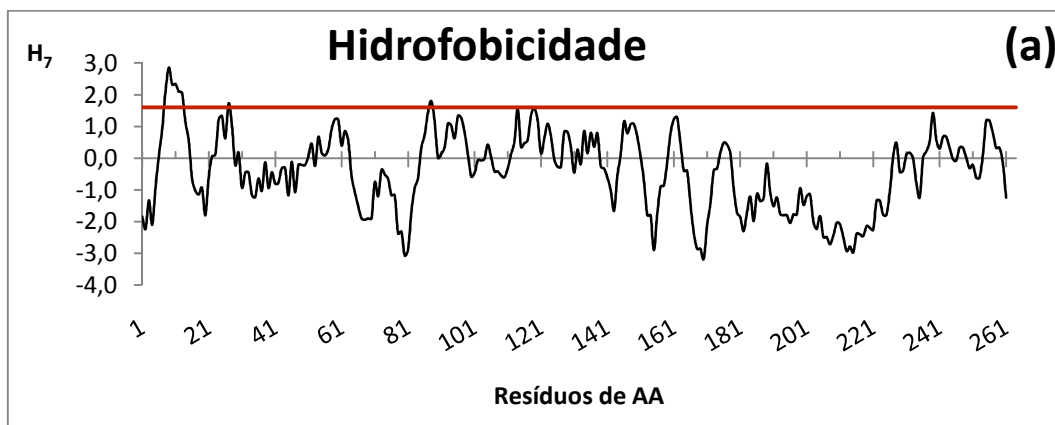
Observa-se que a ovalbumina possui cargas reais positivas entre os intervalos 27-47 (Gráfico 33a), 201-222 (Gráfico 33b) e 258-273 (Gráfico 33b). Existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores (Gráfico 33).

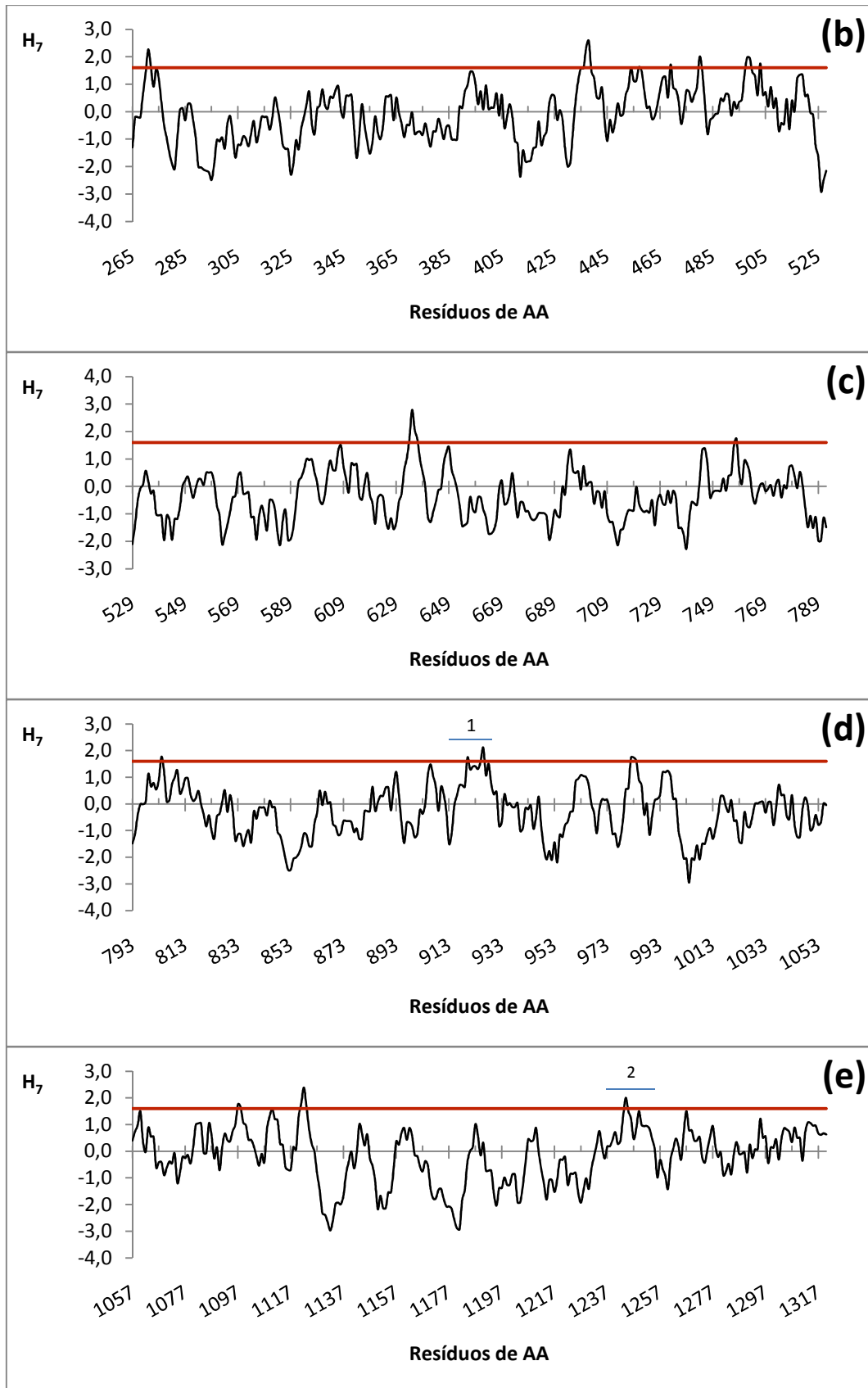
Observa-se igualmente que as cargas absolutas da ovalbumina mostram que ela tem um comportamento polar positivo ao longo de toda a cadeia polipeptídica, essencialmente entre os intervalos: 155-200 (Gráfico 33a) e 228-281 (Gráfico 33b), onde os valores das cargas são mais elevados. Isto significa que a ovalbumina é sensível a corrente eléctrica contínua.

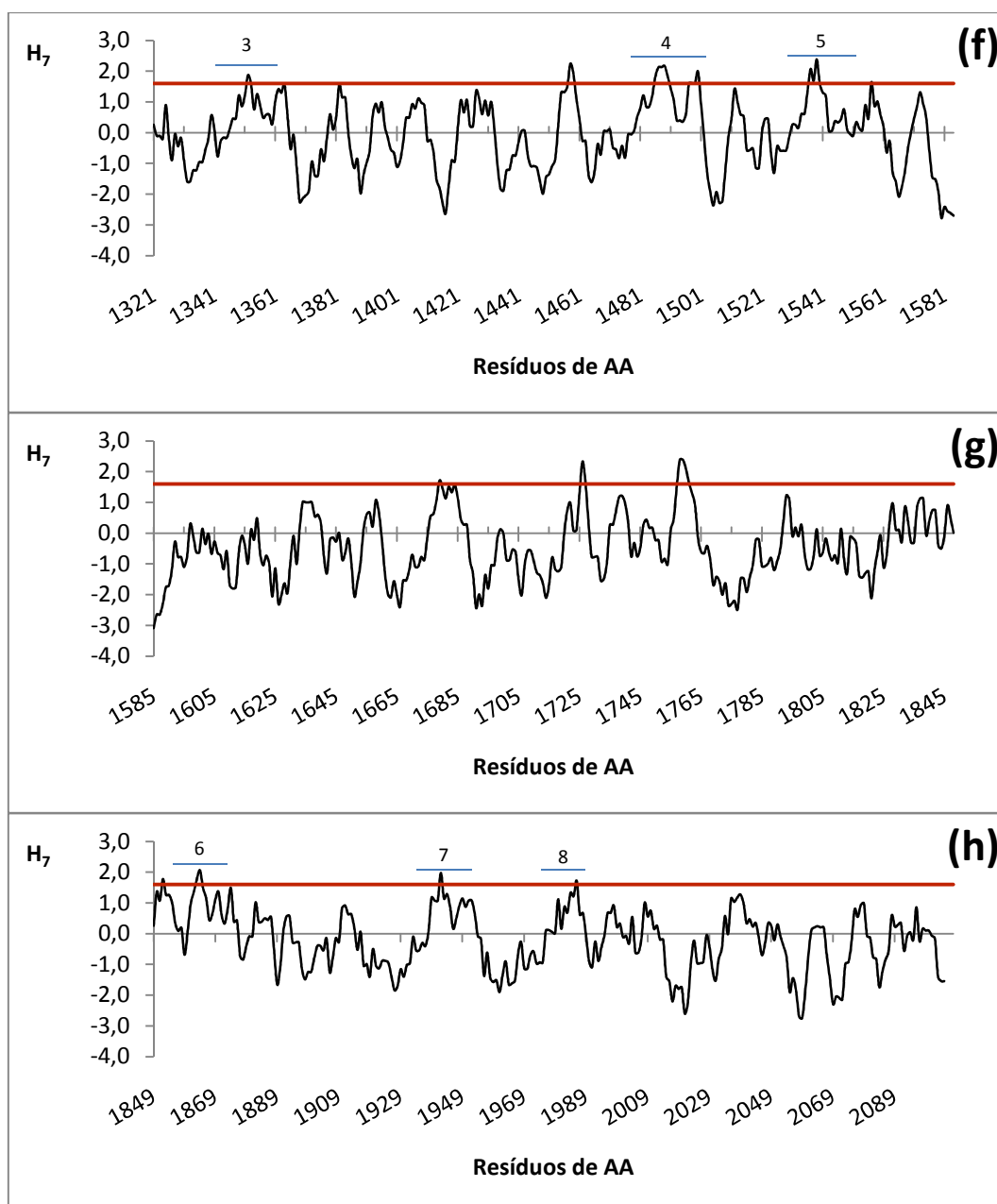
#### 4.2.2.3. Ovomucina

Devido ao facto da subunidade  $\alpha$  da ovomucina ser constituída por 2108 resíduos de aminoácidos, foi necessário dividir em oito partes os diferentes gráficos ( $H_7$ ,  $H_\alpha$  e  $H_{Ag}$ ). A subunidade B da ovomucina é constituída por 827 resíduos de aminoácidos, o que dá um total de 2935 resíduos de aminoácidos.

A hidrofobicidade da ovomucina (subunidade  $\alpha$ ) pode ser demonstrada através do gráfico 34.





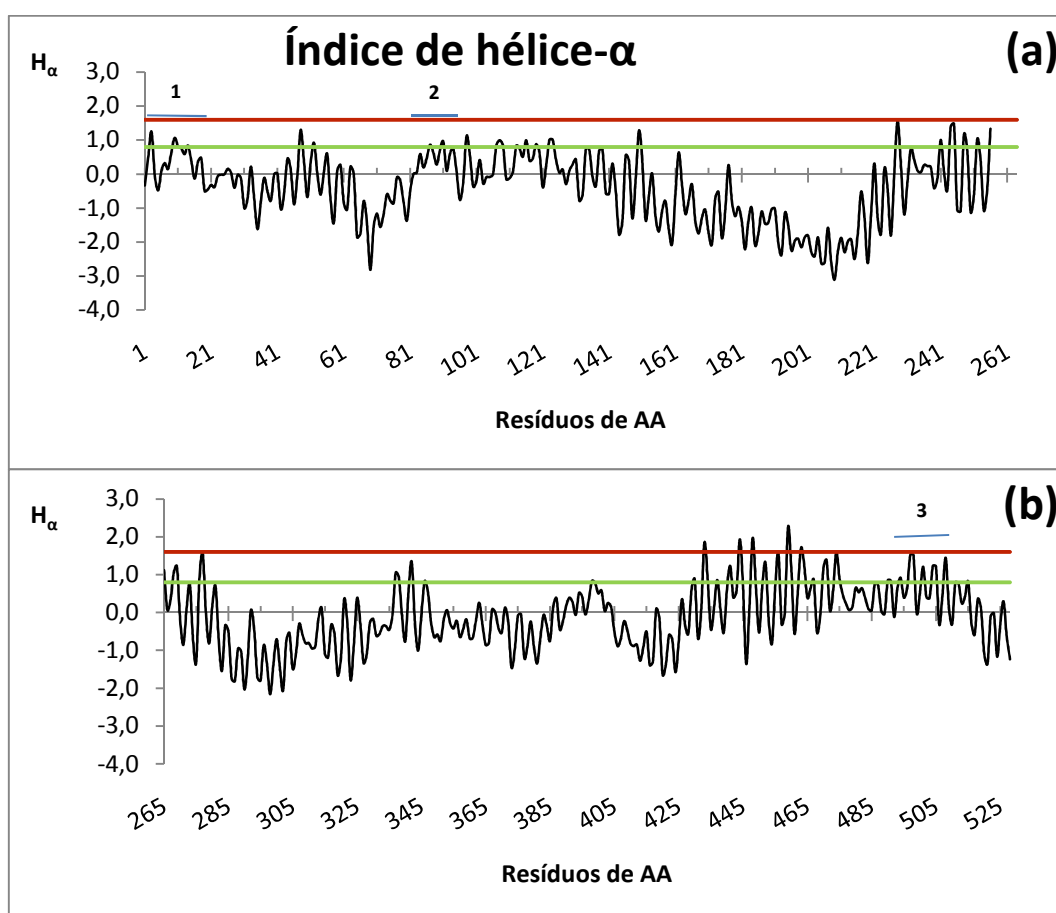


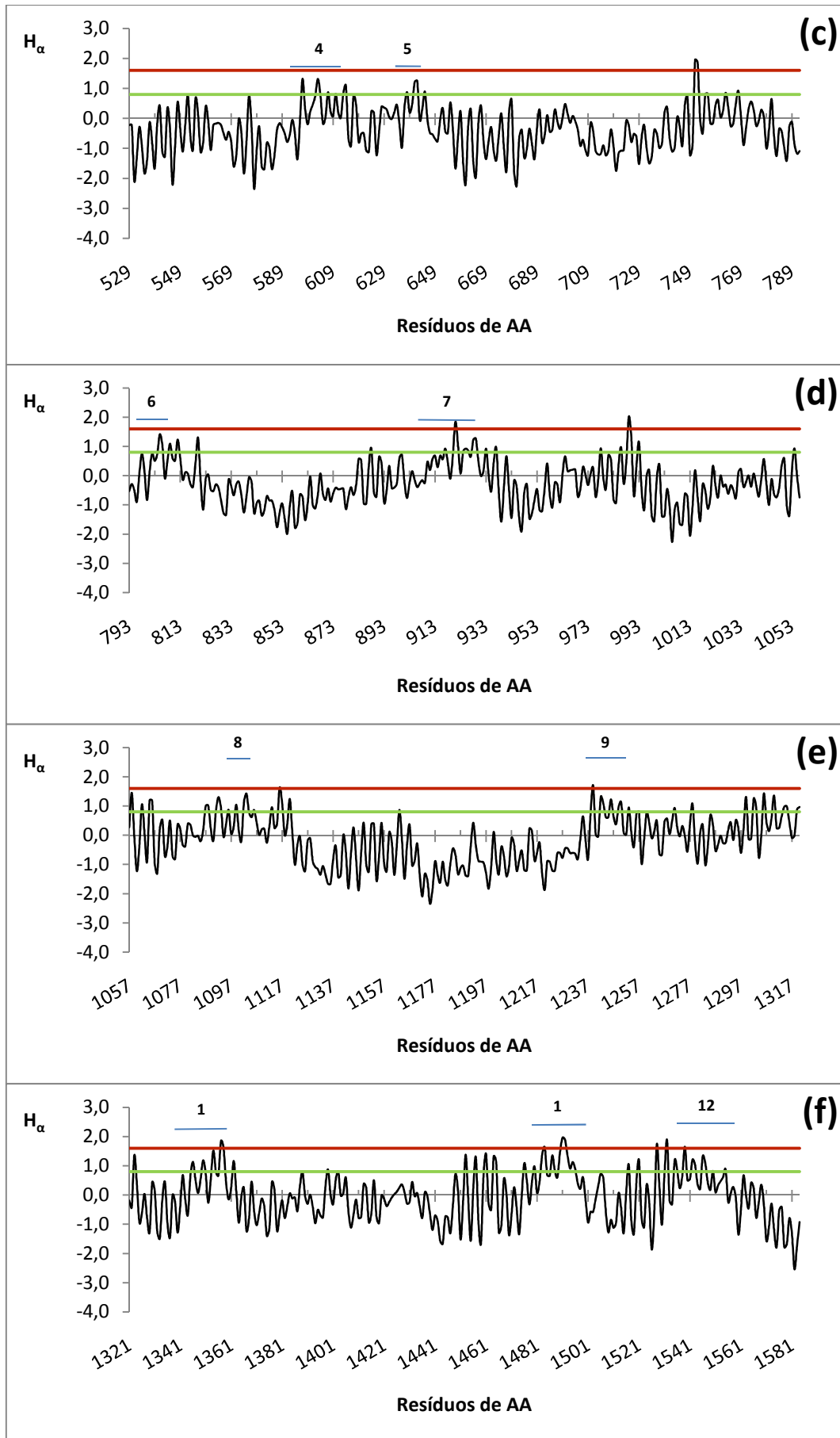
**Gráfico 34:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a subunidade  $\alpha$  da ovomucina (2108 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho). (a) resíduos 1-264; (b) resíduos 265-528; (c) resíduos 529-792; (d) resíduos 793-1056; (e) resíduos 1057-1320; (f) resíduos 1321-1584; (g) resíduos 1585-1848 e (h) resíduos 1849-2108.

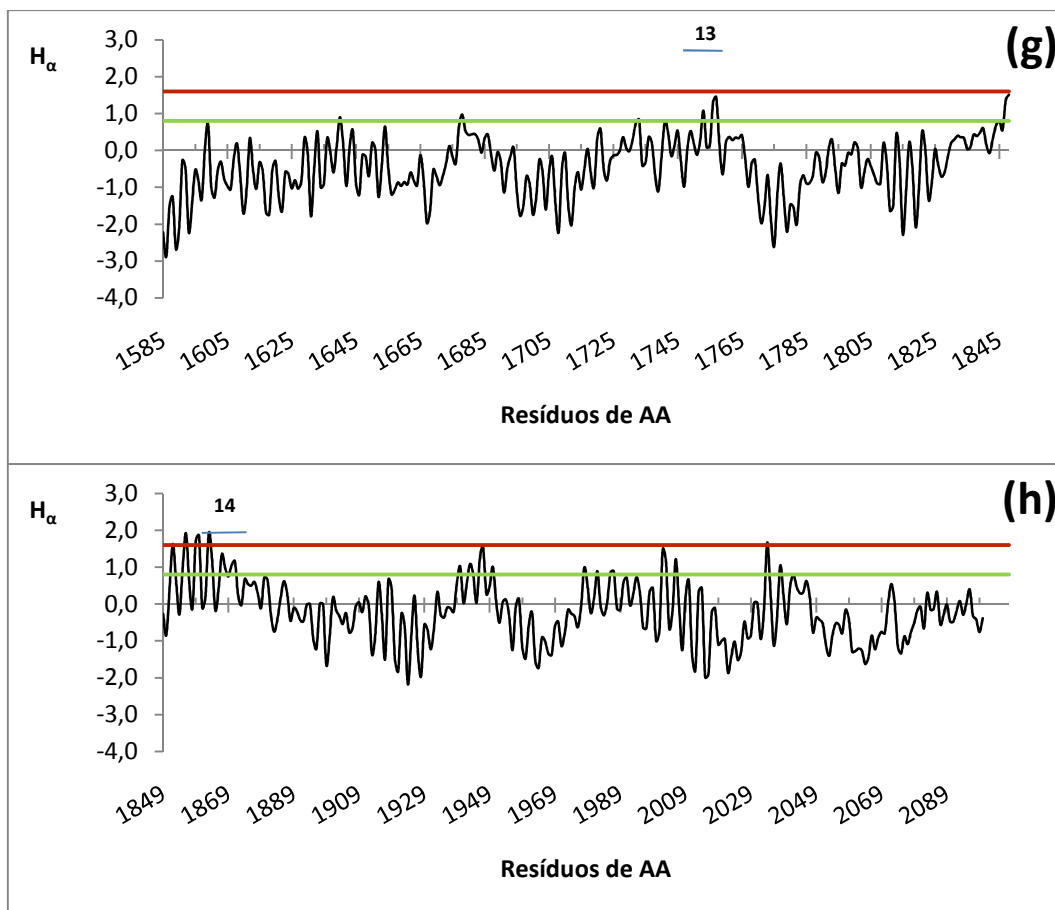
Entre os resíduos 913 e 934 (Gráfico 34d), observa-se um pico (1), correspondente a uma zona hidrofóbica. Apesar de haver apenas três médias de sete resíduos neste intervalo com um grau de hidrofobicidade superior ou igual a 1,6 são necessários 21 resíduos para efectuar os cálculos que permitem desenhar este pico. Como a vizinhança próxima também tem um comportamento hidrofóbico, apesar de ser inferior a 1,6 considera-se uma zona hidrofóbica, visto que possui mais de 19 resíduos de aminoácidos. Entre os resíduos 1237 e 1255 (Gráfico 34e), observa-se um segundo pico (2) correspondente a uma zona hidrofóbica. Mais uma vez, apesar de haver apenas um ponto com uma média de sete resíduos neste

intervalo, com um grau de hidrofobicidade superior a 1,6 são necessários 24 resíduos para efectuar os cálculos que permitem desenhar este pico. Como a vizinhança próxima também tem um comportamento hidrofóbico, apesar de ser inferior a 1,6 considera-se uma zona hidrofóbica, visto que possui mais de 19 resíduos de aminoácidos. O mesmo tipo de raciocínio pode aplicar-se aos picos situados nos intervalos de aminoácidos: 1346-1365, 1479-1501 e 1531-1550, correspondentes aos picos 3; 4 e 6 (Gráfico 34f); 1860-1876, 1938-1953 e 1976-1989, correspondentes aos picos 6; 7 e 8 (Gráfico 34h). Sabe-se que a ovomucina é composta por uma fracção solúvel e uma fracção insolúvel (Jedrychowski & Wichers, 2009; Mine, 2008). O gráfico 34 mostrar uma previsão de oito zonas hidrofóbicas. Deste facto pode deduzir-se que a fracção  $\alpha$  seja a fracção insolúvel.

O índice de hélice- $\alpha$  pode ser representado no gráfico 35.



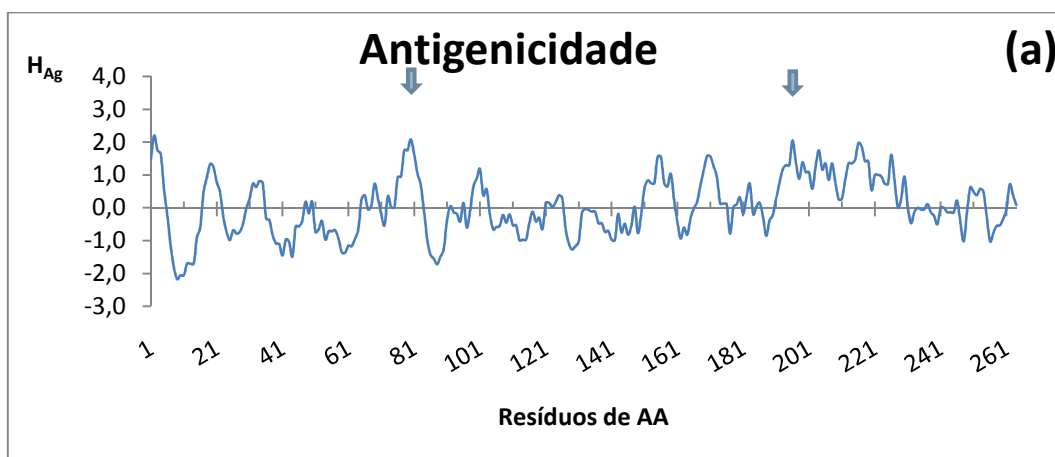


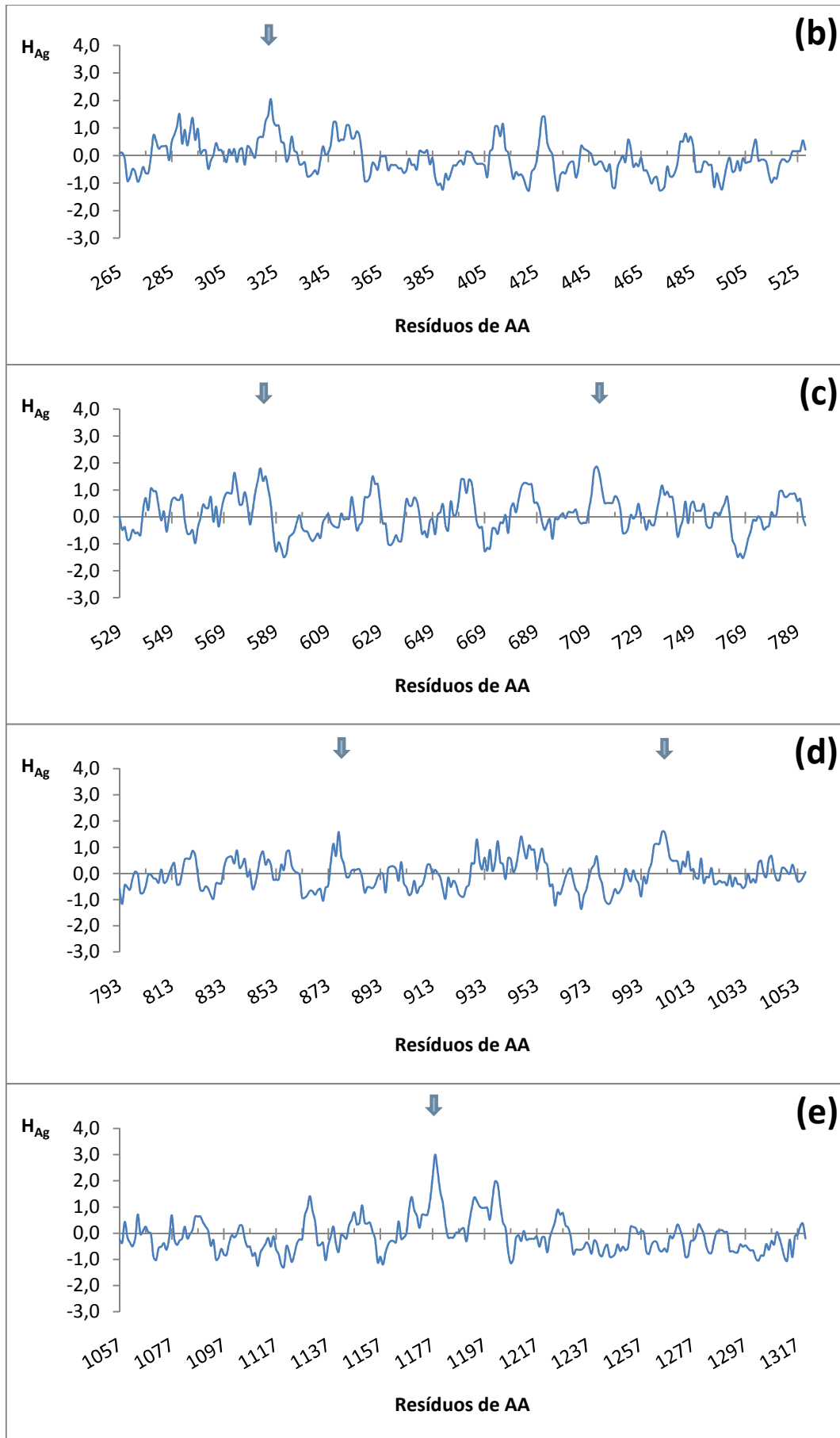


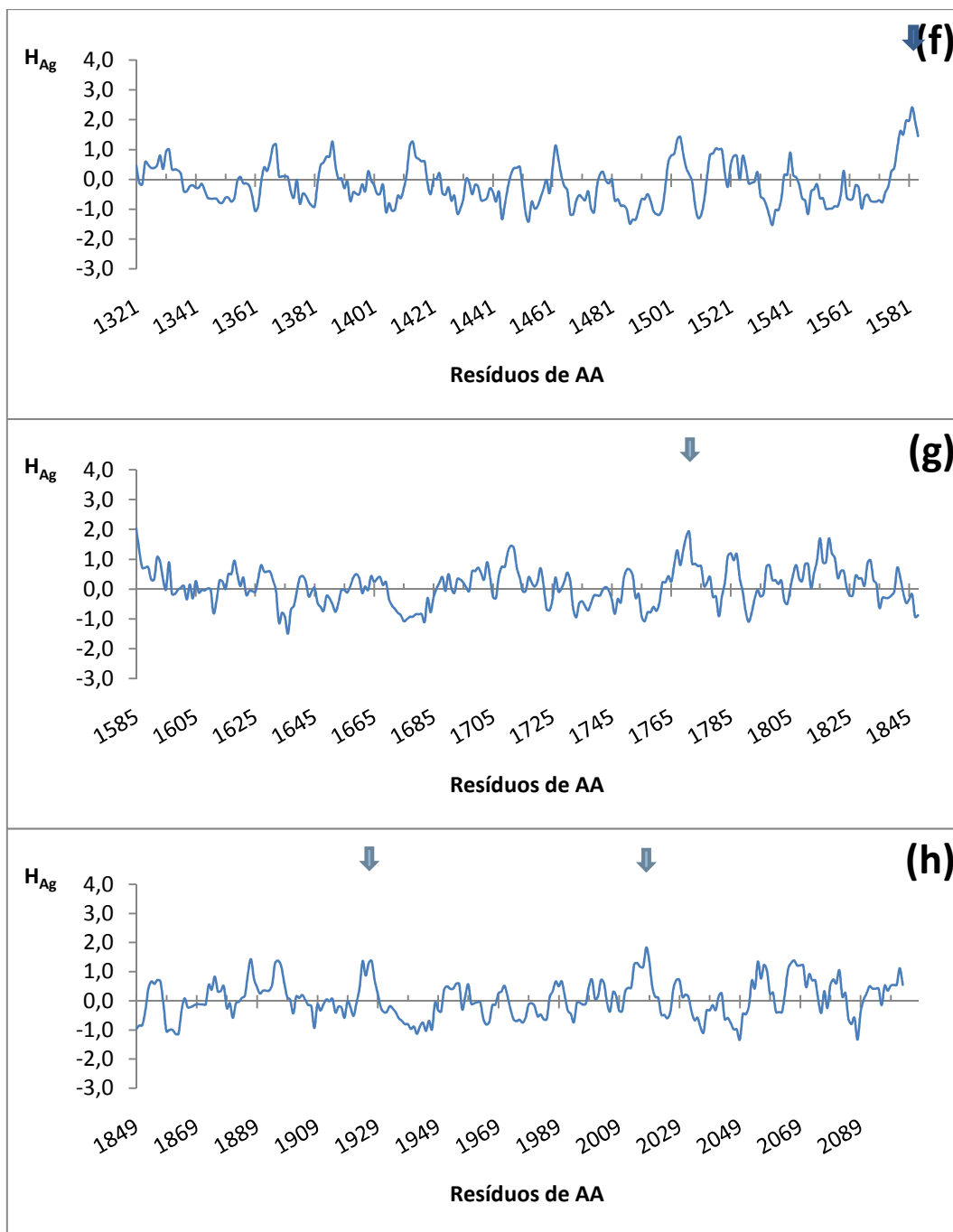
**Gráfico 35:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a ovomucina (2108 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro). (a) resíduos 1-264; (b) resíduos 265-528; (c) resíduos 529-792; (d) resíduos 793-1056; (e) resíduos 1057-1320; (f) resíduos 1321-1584; (g) resíduos 1585-1848 e (h) resíduos 1849-2108.

Observa-se que no índice de hélice- $\alpha$  constam 14 picos (Gráfico 35) que indicam a possível presença de hélices- $\alpha$  anfifílicas. Este tipo de hélice é caracterizado por possuírem resíduos de aminoácido alternadamente hidrofóbicos e hidrofílicos a cada três ou quatro monómeros (Murray *et al.*, 2002).

A antigenicidade da ovomucina pode ser representada segundo o gráfico 36.





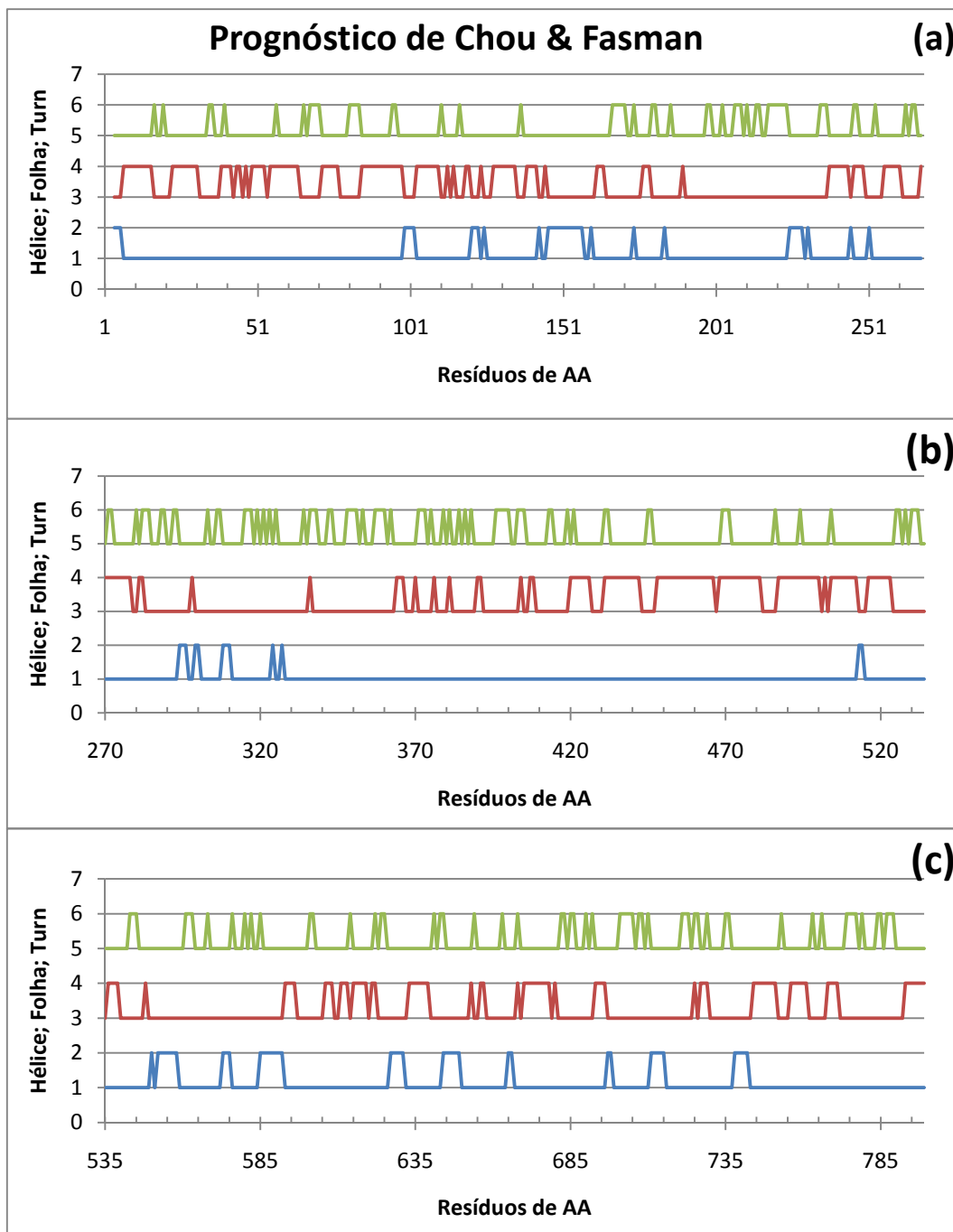


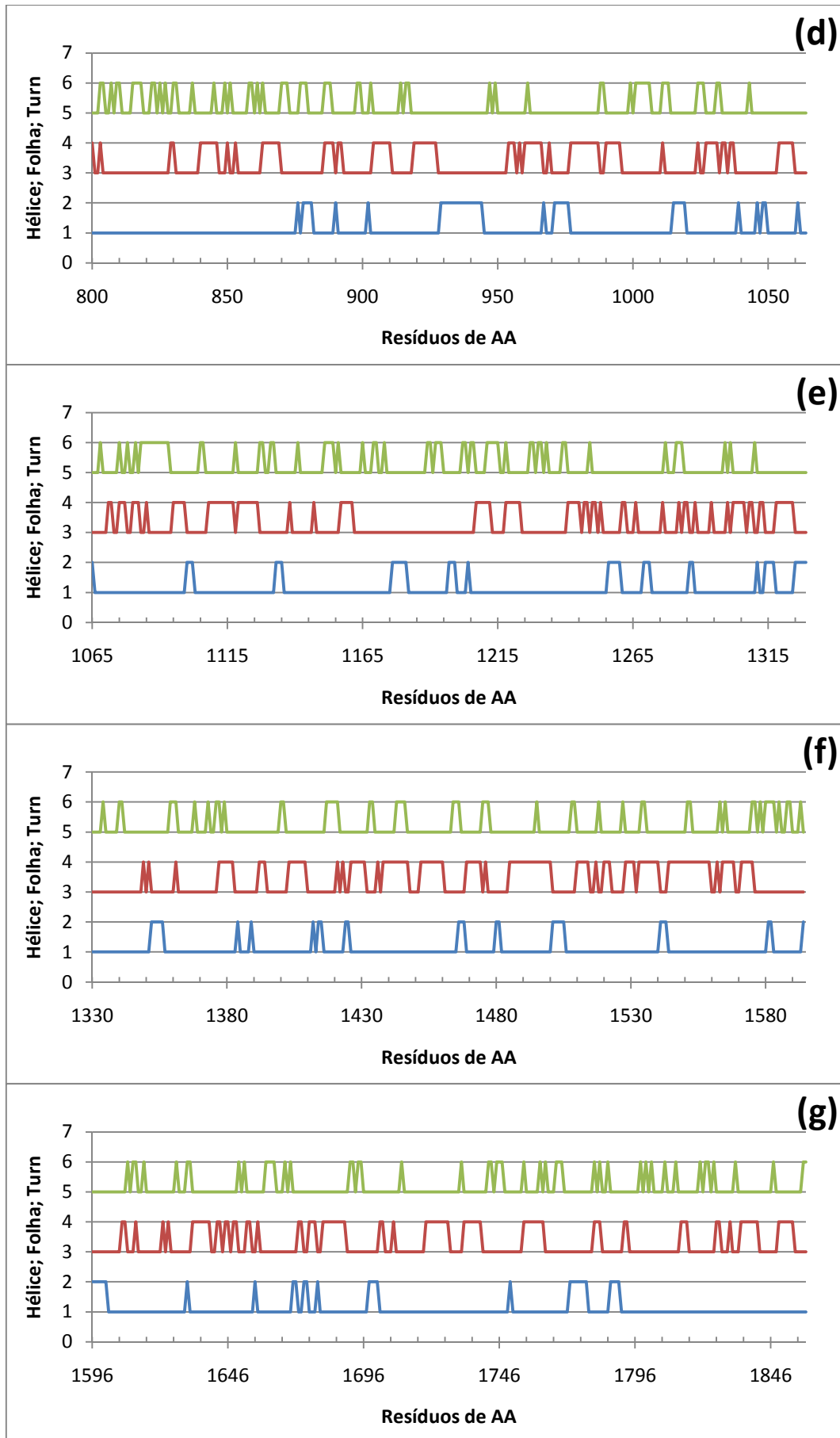
**Gráfico 36:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a ovomucina, usando a escala de Hopp & Wood (azul). (a) resíduos 1-264; (b) resíduos 265-528; (c) resíduos 529-792; (d) resíduos 793-1056; (e) resíduos 1057-1320; (f) resíduos 1321-1584; (g) resíduos 1585-1848 e (h) resíduos 1849-2108. Zonas antigénicas possíveis (setas azuis).

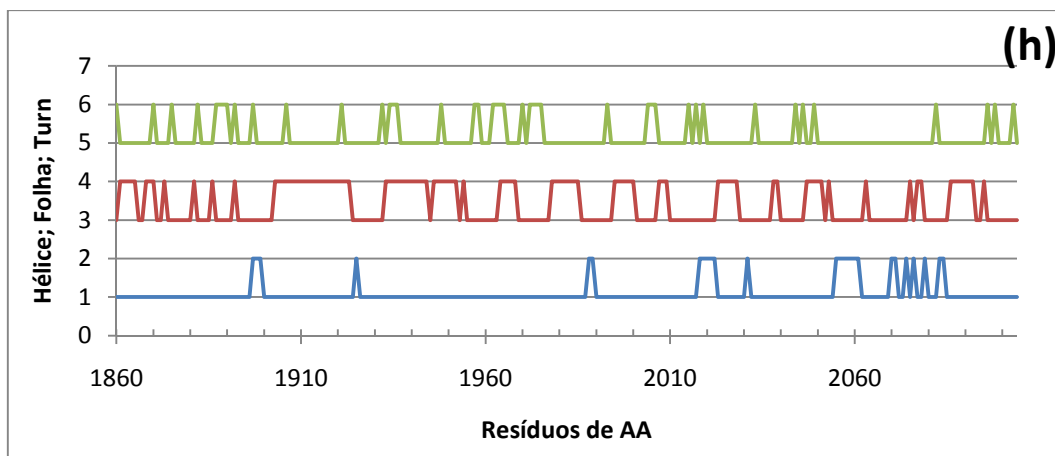
Ao nível dos antígenos (Gráfico 36), a ovomucina mostra ter doze regiões antigénicas. É certo que se trata de uma previsão, pois a alergenicidade da ovomucina é considerada minoritária quando comparada com outros alérgenos como a ovomucóide e a ovalbumina (Jedrychowski & Wichers, 2009).

Observa-se também que a maioria destes picos corresponde às zonas mais hidrofílicas desta proteína, visto que são regiões externas.

O gráfico 37 revela as previsões do Método de Chou & Fasman para os três elementos principais da estrutura secundária. Para uma melhor leitura e visto que a ovomucina é uma proteína de grande dimensão, o gráfico de Chou & Fasman foi dividido em oito partes (a; b; c; d; e; f; g e h).



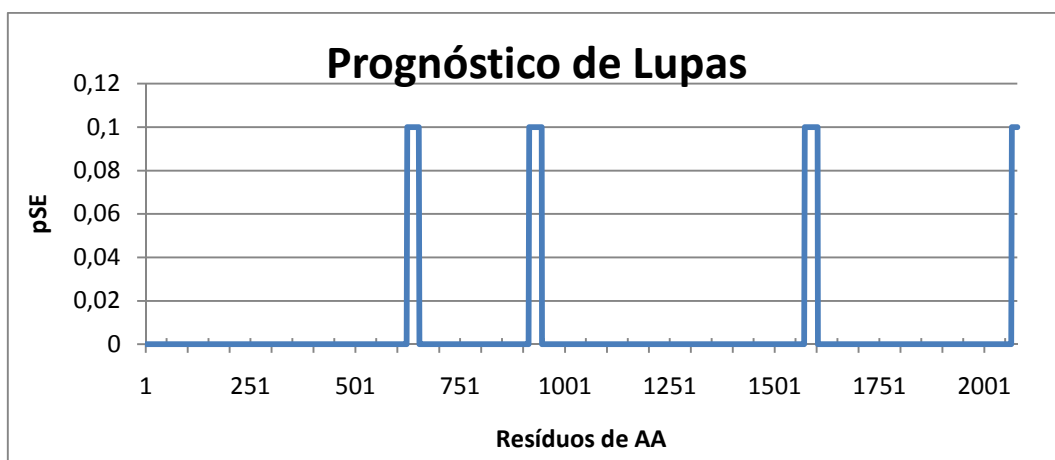




**Gráfico 37:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na ovomucina. Cada pico corresponde à presença de um possível elemento. (a) resíduos 1-269; (b) resíduos 270-534; (c) resíduos 535-799; (d) resíduos 800-1064; (e) resíduos 1065-1329; (f) resíduos 1330-1595; (g) resíduos 1596-1859 e (h) resíduos 1860-2104.

Observa-se que existe uma grande probabilidade de encontrar hélices- $\alpha$  ao longo da cadeia polipeptídica da ovomucina (Gráfico 37). O mesmo se pode afirmar relativamente às voltas e folhas- $\beta$ .

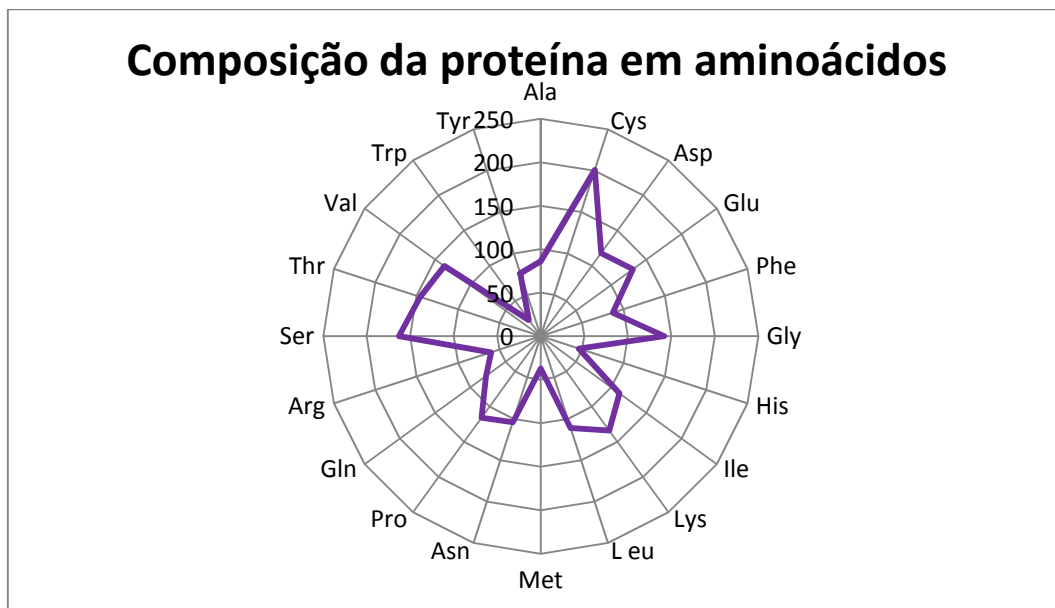
O prognóstico de Lupas é apresentado no gráfico 38.



**Gráfico 38:** Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na ovomucina.

Observa-se que existe a probabilidade de encontrar quatro super-hélice- $\alpha$  na ovomucina (Gráfico 38). Uma primeira entre os resíduos 626-652, uma segunda entre os resíduos 915-945 e uma terceira entre os resíduos 1572-1603. Há fortes indícios para que haja uma quarta estrutura a partir dos resíduos 2066.

A composição em aminoácidos da ovomucina pode ser apresentada num gráfico em radar (Gráfico 39).

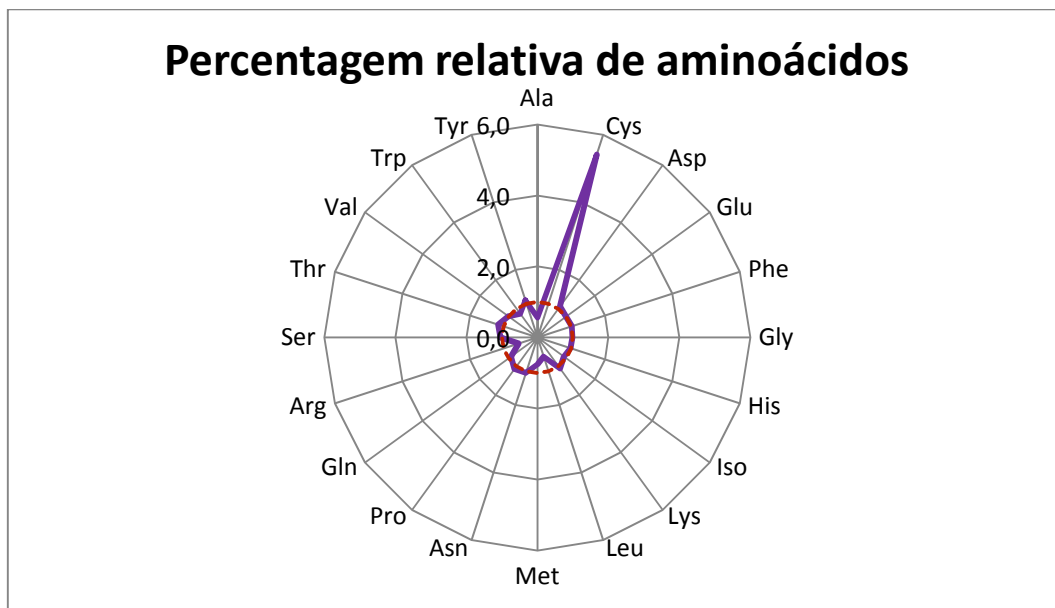


**Gráfico 39:** Representação gráfica da quantidade real de cada aminoácido na ovomucina (2108 resíduos de aminoácido).

A ovomucina é constituída pelos vinte aminoácidos principais (Gráfico 39). Sendo assim, a ovomucina é constituída maioritariamente por (ordem decrescente): cisteína; serina; glicina; treonina; valina; lisina; glutamato; aspartato; prolina; isoleucina; leucina e asparagina. Possui valores mais baixos de: alanina; fenilalanina; glutamina; tirosina; arginina; histidina; metionina e triptofano.

Visto que a ovomucina é relativamente rica em glutamato e leucina, existe uma grande probabilidade para que ela tenha pelo menos uma hélice- $\alpha$  na sua estrutura, visto que estes dois aminoácidos são frequentemente encontrados neste tipo de elementos (Berg et al., 2002). O elevado número de resíduos de isoleucina e valina revela também a possível presença de folhas- $\beta$ . A presença de voltas é igualmente esperada devido à existência de glicina; asparagina e prolina (Berg et al., 2002).

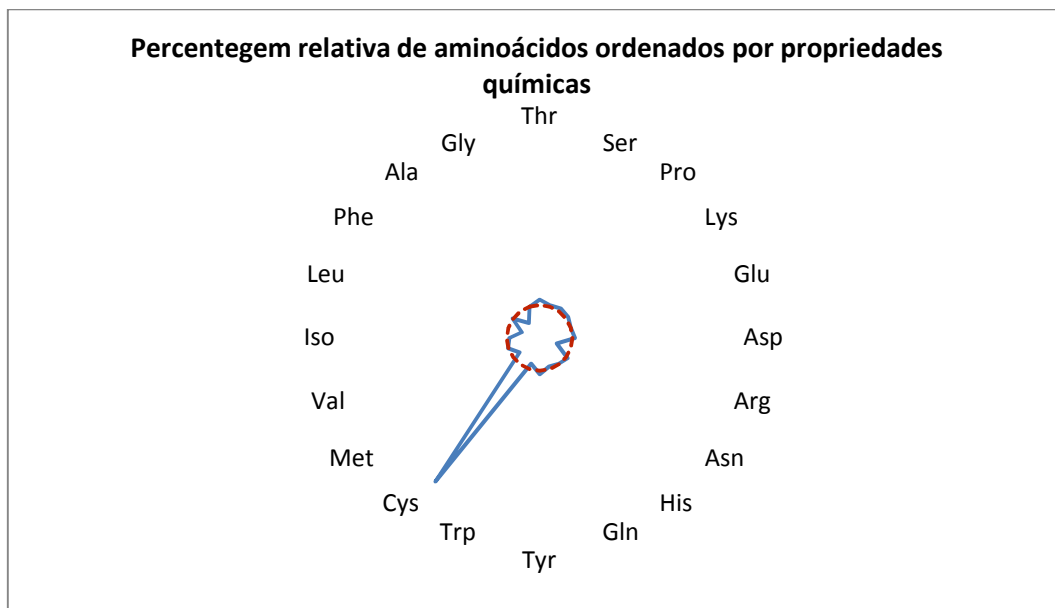
O gráfico 40 mostra a relatividade entre a percentagem de aminoácidos da ovomucina e da GenPep#97.



**Gráfico 40:** Representação gráfica da percentagem de cada aminoácido na ovomucina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína (2108 resíduos de aminoácido).

Observa-se que a ovomucina possui dez aminoácidos com uma percentagem relativa acima da média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína), representada a tracejado vermelho (Gráfico 40). Desses dez, o mais abundante é a cisteína, seguida da treonina; tirosina; prolina; lisina; aspartato; asparagina; serina; fenilalanina e valina. Os restantes aminoácidos apresentam-se num número abaixo da média.

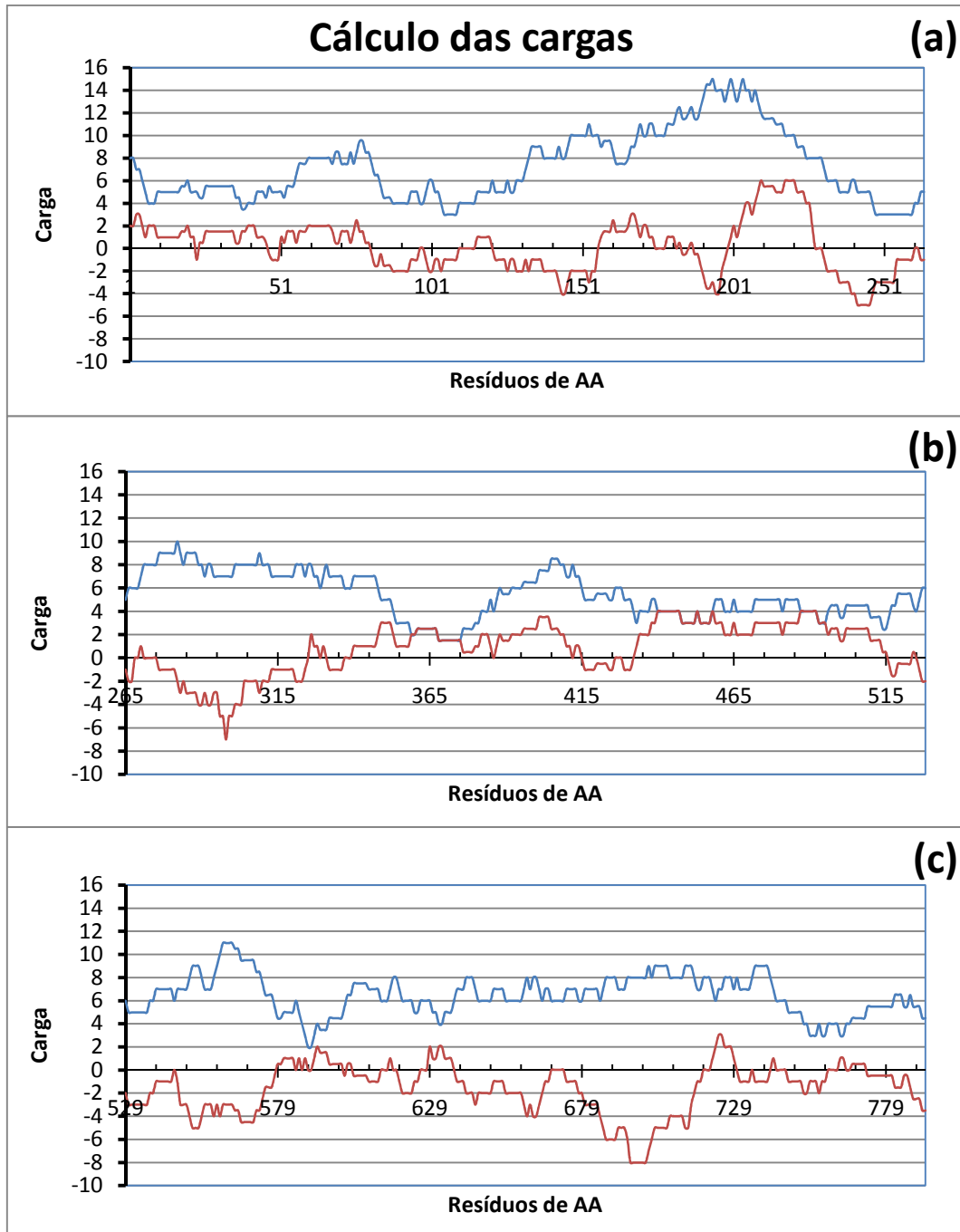
O gráfico 41 representa a distribuição dos aminoácidos mediante as suas propriedades químicas.

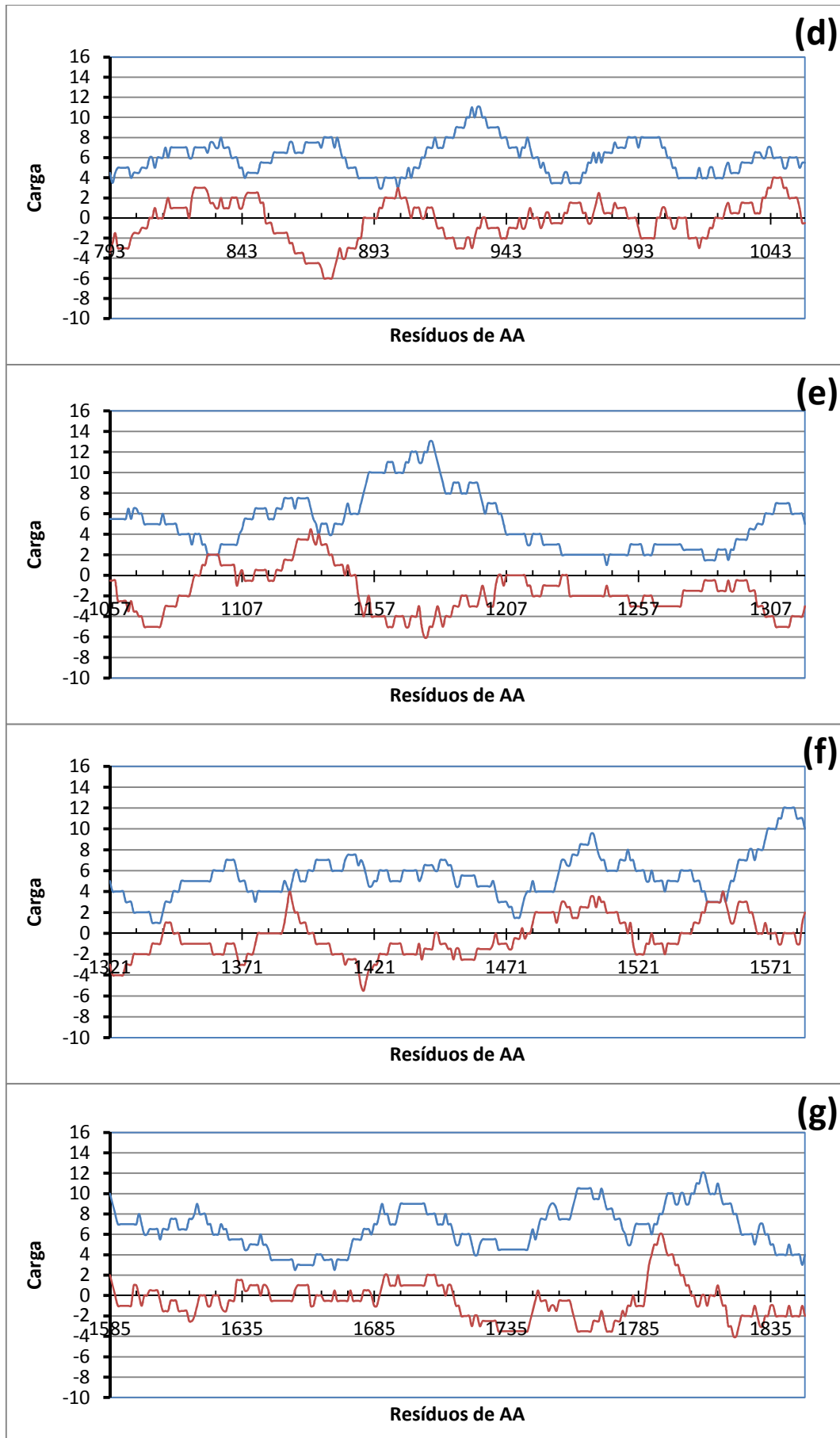


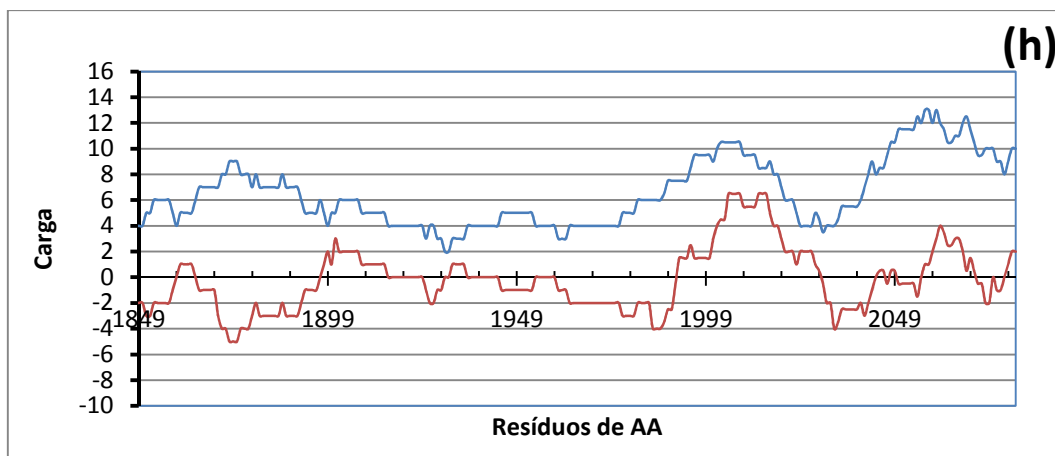
**Gráfico 41:** Representação gráfica da percentagem relativa dos aminoácidos da ovomucina, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão situados à direita e os mais hidrofóbicos estão situados à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

A ovomucina mostra claramente um comportamento hidrofóbico devido à grande percentagem relativa de cisteína (aminoácido hidrofóbico), que a compõe (Gráfico 41). Tendo em conta que a ovomucina é composta por duas subunidades, uma subunidade solúvel e uma subunidade insolúvel (Jedrychowski & Wichers, 2009; Mine, 2008), confirma-se que esta subunidade (fracção  $\alpha$ ) é insolúvel. Não se pode desprezar as quantidades acima da média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína e representada a tracejado vermelho) de aminoácidos hidrofílicos (carregados) tais como a prolina; lisina; aspartato; asparagina e serina. A razão pela qual a percentagem relativa destes aminoácidos não é tão flagrante ao nível do gráfico, deve-se ao facto da ovomucina (fracção  $\alpha$ ) ter uma cadeia muito extensa. A treonina e a tirosina, situam-se no limiar das duas propriedades, (hidrofóbico, carregados) e por essa razão não devem interferir no comportamento global da proteína. A fenilalanina e a valina, também estão acima da média e favorecem a hidrofobicidade.

As cargas absolutas e reais para a ovomucina estão representadas no gráfico 42.







**Gráfico 42:** Representação gráfica das cargas da ovomucina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos. (a) resíduos 1-264; (b) resíduos 265-528; (c) resíduos 529-792; (d) resíduos 793-1056; (e) resíduos 1057-1320; (f) resíduos 1321-1584; (g) resíduos 1585-1848; (h) resíduos 1849-2081.

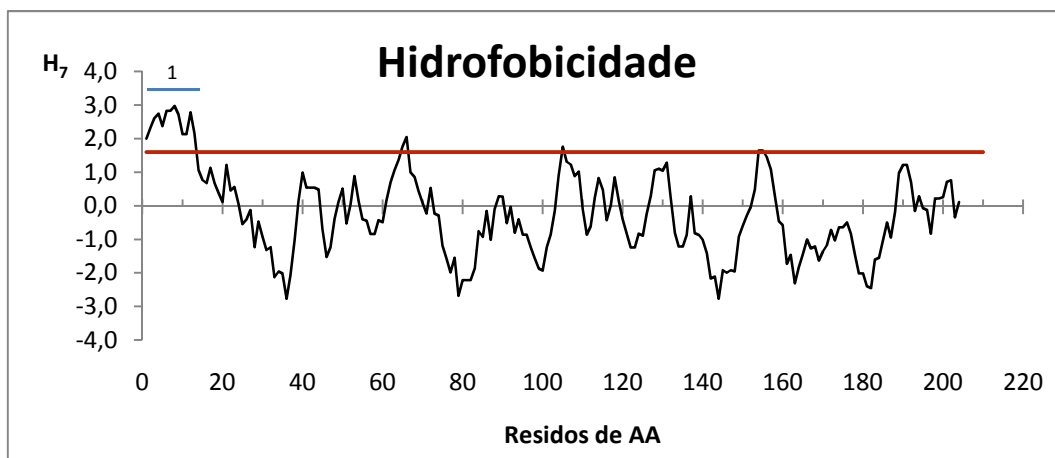
Observa-se que a ovomucina possui cargas reais positivas entre os intervalos 200-227 (Gráfico 42a), 433-515 (Gráfico 42b) 1120-1145 (Gráfico 42e), 1481-1516 (Gráfico 42f), 1788-1804 (Gráfico 42g) e 1992-2029 (Gráfico 42h). Existem outros intervalos, onde as cargas reais também são positivas, mas são menores.

Observa-se igualmente que as cargas absolutas da ovomucina mostram que ela tem um comportamento polar positivo ao longo de toda a cadeia polipeptídica. Destaca-se no entanto um aumento dessas mesmas cargas nos intervalos: 116-239 (Gráfico 42a), 269-352 (Gráfico 42b), 907-958 (Gráfico 42d), 1145-1206 (Gráfico 42e), 1555-1645 (Gráfico 42f, g), 1783-1829 (Gráfico 42g) e 2035-2081 (Gráfico 42h) Isto significa que se a subunidade  $\alpha$  da ovomucina for colocada numa solução com uma corrente eléctrica, ela deslocar-se-á para o cátodo, (pólo negativo).

È de notar igualmente, que em algumas zonas as cargas reais e as cargas absolutas confundem-se. É o que acontece nos intervalos: 359-375, 440-458 e 487-495 (Gráfico 42b); 1094-1098 (Gráfico 42e); 1547-1554 (Gráfico 42f), assim como alguns pontos isolados: 902 (Gráfico 42d), 1136 (Gráfico 42e) e 1389 (Gráfico 42f).

#### 4.2.2.4. Ovomucóide

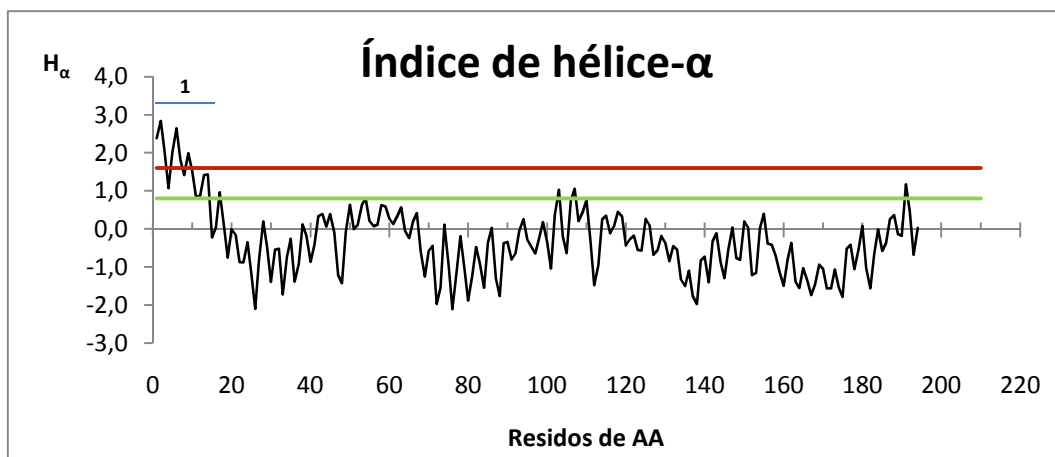
O cálculo de  $H_7$  para a ovomucóide é apresentado no gráfico 43.



**Gráfico 43:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a ovomucóide (210 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).

Com exceção do pico 1 que mostra uma zona bastante hidrofóbica, a ovomucóide apresenta um comportamento bastante hidrofílico ao longo da sua cadeia polipeptídica (Gráfico 43).

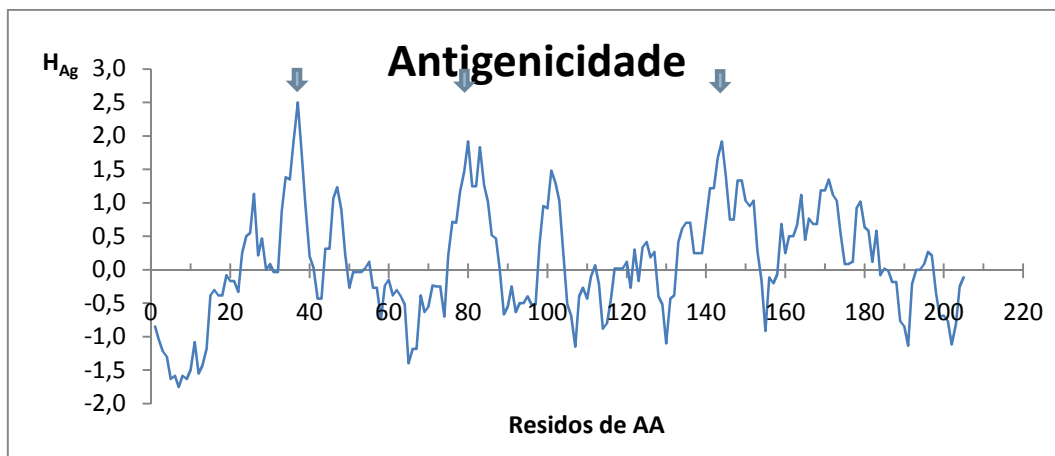
No gráfico 44, observa-se o resultado dos cálculos para o índice de hélice- $\alpha$ .



**Gráfico 44:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a ovomucóide (210 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

O gráfico do índice de hélice- $\alpha$  (Gráfico 44) mostra igualmente um pico inicial hidrofóbico (1), certamente representativo de uma hélice- $\alpha$  anfifílica interna da ovomucóide.

A antigenicidade da ovomucóide encontra-se no gráfico 45.



**Gráfico 45:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a ovomucóide, usando a escala de Hopp & Wood (azul). Zonas antigénicas possíveis (setas azuis).

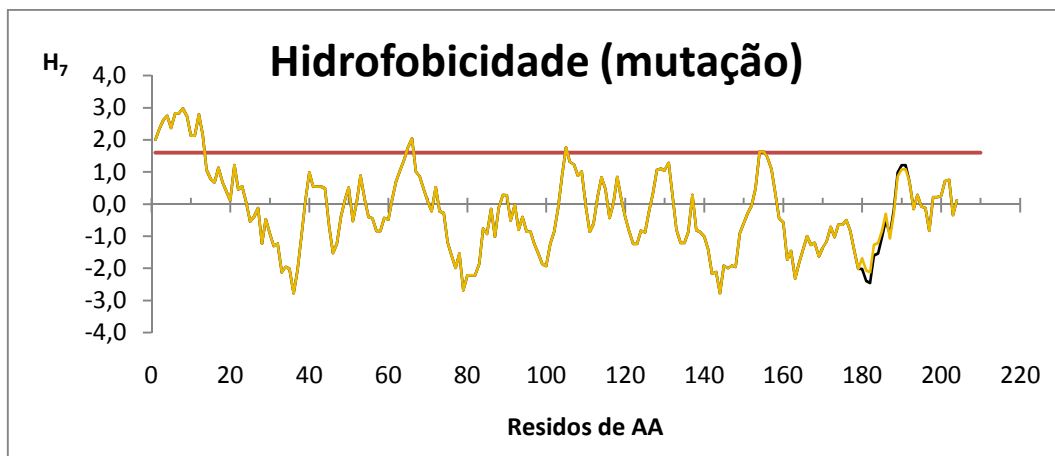
Devido ao comportamento evidenciado pela ovomucóide relativamente à sua alergenicidade dominante em comparação às outras proteínas da clara do ovo (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009), seria de esperar um número importante de antígenos. No entanto, a ovomucóide apresenta três picos antigénicos importantes (Gráfico 45). Esta ocorrência é certamente devida ao facto da ovomucóide ser de pequena dimensão (210 resíduos de aminoácido).

É preciso não esquecer que a ovomucóide é composta por três domínios (Matsuda et al., 1981; Jedrychowski & Wichers, 2009) e que provavelmente, cada domínio possui o seu próprio antígeno.

Tendo em conta que a ovomucóide é uma das proteínas da clara do ovo mais alergénicas (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009), Mine *et al.* (2003), tentaram diminuir a antigenicidade e a alergenicidade da ovomucóide, através de mutações genéticas, que deram origem a alterações na sequência dos resíduos de aminoácidos. Assim substituíram no domínio III da ovomucóide, a glicina da posição 32 por metionina e a fenilalanina da posição 37 também por metionina.

Com o objectivo de verificar se as conclusões de Mine *et al.* (2003) podem ser previstas nos cálculos efectuados para a Ovomucóide, foram efectuadas as mutações referidas anteriormente no domínio III e efectuados de novo os cálculos para  $H_7$  (Gráfico 46),  $H_a$  (Gráfico 48) e  $H_{Ag}$  (Gráfico 50).

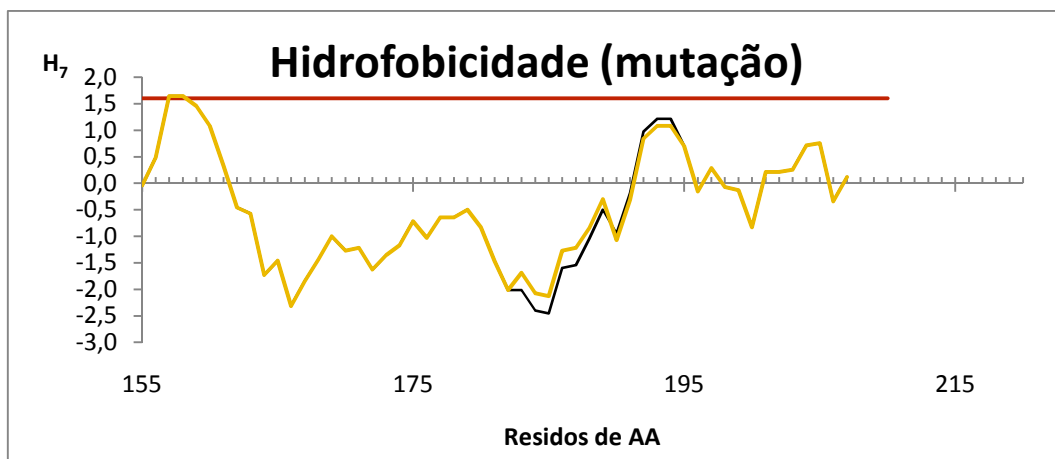
A hidrofobicidade da ovomucóide mutada pode ser observada no gráfico 46.



**Gráfico 46:** Representação gráfica dos cálculos da hidrofobicidade  $H_7$  para a ovomucóide, usando a escala de Kyte & Doolittle para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Esta mutação consiste em substituir a glicina 32 (G32) e a fenilalanina 37 (F37) pela metionina. Linha de referência de 1,6 (vermelho).

Observa-se umas ligeiras variações da hidrofobicidade entre os resíduos 183-194 na ovomucóide mutada (Gráfico 46). Estas variações situam-se no terceiro domínio da ovomucóide (Matsuda et al., 1981; Jedrychowski & Wichers, 2009).

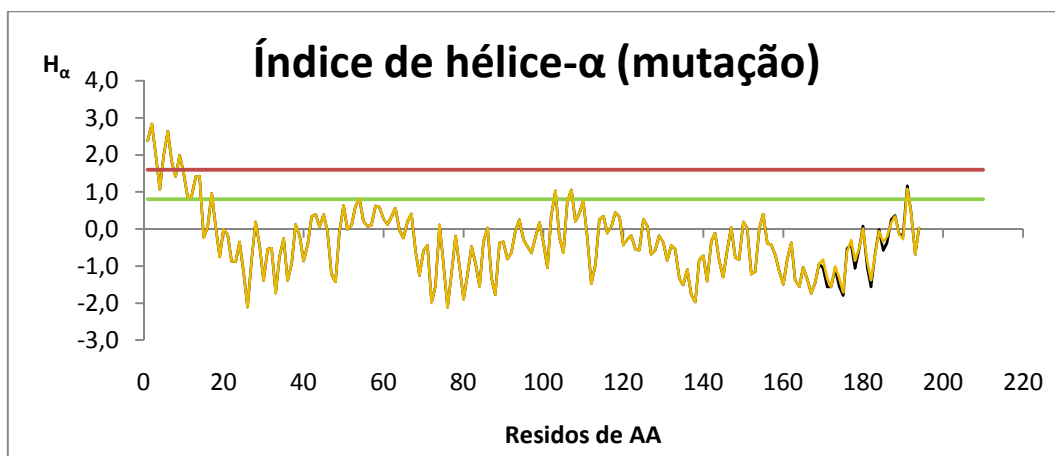
A ampliação do terceiro domínio da ovomucóide nativa e mutada, permite visualizar detalhadamente as alterações de hidrofobicidade  $H_7$ .



**Gráfico 47:** Representação gráfica pormenorizada da hidrofobicidade  $H_7$  para a ovomucóide, usando a escala de Kyte & Doolittle para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Janela de apenas 56 resíduos de aminoácido (Terceiro domínio completo). Linha de referência de 1,6 (vermelho).

Observa-se um ligeiro aumento da hidrofobicidade da ovomucóide mutada entre os resíduos de aminoácido 183-189 e uma pequena diminuição entre os resíduos 190-194 (Gráfico 47).

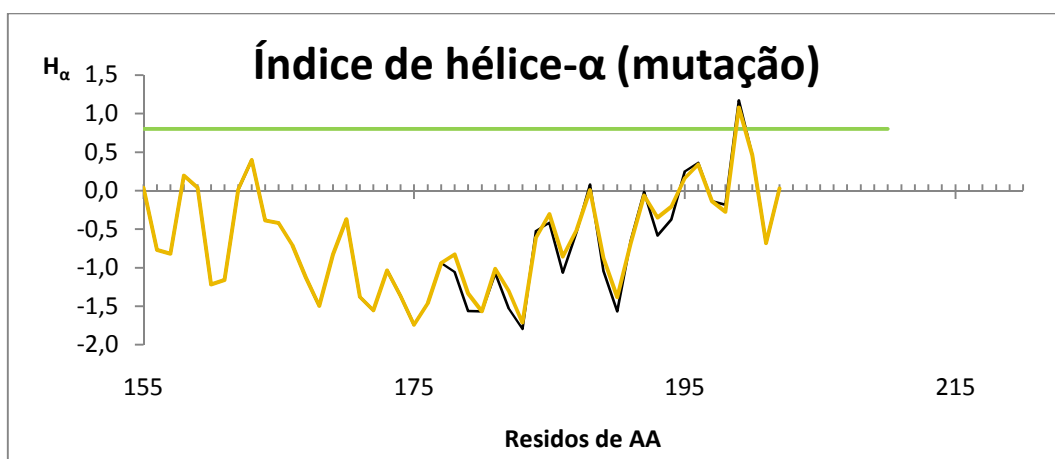
O índice de hélice- $\alpha$  da ovomucóide mutada e nativa pode ser observado no gráfico 48.



**Gráfico 48:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a ovomucóide, usando a escala de Kyte & Doolittle para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Esta mutação consiste em substituir a glicina 32 (G32) e a fenilalanina 37 (F37) pela metionina. (210 resíduos de aminoácido). Linha de referência de 0,8 (verde claro).

Observa-se umas ligeiras variações no índice de hélice- $\alpha$  entre os resíduos 178-199 na ovomucóide mutada (Gráfico 48). Estas variações situam-se no terceiro domínio da ovomucóide (Matsuda et al., 1981; Jedrychowski & Wichers, 2009).

A ampliação do terceiro domínio da ovomucóide nativa e mutada, permite visualizar detalhadamente as alterações de hidrofobicidade  $H_7$ .



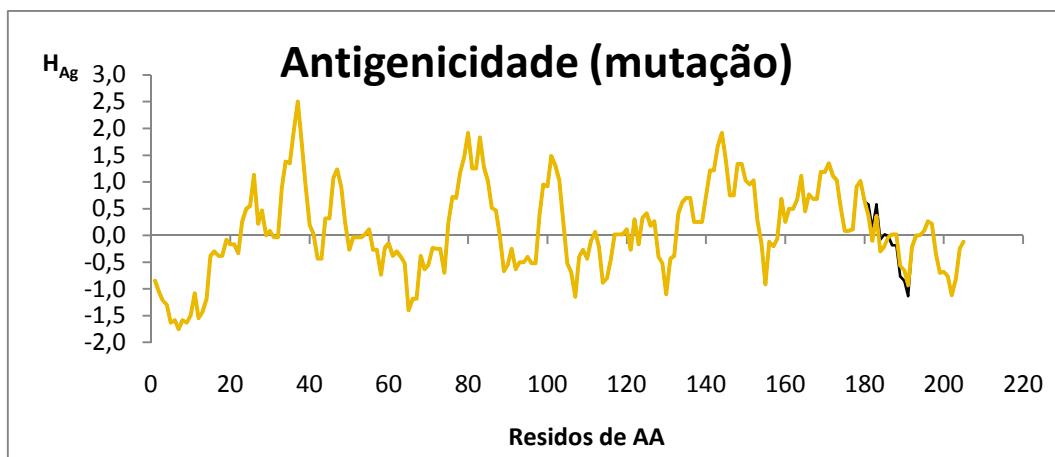
**Gráfico 49:** Representação gráfica pormenorizada do índice de hélice- $\alpha$  para a ovomucóide, usando a escala de Kyte & Doolittle para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Janela de apenas 56 resíduos de aminoácido (Terceiro domínio completo). Linha de referência de 0,8 (verde claro).

Observa-se um ligeiro aumento no índice de hélices- $\alpha$  da ovomucóide mutada entre os resíduos de aminoácido 178-194 e uma pequena diminuição entre os resíduos 195-199 (Gráfico 49).

Apesar do estudo feito pela equipa de Yoshinori Mine ter dado resultados promissores, quando se faz a mesma substituição na sequência de resíduos de aminoácidos da ovomucóide

usando o método da escala de hidrofobicidade de Kyte & Doolittle, os resultados obtidos não são muito significativos. Como se pode observar, tanto na hidrofobicidade  $H_7$  (Gráficos 46 e 47), como no índice de hélices- $\alpha$  (Gráficos 48 e 49), não se observa grandes alterações relativamente à proteína original.

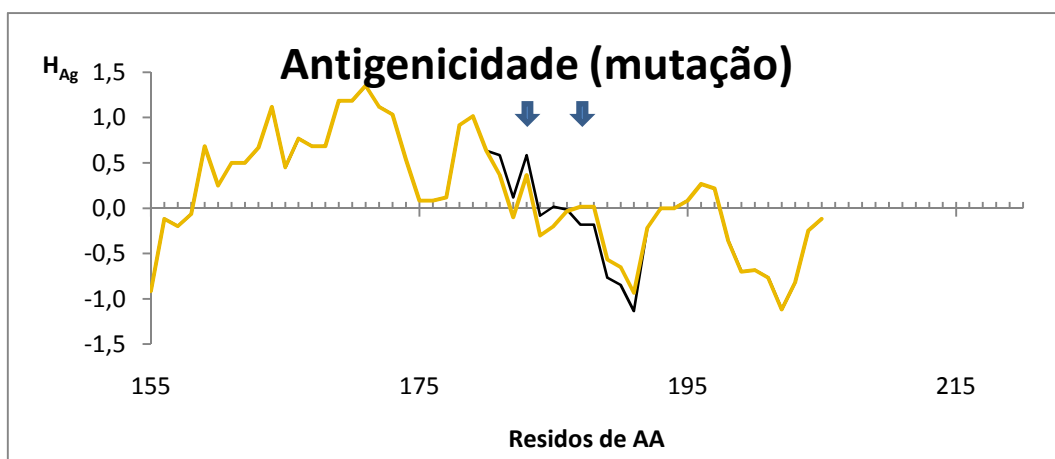
A antigenicidade da ovomucóide nativa e mutada pode ser observada no gráfico 50.



**Gráfico 50:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a ovomucóide, usando a escala de Hopp & Wood para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Esta mutação consiste em substituir a glicina 32 (G32) e a fenilalanina 37 (F37) pela metionina.

Mesmo que não seja de uma grande evidência, é de notar uma pequena diminuição da antigenicidade entre os resíduos 180 e 200 (Gráfico 50).

Para facilitar a observação das diferenças provocadas pela mutação, elaborou-se uma ampliação do gráfico da antigenicidade representando exclusivamente o terceiro domínio da ovomucóide (Gráfico 51).

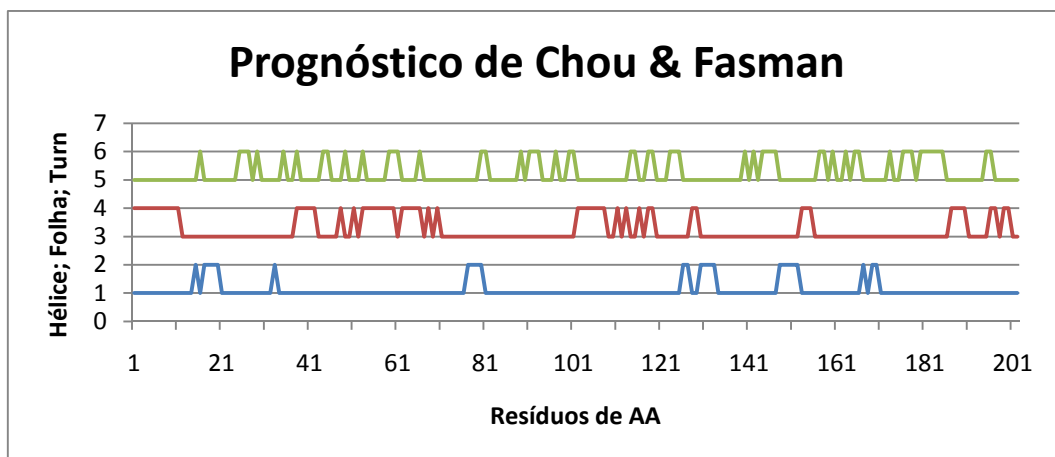


**Gráfico 51:** Representação gráfica pormenorizada da antigenicidade  $H_{Ag}$  para a ovomucóide, usando a escala de Hopp & Wood para a proteína inata (preto) e usando a mesma escala para a proteína mutada (laranja) nos resíduos 32 e 37 do terceiro domínio. Janela de apenas 56 resíduos de aminoácido (Terceiro domínio completo). Zonas de divergência (setas azuis).

Na região mutada verifica-se a diminuição da antigenicidade do primeiro pico. No segundo pico, observa-se o comportamento inverso (Gráfico 51). À luz dos resultados obtidos,

verifica-se que a substituição da glicina pela metionina, diminui a antigenicidade. No entanto, a substituição da fenilalanina também pela metionina, aumenta a antigenicidade e consecutivamente a alergenicidade da ovomucóide. Contudo, não é possível quantificar esta diminuição nem quantificar a intensidade de cada reacção alérgica.

O gráfico 52 demonstra a aplicação do método de Chou & Fasman à ovomucóide.



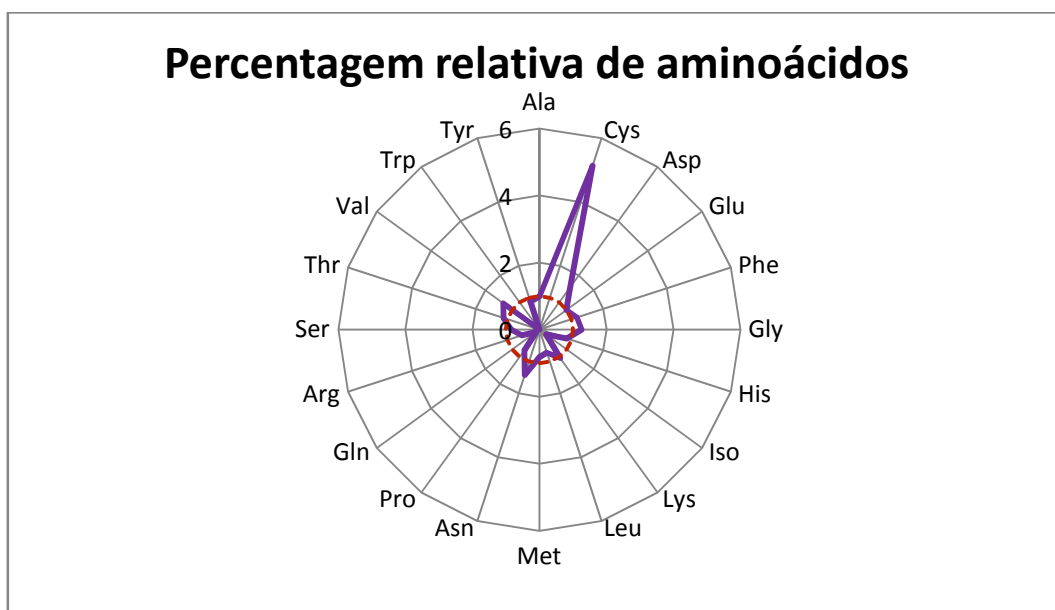
**Gráfico 52:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na ovomucóide. Cada pico corresponde à presença de um possível elemento estrutural.

Observa-se que existe a probabilidade de encontrar cinco hélices- $\alpha$  (trapézio isósceles) ao longo da cadeia polipeptídica da ovomucóide (Gráfico 52). Observa-se igualmente que existe alguma probabilidade de que a ovomucóide possua inúmeras voltas e inúmeras folhas- $\beta$ . Isto confirma o que é descrito na literatura, visto que a ovomucóide possui: 26% de hélices- $\alpha$ , 46% de folhas- $\beta$  e 10% de voltas (Belitz, 2009). Este método demonstrou ser suficientemente sensível para detectar os três tipos de estrutura, no entanto, as proporções parecem estar invertidas relativamente à bibliografia. Apesar de já se ter referido que o método de Chou & Fasman não mostra ser muito adequado na detecção de folhas- $\beta$ , para esta proteína ele foi bastante sensível.



tenha pelo menos uma hélice- $\alpha$  na sua estrutura (Berg et al., 2002). No entanto, a grande quantidade de valina; aspartato; serina; asparagina; treonina e serina, pode perturbar a formação dessas mesmas hélices- $\alpha$  e destabiliza-las caso elas existam (Berg et al., 2002). O valor elevado em valina e médio em treonina revelam também a possível presença de folhas- $\beta$  (Berg et al., 2002). A presença de voltas é igualmente possível devido à existência de valores altos e médios de glicina e asparagina, respectivamente (Berg et al., 2002). O facto de a ovomucóide não possuir triptofano na sua composição, não influencia de modo algum a formação ou a perturbação das três estruturas: hélices- $\alpha$ ; folhas- $\beta$  e voltas.

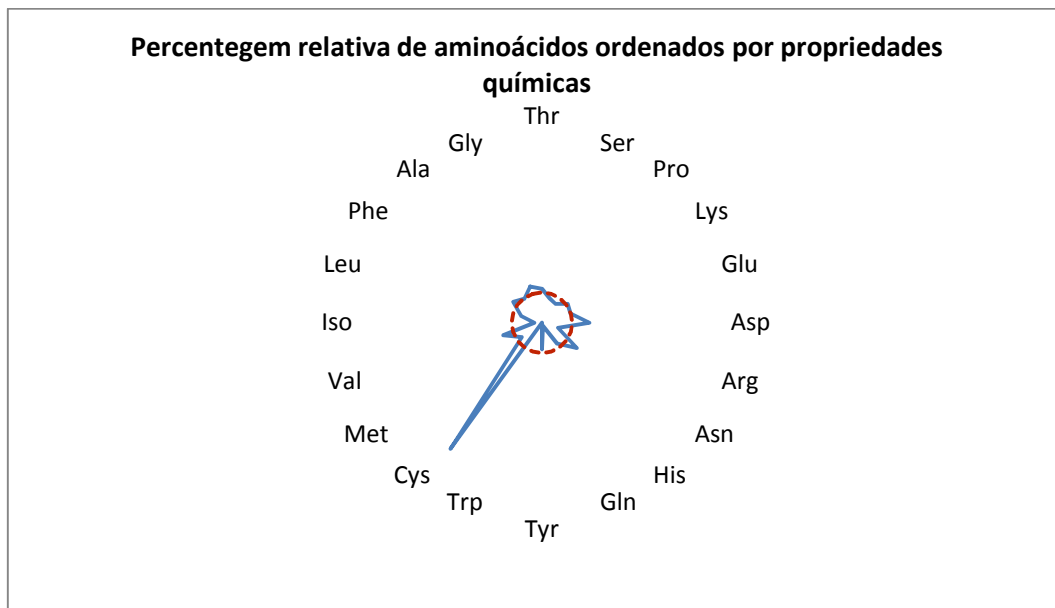
Através de um gráfico em radar é possível comparar a percentagem de aminoácidos da ovomucóide relativamente à GenPep#97 (Gráfico 55).



**Gráfico 55:** Representação gráfica da percentagem de cada aminoácido na ovomucóide, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 19 aminoácidos presentes na proteína.

Observa-se que a ovomucóide possui oito aminoácidos com uma percentagem relativa acima da média, sendo esta calculada a partir das percentagens relativas dos 19 aminoácidos presentes na proteína e representada a tracejado vermelho (Gráfico 55). Desses oito, o mais evidente é a cisteína, seguida do aspartato; asparagina; valina; glicina; fenilalanina; treonina e lisina. Os restantes aminoácidos têm um valor abaixo da média.

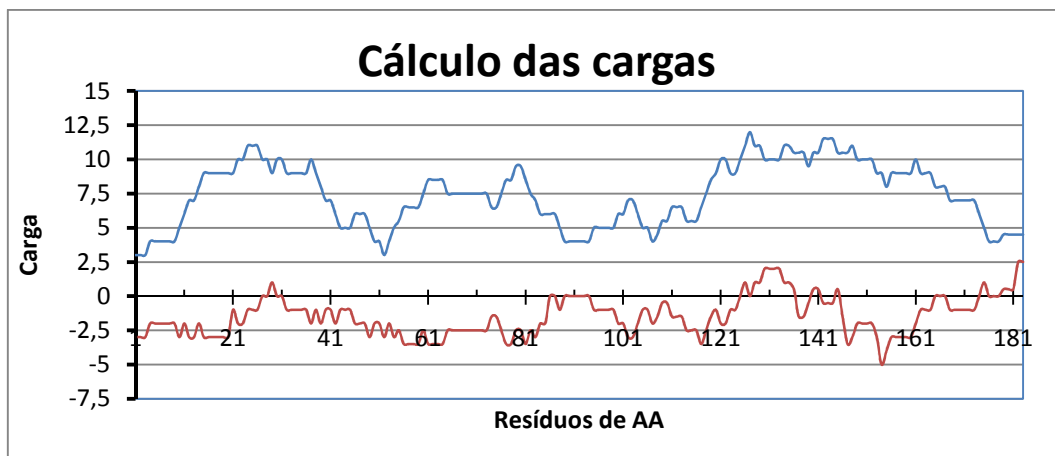
Uma das formas de verificar o comportamento da ovomucóide, é dispor a percentagem de aminoácidos de modo a que os aminoácidos hidrofóbicos fiquem à esquerda e os carregados fique à direita.



**Gráfico 56:** Representação gráfica da percentagem relativa dos aminoácidos da ovomucóide, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 19 aminoácidos presentes na proteína.

Apesar de nada se saber à luz da teoria quanto à hidrossolubilidade da ovomucóide, ela mostra comportar-se de maneira hidrofóbica (Gráfico 56). A ovomucóide possui acima da média três aminoácidos carregados, aspartato; asparagina e lisina, que interferem na característica hidrossolúvel desta proteína. No entanto, para além da cisteína, a ovomucóide também possui acima da média, três aminoácidos hidrofóbicos, tais como: a valina; glicina e a fenilalanina. A treonina também se encontra acima da média mas não interfere no comportamento global da ovomucóide.

As cargas reais e absolutas da ovomucóide são apresentadas no gráfico 57.



**Gráfico 57:** Representação gráfica das cargas da ovomucóide. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

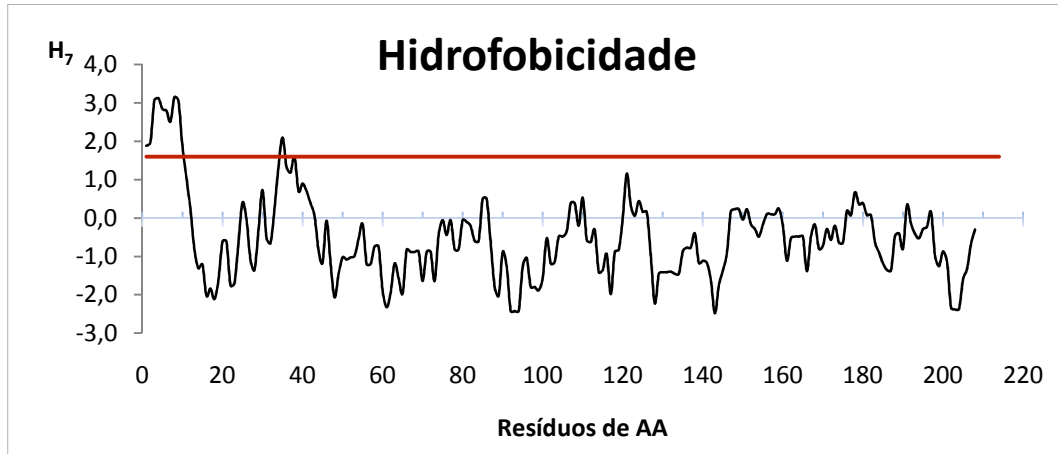
Observa-se que a ovomucóide possui poucas cargas reais positivas (Gráfico 57). As únicas que se destacam encontram-se nos intervalos 128-136 e 179-183. Existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores, como se fossem cargas “residuais”.

Observa-se igualmente que as cargas absolutas da ovomucóide, mostram que ela tem um comportamento polar positivo ao longo de toda a cadeia polipeptídica. Destacam-se três zonas distintas nas quais as cargas absolutas são maiores. Elas situam-se nos intervalos: 9-49; 53-88 e 116-175. Isto significa que se a ovomucóide for colocada numa solução com uma corrente eléctrica, ela deslocar-se-á para o cátodo (pólo negativo).

### 4.2.3. Proteínas do leite

#### 4.2.3.1. $\alpha$ -caseína s1

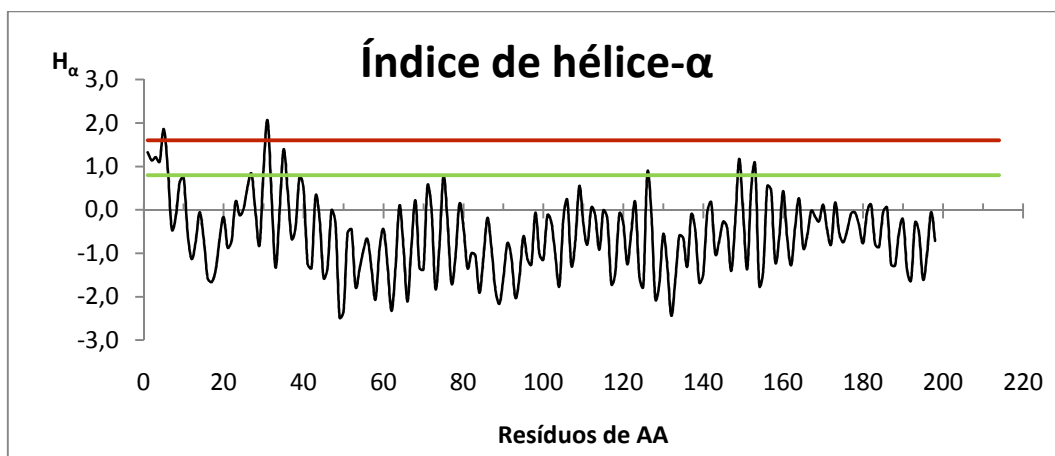
No gráfico 58 encontra-se o resultado do estudo de hidrofobicidade ( $H_7$ ) para a  $\alpha$ -caseína s1.



**Gráfico 58:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a  $\alpha$ -caseína s1 (214 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).

A  $\alpha$ -caseína s1 é uma proteína do leite com o centro hidrofóbico e a periferia hidrofílica como todas as caseínas do leite (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009), no entanto ela não é solúvel em água. Poderia pensar-se que o pico que se observa no início da sua cadeia de aminoácidos (Gráfico 58), seria uma zona hidrofóbica no entanto, esse pico não é muito significativo quanto à sua hidrofobicidade visto que o número de aminoácidos que o compõe é inferior a 19. Apesar de ultrapassar largamente o nível teórico de 1,6 este pico não parece influenciar a solubilidade da  $\alpha$ -caseína s1. Possivelmente é uma região que se encontra no interior da  $\alpha$ -caseína s1 visto que, como já foi referido acima, o seu interior é hidrofóbico.

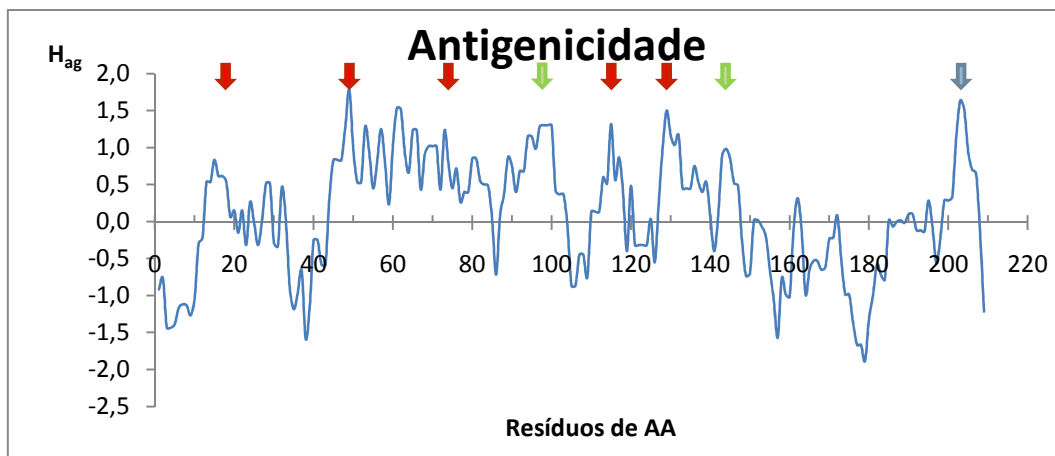
No gráfico 59, observa-se o resultado dos cálculos para o índice de hélice- $\alpha$  da  $\alpha$ -caseína s1.



**Gráfico 59:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a  $\alpha$ -caseína s1 (214 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

Observa-se que a nível do índice de hélices- $\alpha$ , o gráfico 59 não é muito elucidativo visto que não indica qualquer pico susceptível de manifestar a presença de hélices- $\alpha$  anfifílicas. Para além disso, não foram encontrados dados literários que corroborem ou invalidem as informações apresentadas nesse mesmo gráfico.

A antigenicidade da  $\alpha$ -caseína s1 pode ser observada no gráfico 60.

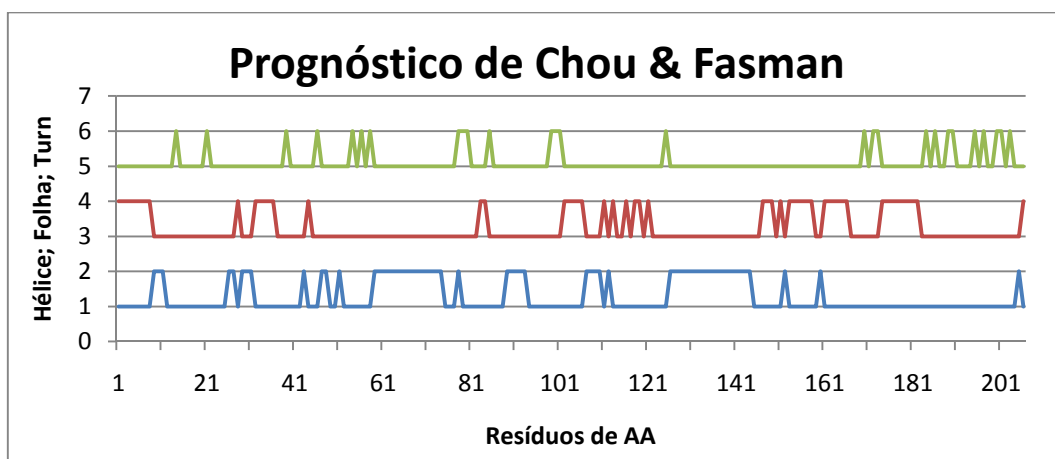


**Gráfico 60:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a  $\alpha$ -caseína s1, usando a escala de Hopp & Wood (azul). Epítomos imunodominantes (setas verdes), ligações à IgE (setas vermelhas) e ligação antigénica provável (seta azul).

Prevê-se que a  $\alpha$ -caseína s1 possua oito zonas antigénicas correspondendo a zonas hidrofílicas (Gráfico 60). Observando o gráfico vê-se que a maior parte dos resíduos são antigénicos porque estão na parte positiva. A presença de um grande número de picos antigénicos era previsível visto que a  $\alpha$ -caseína s1 é uma das proteínas descrita como responsável por alergias mediadas pela IgE (Jedrychowski & Wichers, 2009). Foram identificados epítomos alergénicos nas regiões 19-30, 93-98 e 141-150 da  $\alpha$ -caseína s1, como

sendo um epítipo imunodominante (setas verdes). Foram reconhecidas algumas regiões de ligação às IgE nos aminoácidos 17-36, 39-48, 69-78, 93-102, 109-120, 123-132, 139-154, 159-174 e 173-194 (setas vermelhas) (Spuergin *et al.* 1997). É de notar igualmente que existe uma probabilidade de encontrar um epítipo no final da sequência polipeptídica (Gráfico 60, seta azul).

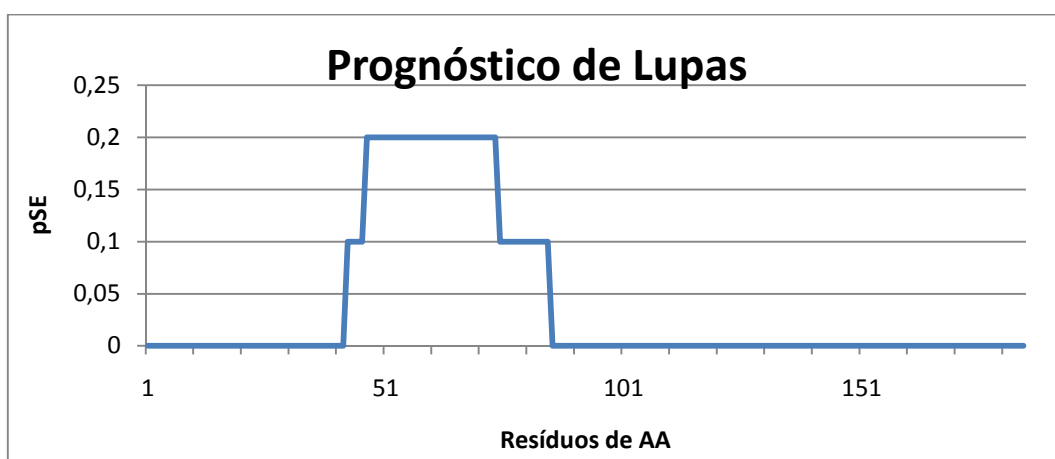
O gráfico 61 demonstra a aplicação do método de Chou & Fasman para a  $\alpha$ -caseína s1.



**Gráfico 61:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha-B (vermelho) e volta (verde), na  $\alpha$ -caseína s1. Cada pico corresponde à presença de um possível elemento.

Observa-se que existe uma grande probabilidade de encontrar hélices- $\alpha$ ; folhas-B e voltas ao longo da cadeia polipeptídica da  $\alpha$ -caseína s1 (Gráfico 61). Este facto não pode ser confirmado pela literatura visto que nada foi encontrado sobre o assunto, para esta proteína. Todavia, o método de Chou & Fasman mostrou ser pouco fiável relativamente à BSA, na detecção de folhas-B. Nada indica que o mesmo não aconteça com a  $\alpha$ -caseína s1.

O algoritmo de Lupas *et al.* pode ser observado através do seguinte gráfico.



**Gráfico 62:** Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na  $\alpha$ -caseína s1. A cada pico corresponder a possível presença dessa estrutura.

Observa-se que existe uma probabilidade em encontrar uma super-hélice- $\alpha$  na  $\alpha$ -caseína s1 (Gráfico 62). Essa região está delimitada entre os resíduos 43-85 mas tem uma probabilidade ainda maior entre os resíduos 47-74.

O gráfico 63 proporciona uma visão global da composição em aminoácidos da  $\alpha$ -caseína s1.

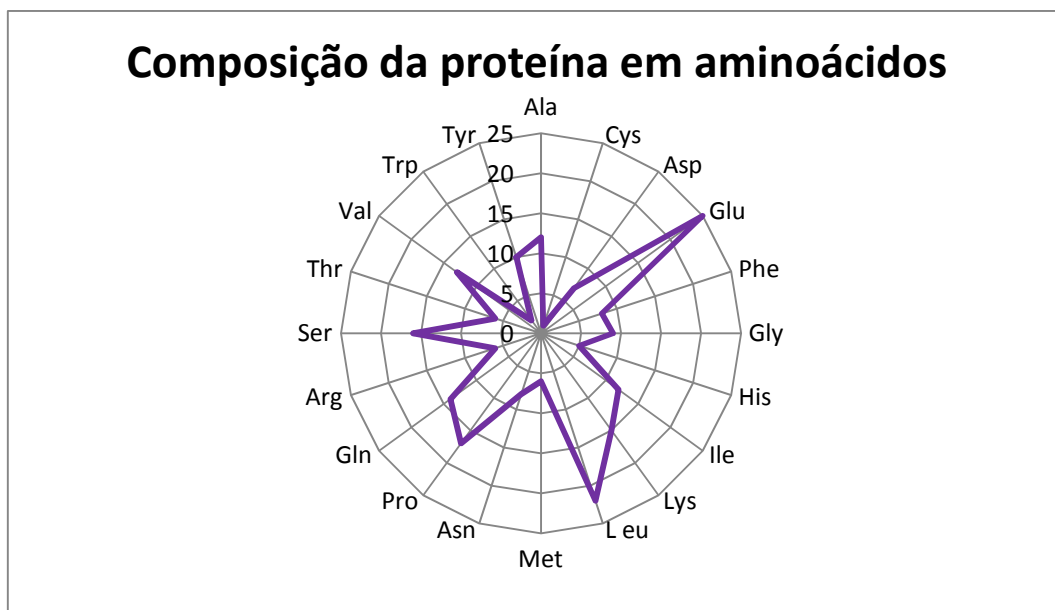
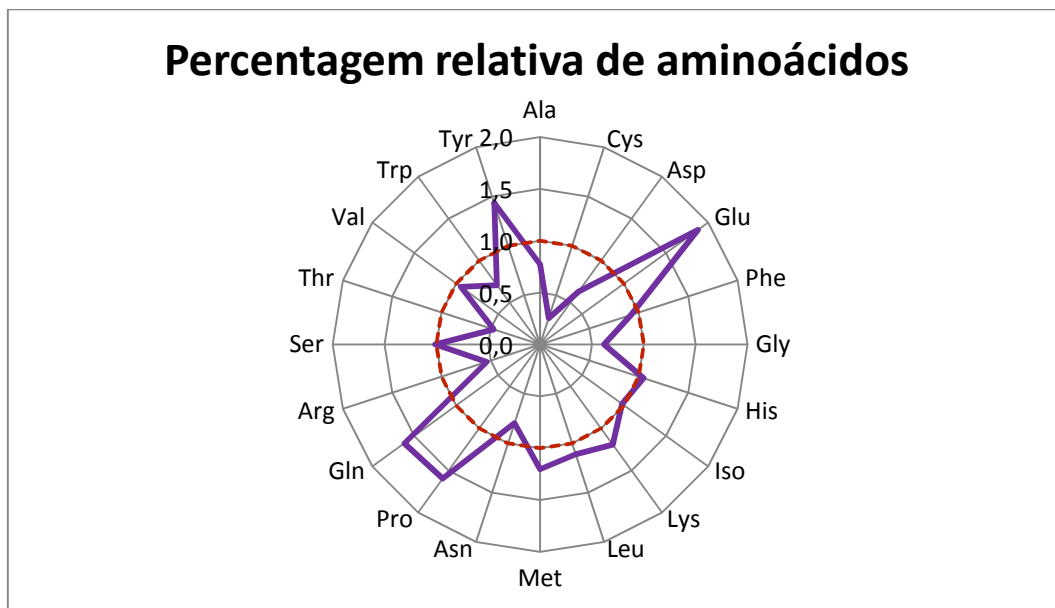


Gráfico 63: Representação gráfica da quantidade real de cada aminoácido na  $\alpha$ -caseína s1 (214 resíduos de aminoácido).

A  $\alpha$ -caseína s1 é constituída pelos vinte principais aminoácidos (Gráfico 63). Sendo assim, a  $\alpha$ -caseína s1 é constituída maioritariamente por (ordem decrescente): glutamato, leucina, prolina, serina, lisina, glutamina, valina, alanina, isoleucina e tirosina. Tem uma quantidade média de: glicina, fenilalanina, asparagina, aspartato, metionina, arginina, treonina e histidina. A  $\alpha$ -caseína s1 possui igualmente alguns aminoácidos residuais tais como: triptofano e cisteína.

Visto que a  $\alpha$ -caseína s1 é relativamente rica em, glutamato e leucina, existe uma forte possibilidade de que ela apresente pelo menos uma hélice- $\alpha$  na sua estrutura (Berg et al., 2002).

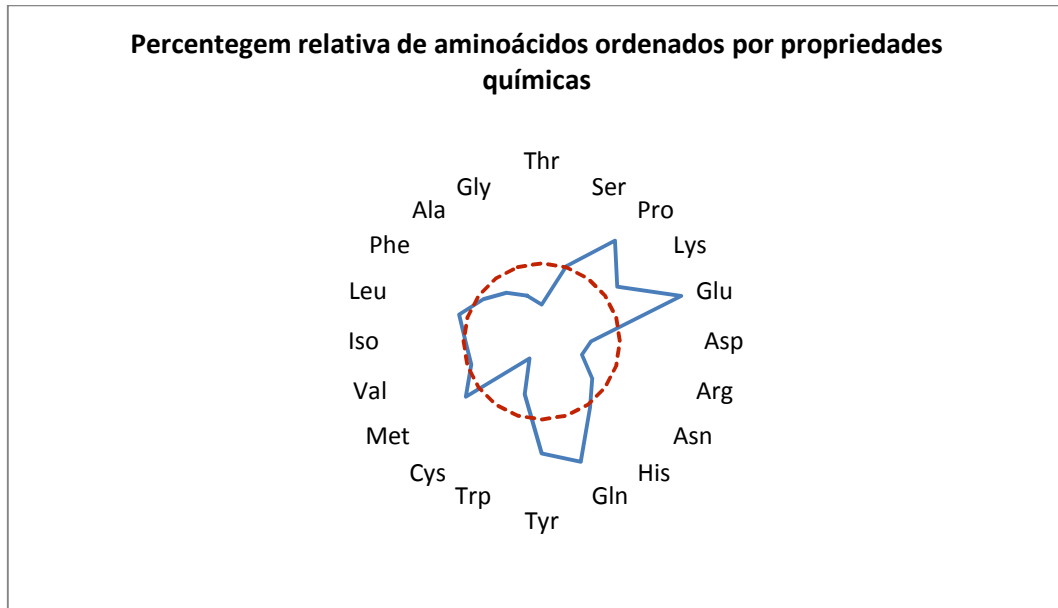
O gráfico 64 mostra a percentagem relativa de aminoácidos da  $\alpha$ -caseína s1 em relação à GenPep#97.



**Gráfico 64:** Representação gráfica da percentagem de cada aminoácido na  $\alpha$ -caseína s1, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

Observa-se que a  $\alpha$ -caseína s1 possui nove aminoácidos com uma percentagem relativa acima da média, sendo esta calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína e representada a tracejado vermelho (Gráfico 64). Desses nove, o mais evidente é o glutamato seguido de glutamina, prolina, tirosina, metionina, lisina, leucina, histidina e no limiar da média encontra-se a serina. Os restantes aminoácidos encontram-se num número abaixo da média.

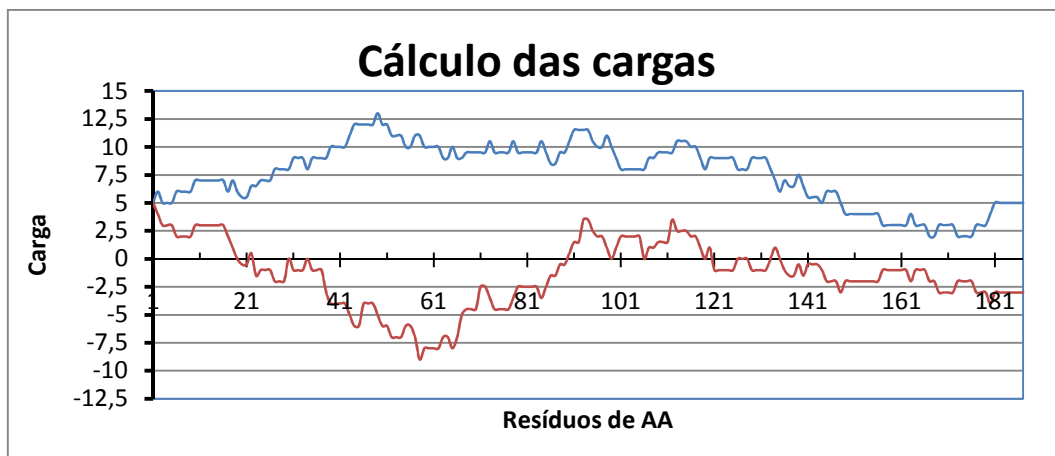
A distribuição da percentagem relativa de aminoácidos segundo as suas propriedades químicas, revela algumas características da  $\alpha$ -caseína s1 (Gráfico 65).



**Gráfico 65:** Representação gráfica da percentagem relativa dos aminoácidos da  $\alpha$ -caseína s1, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão situados à direita e os mais hidrofóbicos estão situados à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

De acordo com o gráfico 65, a  $\alpha$ -caseína s1 mostra ser uma proteína hidrofílica. Essa ideia é reforçada pela presença de uma percentagem relativa acima da média de aminoácidos hidrofílicos tais como: glutamato, glutamina, prolina, lisina, histidina e serina, mesmo que os seus efeitos estejam limitados pelos aminoácidos hidrofóbicos. Dos aminoácidos hidrofóbicos que compõem a  $\alpha$ -caseína s1, metionina e leucina têm valores acima da média e contribuem ao carácter hidrofóbico desta proteína visto que o centro das caseínas é hidrofóbico (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009). A tirosina também tem valores acima da média mas não afecta o comportamento global da  $\alpha$ -caseína s1.

As cargas reais e absolutas da  $\alpha$ -caseína s1 encontram-se representadas no gráfico 66.



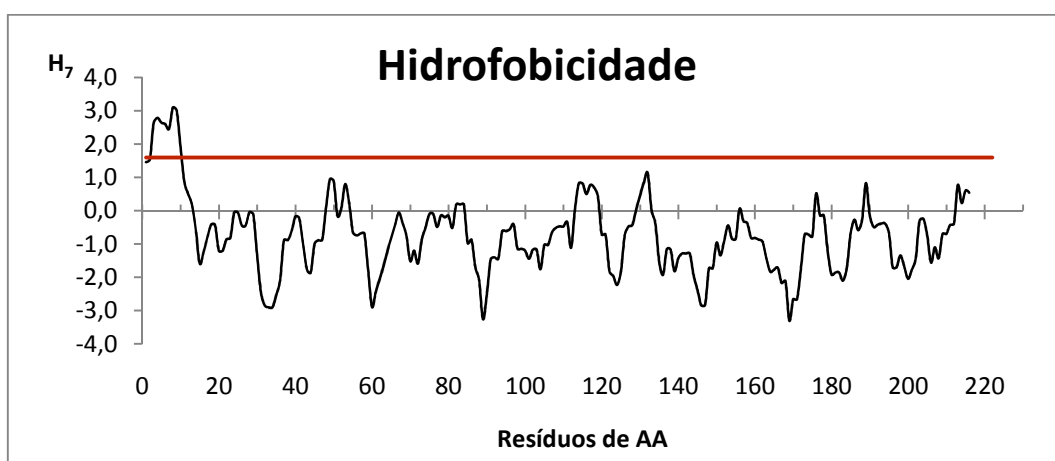
**Gráfico 66:** Representação gráfica das cargas da  $\alpha$ -caseína s1. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

Observa-se que a  $\alpha$ -caseína s1 possui poucas cargas reais positivas (Gráfico 66). As únicas que se destacam encontram-se nos intervalos: 1-18 e 90-118. Existem outros intervalos onde as cargas reais também são positivas mas são apenas cargas residuais.

Observa-se igualmente que as cargas absolutas da  $\alpha$ -caseína s1, mostram que ela tem um comportamento polar positivo ao longo de toda a cadeia polipeptídica, destacando-se apenas o final da cadeia em que as cargas absolutas não são tão elevadas como no resto da cadeia.

#### 4.2.3.2. $\alpha$ -caseína s2

A hidrofobicidade da  $\alpha$ -caseína s2 pode representar-se graficamente para uma melhor visualização (Gráfico 67).

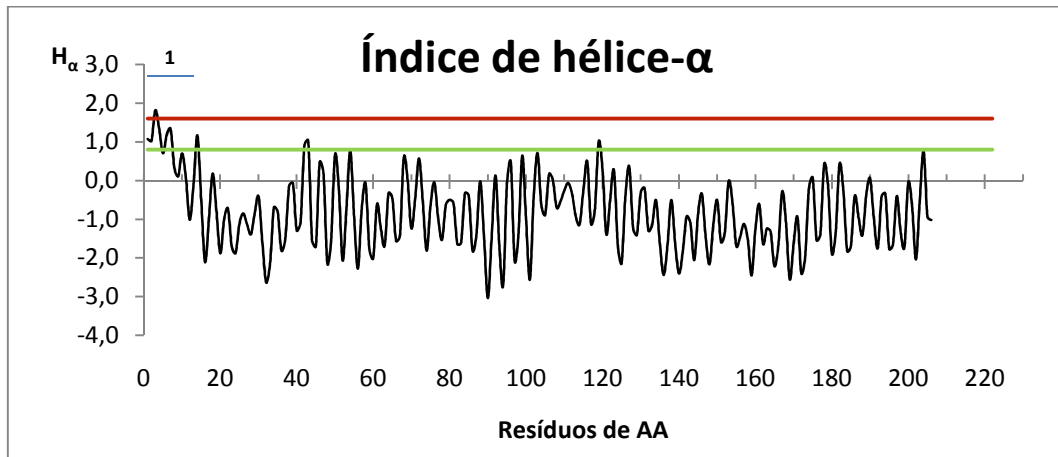


**Gráfico 67:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a  $\alpha$ -caseína s2 (222 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).

Tal como a  $\alpha$ -caseína s1, a  $\alpha$ -caseína s2 também não é hidrossolúvel, apesar de a sua periferia ser hidrofílica (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills,

2009). Por essa razão a semelhança dos seus comportamentos é notável (Gráfico 67). O início da sequência de aminoácidos da  $\alpha$ -caseína s2 mostra um pico hidrofóbico, tal como acontece com a  $\alpha$ -caseína s1 (Gráfico 58). Poderia pensar-se que o pico que se observa no início da sua cadeia de aminoácidos, seria uma zona hidrofóbica no entanto, esse pico não é muito significativo quanto à sua hidrofobicidade visto que o número de aminoácidos que o compõe é inferior a 19. Apesar de ultrapassar largamente o nível teórico de 1,6. Exceptuando esse pico inicial, a  $\alpha$ -caseína s2 mostra ter um comportamento hidrofílico ao longo da sua cadeia.

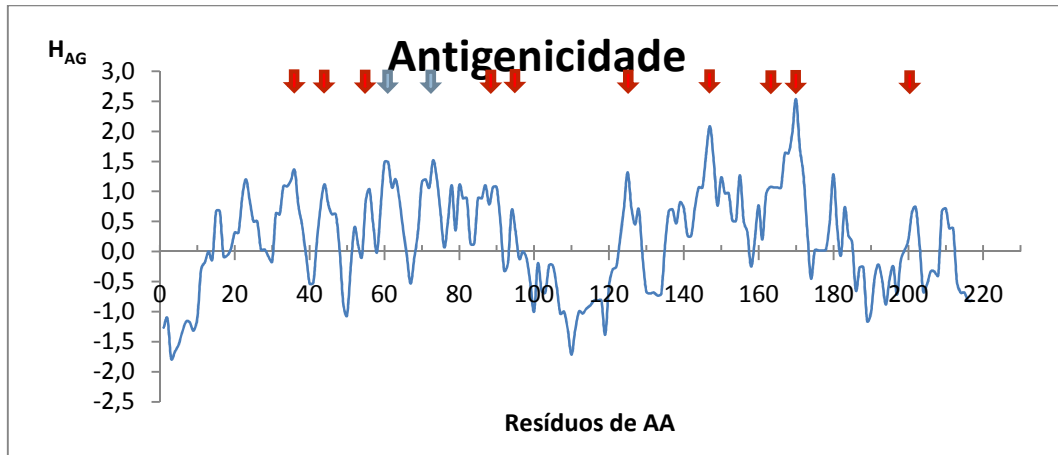
As hélices- $\alpha$  podem ser detectadas através do índice de hélices- $\alpha$  (Gráfico 68).



**Gráfico 68:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a  $\alpha$ -caseína s2 (222 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

Observa-se que no início do gráfico 68, existe uma zona anfifílica (pico 1). Este pico pode indicar a presença de uma hélice- $\alpha$  anfifílica. É um tipo de hélice- $\alpha$  particular, conhecida por ter resíduos alternadamente hidrofóbicos e hidrofílicos (Murray *et al.*, 2002). Ao longo do resto da sequência de aminoácidos, a  $\alpha$ -caseína s2, não mostra ter zonas anfifílicas nem hidrofóbicas, mas sim um comportamento hidrofílico, o que reforça a ideia de que a sua periferia é hidrofílica.

A antigenicidade da  $\alpha$ -caseína s2 pode ser representada graficamente do seguinte modo.



**Gráfico 69:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a  $\alpha$ -caseína s2, usando a escala de Hopp & Wood (azul). Ligações fracas à IgE (setas azuis) e ligações fortes à IgE (setas vermelhas).

A  $\alpha$ -caseína s2 é uma proteína responsável por algumas alergias, (Jedrychowski & Wichers, 2009). Sabendo isso, era de esperar a presença de um número elevado de zonas antigénicas como é o caso (Gráfico 69). No entanto, constata-se que através deste método, a  $\alpha$ -caseína s2 manifesta doze regiões antigénicas, quando se deveria observar apenas dez (Jedrychowski & Wichers, 2009). Dos doze picos observados, dez deles (seta vermelha) destacam-se por parecerem ser zonas de ligação à IgE mais intensas. Usando 99 decapeptídeos sintéticos, concluiu-se que a  $\alpha$ -caseína s2 apresentam dez regiões de ligação à IgE e foram identificadas como zonas alergénicas. Essas regiões situam-se nos aminoácidos: 31-44, 43-56, 83-100, 93-108, 105-114, 117-128, 143-158, 157-172, 165-188 e 191-200 (Jedrychowski & Wichers, 2009).

É de notar igualmente, a presença de duas regiões com uma probabilidade de encontrar epítomos antigénicos (setas azuis).

A previsão das hélices- $\alpha$ ; folhas- $\beta$  e voltas podem ser efectuada através do método de Chou & Fasman.

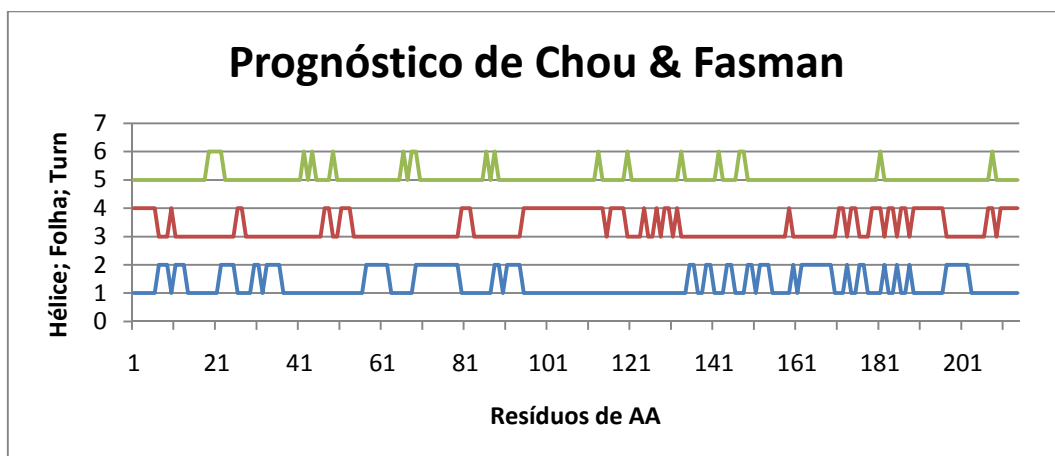


Gráfico 70: Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na  $\alpha$ -caseína s2. Cada pico corresponde à presença de um possível elemento.

Observa-se que existe alguma probabilidade de encontrar hélices- $\alpha$  ao longo da cadeia polipeptídica da  $\alpha$ -caseína s2 (Gráfico 70). Este facto é corroborado à luz de dados teóricos que estipulam a existência de 25% de hélices- $\alpha$  na estrutura secundária da  $\alpha$ -caseína s2 (Jedrychowski & Wichers, 2009). Observa-se igualmente alguma probabilidade de encontrar folhas- $\beta$  e voltas, apesar de não ser possível confirmar a presença destas estruturas através de dados teóricos.

O prognóstico de Lupas pode ser visualizado no gráfico 71.

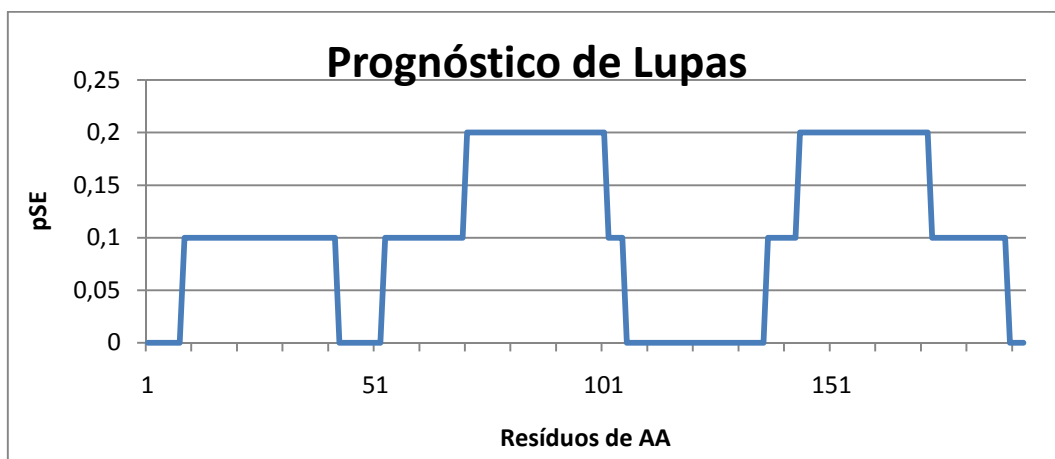
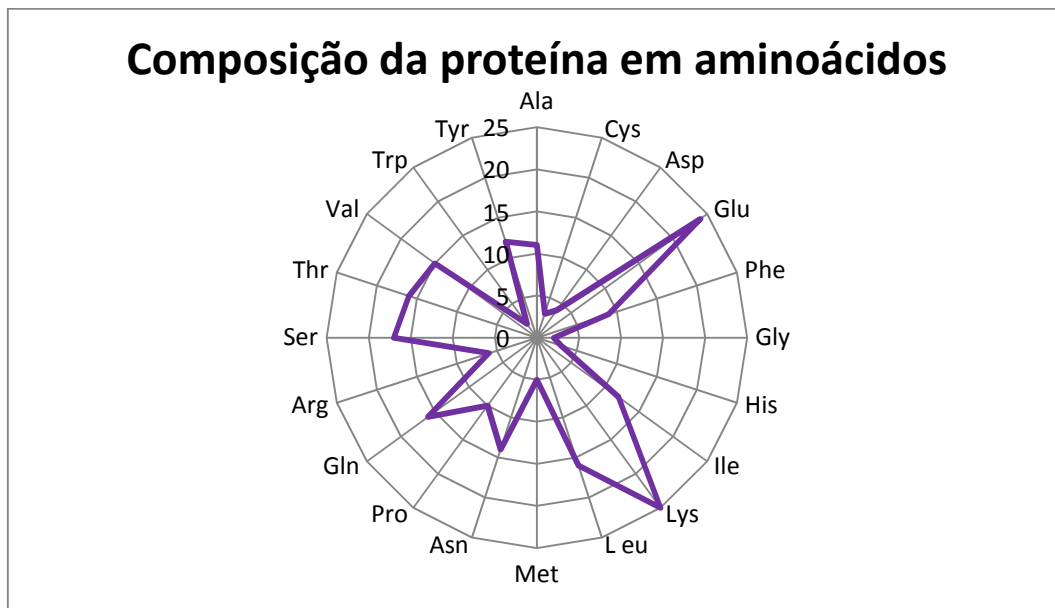


Gráfico 71: Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na  $\alpha$ -caseína s2. A cada pico corresponder a possível presença dessa estrutura.

Observa-se que existe a probabilidade de encontrar três super-hélices- $\alpha$  na  $\alpha$ -caseína s2, (Gráfico 71). Uma primeira entre os resíduos 9-42, uma segunda entre os resíduos 53-105 (com uma probabilidade acentuada entre 71-101) e uma terceira entre os resíduos 137-189 (com uma probabilidade acentuada entre 144-172). Nenhum dado teórico foi encontrado para confirmar a presença de super-hélices- $\alpha$  nesta proteína.

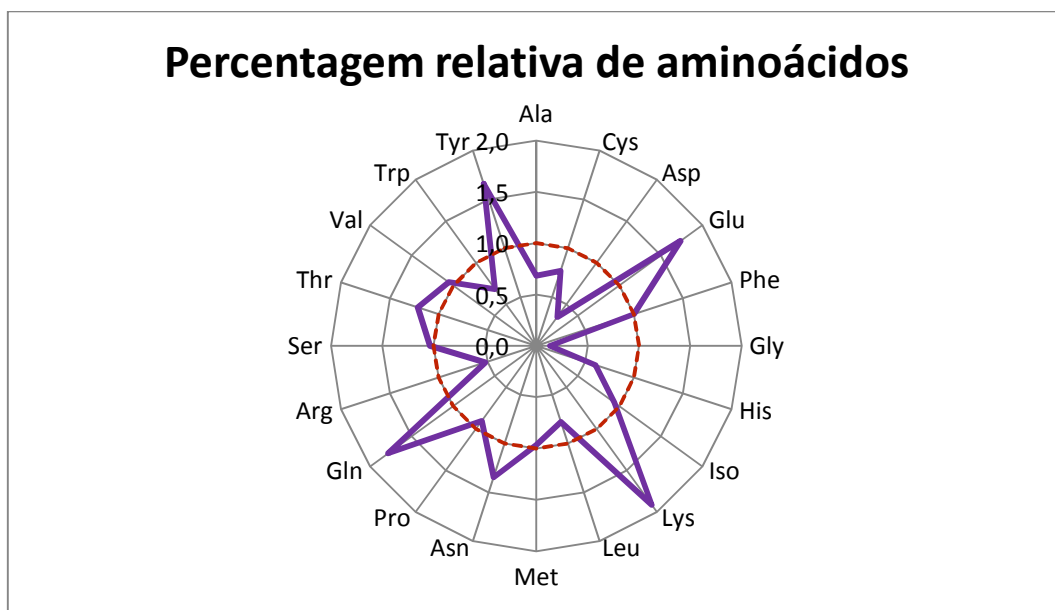
A composição em aminoácidos da  $\alpha$ -caseína s2 pode ser representada por um gráfico em radar (Gráfico 72).



**Gráfico 72:** Representação gráfica da quantidade real de cada aminoácido na  $\alpha$ -caseína s2 (222 resíduos de aminoácido).

A  $\alpha$ -caseína s2 é constituída pelos vinte principais aminoácidos, (Gráfico 72). Sendo assim, a  $\alpha$ -caseína s2 é constituída maioritariamente por (ordem decrescente): lisina, glutamato, serina, leucina, glutamina, treonina, valina, asparagina, isoleucina, tirosina, alanina e prolina. Tem uma quantidade mínima de: fenilalanina, arginina, metionina, aspartato, cisteína, histidina, glicina e triptofano. Visto que a  $\alpha$ -caseína s2 é rica em glutamato e tem uma quantidade média de leucina, reforça a ideia da presença de hélices- $\alpha$  na sua estrutura (Berg et al., 2002). A presença de isoleucina e valina revelam também a possível existência de folhas- $\beta$  (Berg et al., 2002). A presença de voltas é igualmente suspeitada devido à presença de asparagina e prolina (Berg et al., 2002).

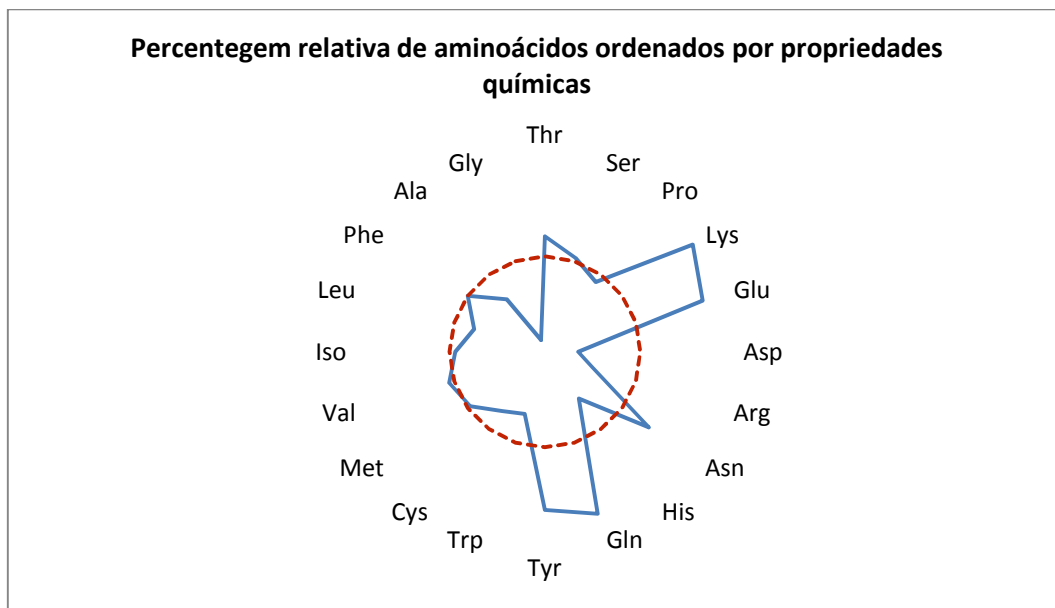
A comparação entre a  $\alpha$ -caseína s2 e a proteína média (GenPep#97) pode ser feita através das percentagens de aminoácidos que compõem cada uma delas.



**Gráfico 73:** Representação gráfica da percentagem de cada aminoácido na  $\alpha$ -caseína s2 (222 resíduos de aminoácido), relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

Observa-se que a  $\alpha$ -caseína s2 possui oito aminoácidos com uma percentagem acima da média relativamente à GenPep#97, média esta que foi calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína e está representada a tracejado vermelho (Gráfico 73). Desses oito aminoácidos, o mais evidente é a lisina, seguida da glutamina, glutamato, tirosina, asparagina, treonina, valina e serina. Os restantes aminoácidos têm um valor abaixo da média.

A distribuição dos aminoácidos segundo as suas propriedades químicas, revela as características hidrofóbicas ou hidrofílicas da  $\alpha$ -caseína s2 (Gráfico 74).

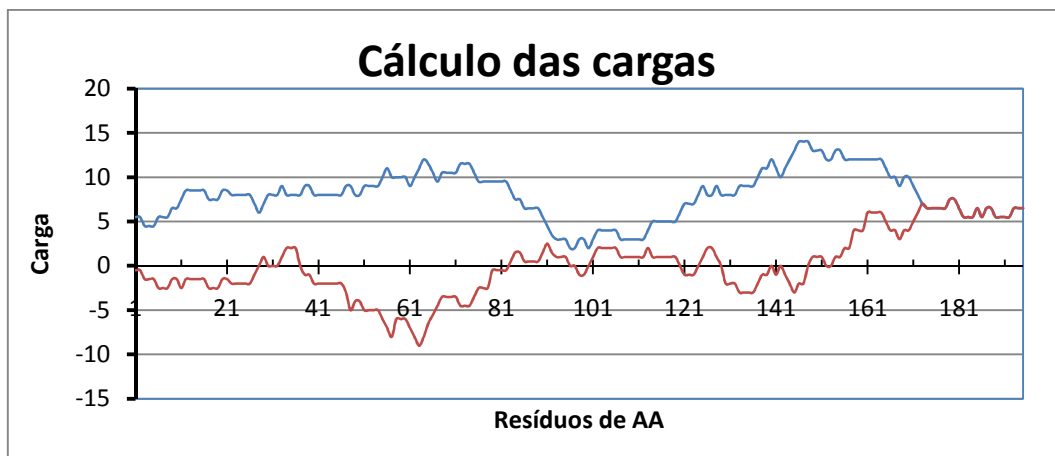


**Gráfico 74:** Representação gráfica da percentagem relativa dos aminoácidos da  $\alpha$ -caseína s2, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa a média das percentagens relativas da proteína  $\alpha$ -caseína s2.

Tal como a  $\alpha$ -caseína s1, a  $\alpha$ -caseína s2 mostra ter um comportamento hidrofílico devido a um maior número de aminoácidos carregados acima da média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína), representada a tracejado vermelho (Gráfico 74). Isto corrobora o facto de que a  $\alpha$ -caseína s2 é a mais hidrofílica de todas as caseínas (Jedrychowski & Wichers, 2009).

Entre os aminoácidos carregados encontram-se: a lisina, glutamina, glutamato, asparagina e serina. A valina é o único aminoácido acima da média de carácter hidrofóbico e que dificilmente rivaliza com os cinco carregados. Este tipo de composição reforça a ideia de que as caseínas têm uma periferia hidrofílica (Jedrychowski & Wichers, 2009; Hoffmann-Sommergruber e Mills, 2009) e das quais a  $\alpha$ -caseína s2 não é excepção. A tirosina e a treonina também têm valores acima da média mas não contribuem de modo significativo no comportamento global da proteína.

As cargas da  $\alpha$ -caseína s2 podem ser observadas através do gráfico 75.



**Gráfico 75:** Representação gráfica das cargas da  $\alpha$ -caseína s2. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

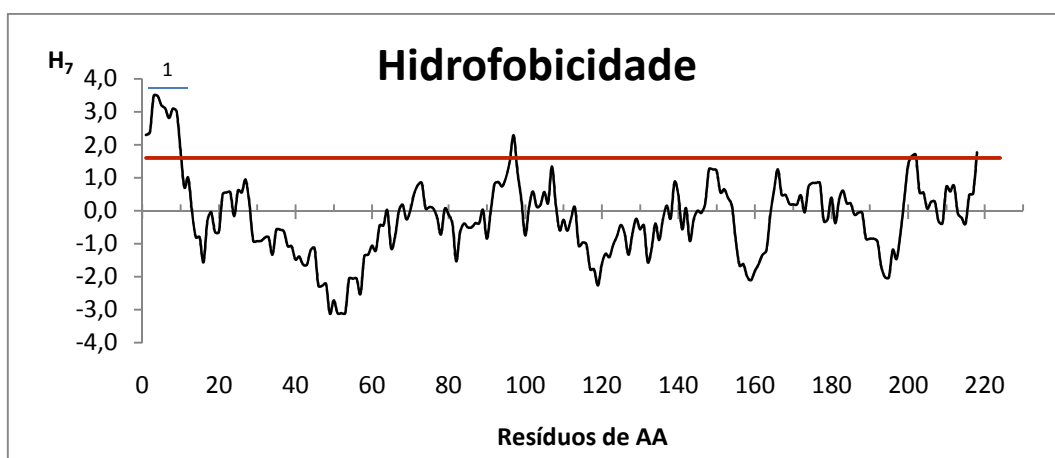
Observa-se que a  $\alpha$ -caseína s2 possui cargas reais positivas entre o intervalo 154-195. Existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores, como se fossem cargas residuais (Gráfico 75).

Observa-se igualmente que as cargas absolutas da  $\alpha$ -caseína s2, mostram que ela tem um comportamento polar positivo ao longo de toda a cadeia polipeptídica. Nota-se uma pequena descida do valor das cargas absolutas no intervalo 91-113. Tal como a  $\alpha$ -caseína s1, se a  $\alpha$ -caseína s2 for colocada numa solução com uma corrente eléctrica contínua, ela deslocar-se-á para o cátodo, (pólo negativo).

È de notar igualmente, que em algumas zonas as cargas reais e as cargas absolutas confundem-se. È o que acontece no intervalo: 173-195 (Gráfico 75).

#### 4.2.3.3. $\beta$ -caseína

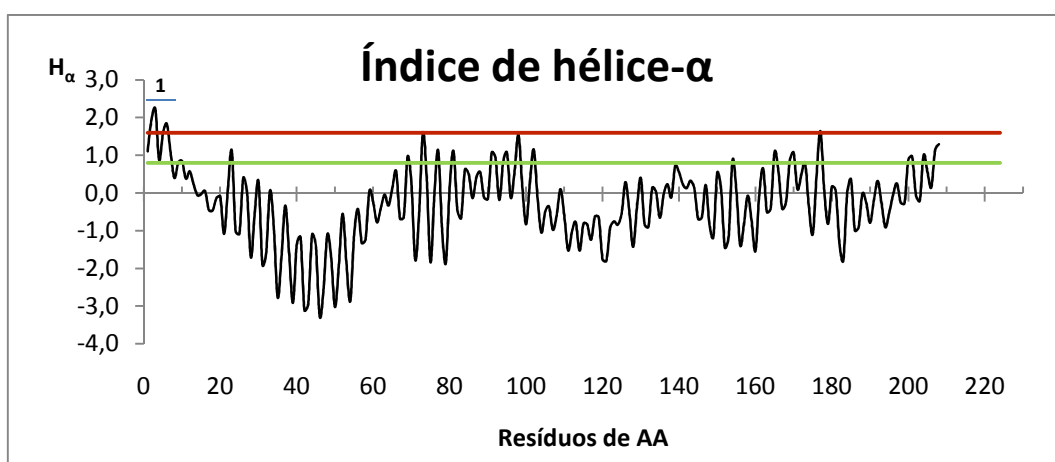
A hidrofobicidade da  $\beta$ -caseína pode ser demonstrada através do gráfico 76.



**Gráfico 76:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a  $\beta$ -caseína (224 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).

Excepto o pico 1, a  $\beta$ -caseína mostra ter um comportamento bastante hidrofílico ao longo da sua cadeia polipeptídica (Gráfico 76). Estes resultados contradizem os dados esperados visto que de todas as caseínas, a  $\beta$ -caseína é a mais hidrofóbica (Jedrychowski & Wichers, 2009; Belitz *et al.*, 2009). É de notar a presença de alguns picos hidrofóbicos a meio e no fim da cadeia de aminoácidos que ultrapassam o valor 1,6. Esses picos não são muito importantes visto que o número de resíduos de aminoácidos que os compõem é inferior a 19 e os aminoácidos vizinhos desses picos também não são muito hidrofóbicos (Han e Tashjian, 2003).

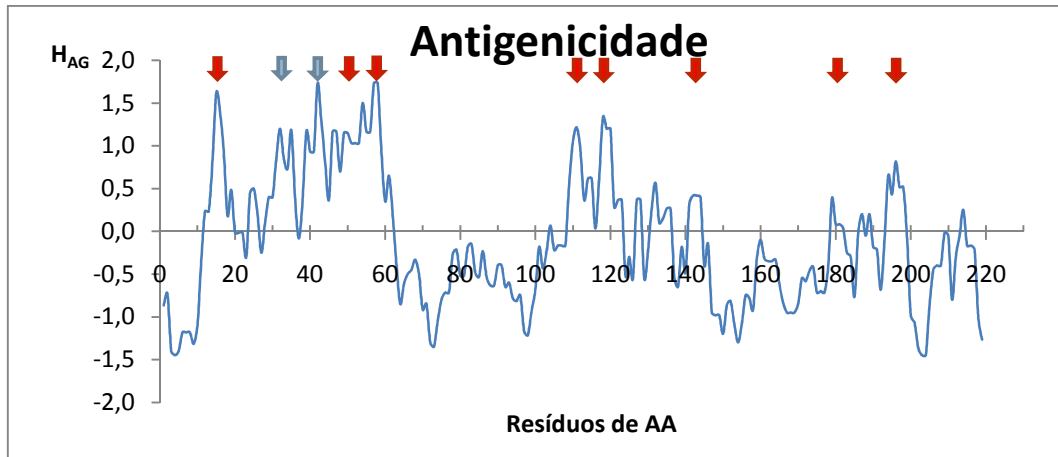
O índice de hélice- $\alpha$  pode ser representado no gráfico 77.



**Gráfico 77:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a  $\beta$ -caseína (224 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

No índice de hélices- $\alpha$  (Gráfico 77), observa-se a presença de um pico hidrofóbico inicial (1), que pode indicar a existência de um tipo particular de hélices- $\alpha$  que se situam no interior das proteínas globulares (Murray *et al.*, 2002). Entre os resíduos 80-110, observa-se um leve aumento da hidrofobicidade mas não existem resíduos suficientes nesse pico para se poder afirmar com toda a certeza de que se trata de uma hélice- $\alpha$  anfifílica.

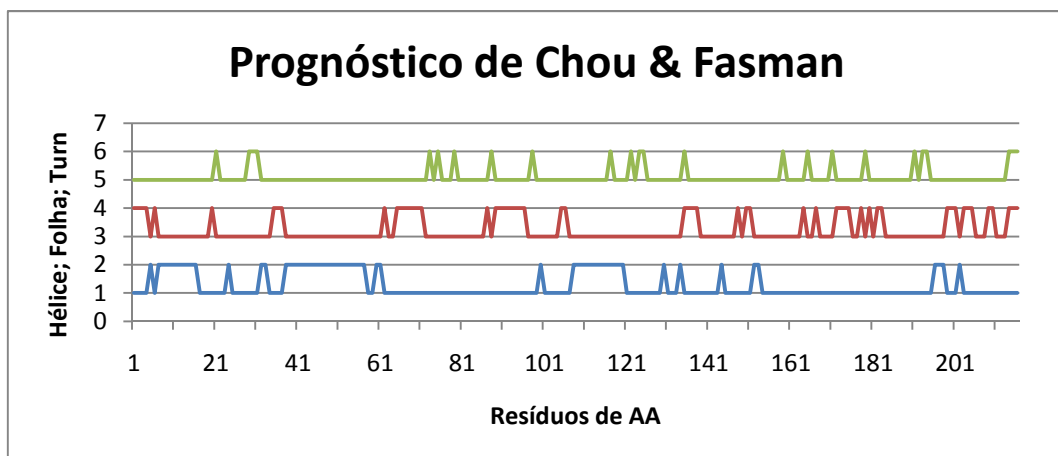
A antigenicidade da  $\beta$ -caseína pode ser representada segundo o gráfico 78.



**Gráfico 78:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a  $\beta$ -caseína, usando a escala de Hopp & Wood (azul). Zonas de ligação à IgE (setas vermelhas) e possíveis zonas de ligação antigénicas (setas azuis).

Observa-se a presença de oito picos (setas vermelhas) correspondentes a oito das nove zonas de ligação à IgE previstas (Gráfico 78). Para esta previsão, foram usados cem decapeptídeos sintéticos emparelhados, de maneira a poder averiguar quais as regiões da  $\beta$ -caseína susceptíveis de se ligarem às IgE. Os aminoácidos: 1-16, 45-54, 55-70, 83-92, 107-120, 135-144, 149-164, 167-184 e 185-208, foram descritos como característicos em pacientes com alergia persistente ao leite de vaca. (Jedrychowski & Wichers, 2009). Existem outros picos que indicam a probabilidade de serem zonas antigénicas (setas azuis).

O gráfico 79 revela as provisões do método de Chou & Fasman para os três elementos principais da estrutura secundária.

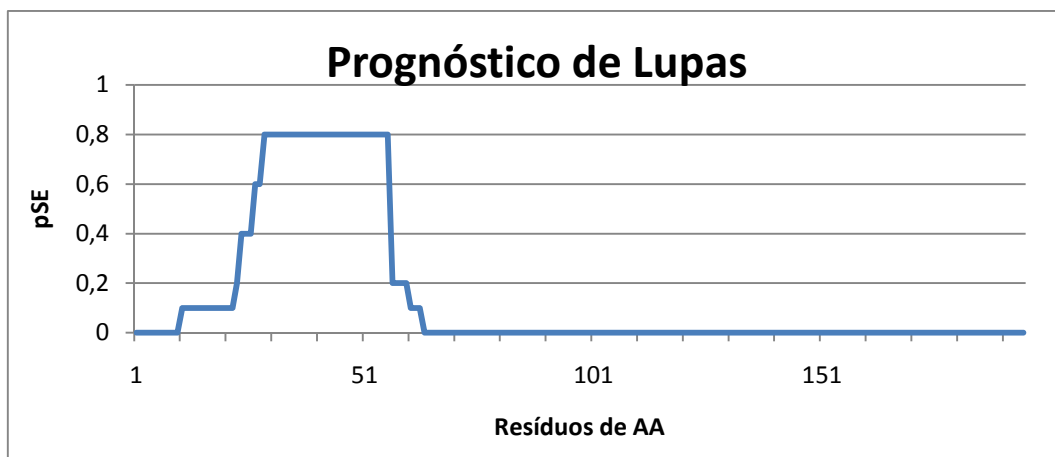


**Gráfico 79:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na  $\beta$ -caseína. Cada pico corresponde à presença de um possível elemento.

Observa-se que existe alguma probabilidade de encontrar hélices- $\alpha$  ao longo da cadeia polipeptídica da  $\beta$ -caseína (Gráfico 79). Este resultado confirma a presença de pelo menos 9% de hélices- $\alpha$  na estrutura da  $\beta$ -caseína (Belitz, 2009). Observa-se igualmente que os

resultados obtidos referentes às folhas- $\beta$ , confirmam o facto de que a  $\beta$ -caseína possui cerca de 25% de folhas- $\beta$  (Belitz, 2009). Quanto às voltas, o gráfico 79 mostra alguma probabilidade de se encontrarem mais do que uma. Todavia, este resultado não pode ser confirmado por dados teóricos.

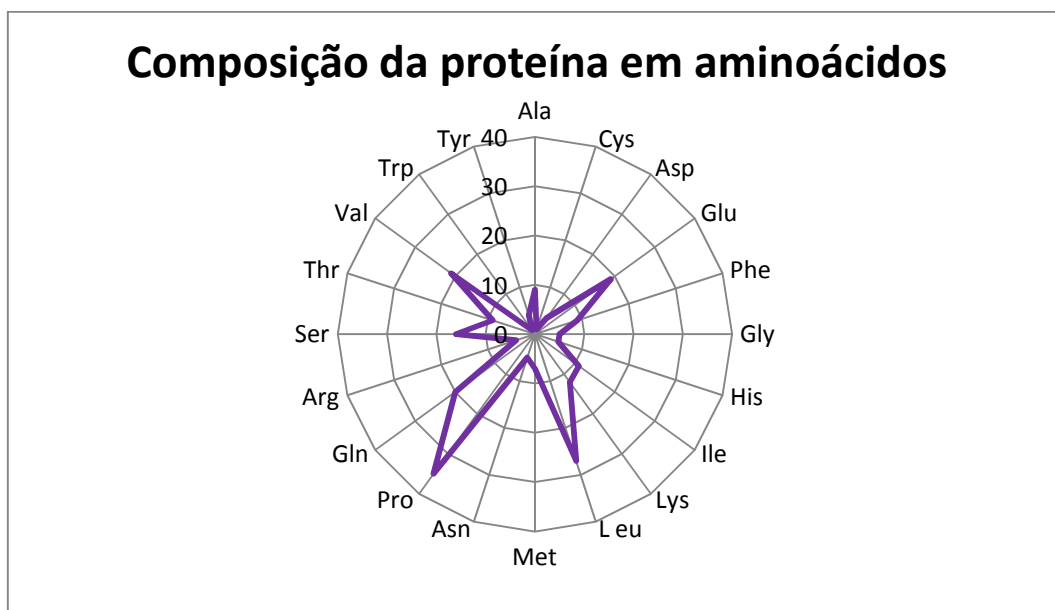
O prognóstico de Lupas pode ser apresentado no gráfico 80.



**Gráfico 80:** Representação gráfica do prognóstico de Lupas para as super-hélice- $\alpha$  na  $\beta$ -caseína. A cada pico corresponde a possível presença dessa estrutura.

Observa-se que existe a probabilidade de encontrar uma super-hélice- $\alpha$  na  $\beta$ -caseína (Gráfico 80). O único pico encontra-se entre os resíduos 11-63 (com uma probabilidade acrescida entre 30-56).

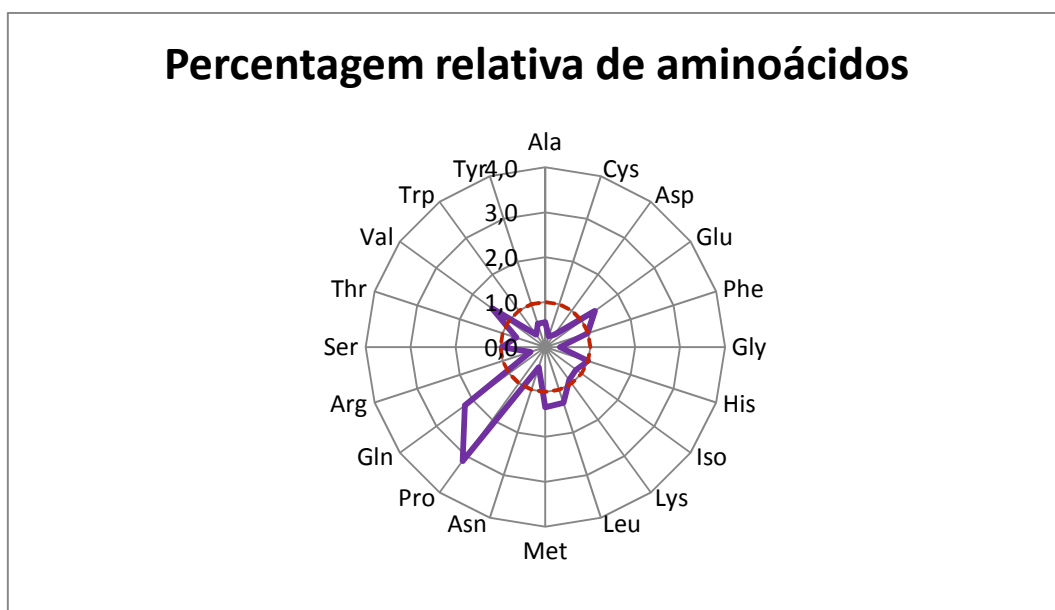
Para facilitar a observação da composição global da  $\beta$ -caseína dispuseram-se os aminoácidos através de um gráfico em radar (Gráfico 81).



**Gráfico 81:** Representação gráfica da quantidade real de cada aminoácido na  $\beta$ -caseína (224 resíduos de aminoácido).

A  $\beta$ -caseína é constituída pelos vinte principais aminoácidos (Gráfico 81). Sendo assim, a  $\beta$ -caseína é constituída maioritariamente por (ordem decrescente): prolina, leucina, valina, glutamina glutamato, serina, lisina e isoleucina. Tem uma quantidade média de: alanina, fenilalanina, treonina, metionina, glicina, histidina, asparagina, aspartato, arginina e tirosina. A  $\beta$ -caseína possui igualmente alguns aminoácidos residuais tais como: cisteína e triptofano. A grande quantidade de prolina vai favorecer a formação de voltas (Berg et al., 2002).

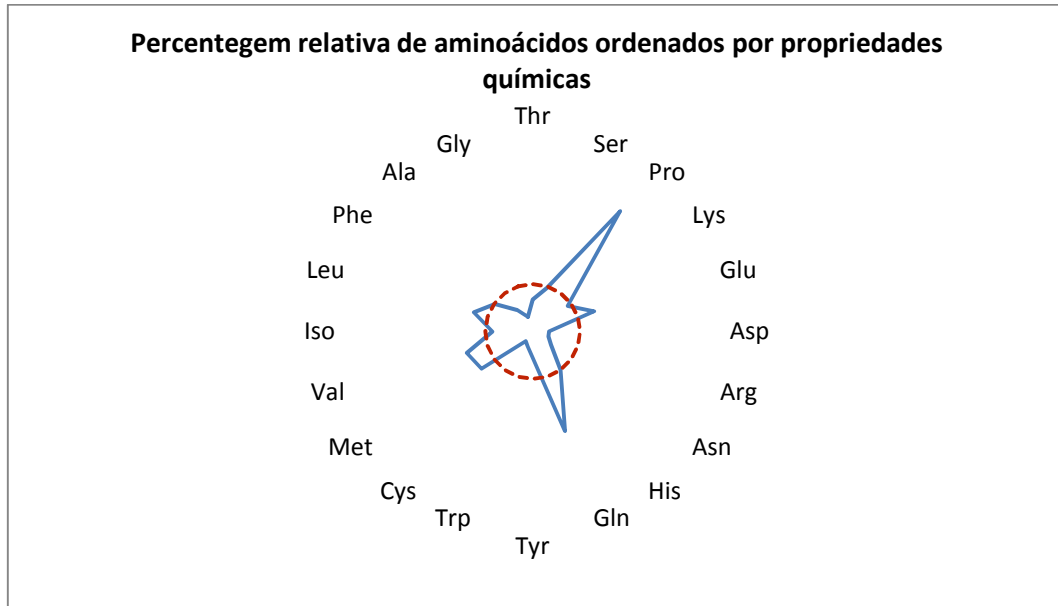
Através de um gráfico em radar é possível comparar a percentagem de aminoácidos da  $\beta$ -caseína relativamente à GenPep#97.



**Gráfico 82:** Representação gráfica da percentagem de cada aminoácido na  $\beta$ -caseína, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína. Isto facilita a observação da composição relativa desta proteína.

Observa-se que a  $\beta$ -caseína possui sete aminoácidos com uma percentagem relativa acima da média, tendo esta sido calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína e está representada a tracejado vermelho (Gráfico 82). Desses sete, o mais evidente é a prolina, seguida da glutamina, valina, glutamato, metionina, leucina e histidina. Os restantes aminoácidos têm um valor abaixo da média.

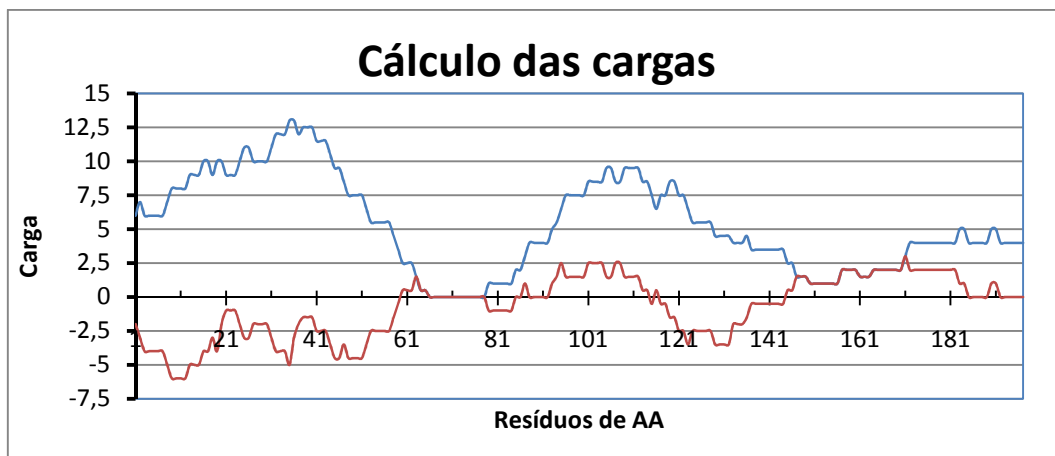
Uma das formas de verificar o comportamento da  $\beta$ -caseína, é dispor a percentagem relativa de aminoácidos de modo a que os aminoácidos hidrofóbicos fiquem à esquerda e os carregados fiquem à direita.



**Gráfico 83:** Representação gráfica da percentagem relativa dos aminoácidos da  $\beta$ -caseína, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

Observa-se que a  $\beta$ -caseína possui na sua composição, uma percentagem relativa de aminoácidos acima da média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína) de prolina, glutamina, valina, glutamato, metionina, leucina e histidina (Gráfico 83). Quatro destes aminoácidos são carregados (prolina, glutamina, glutamato e histidina). Para além disto, a  $\beta$  caseína também possui três aminoácidos acima da média com carácter hidrofóbico. Esses aminoácidos são: valina, metionina e leucina, que apesar de terem valores ligeiramente acima da média, têm os seus efeitos diminuídos pelos aminoácidos carregados.

As cargas reais e absolutas da  $\beta$  caseína são apresentadas no gráfico 84.



**Gráfico 84:** Representação gráfica das cargas da  $\beta$ -caseína. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

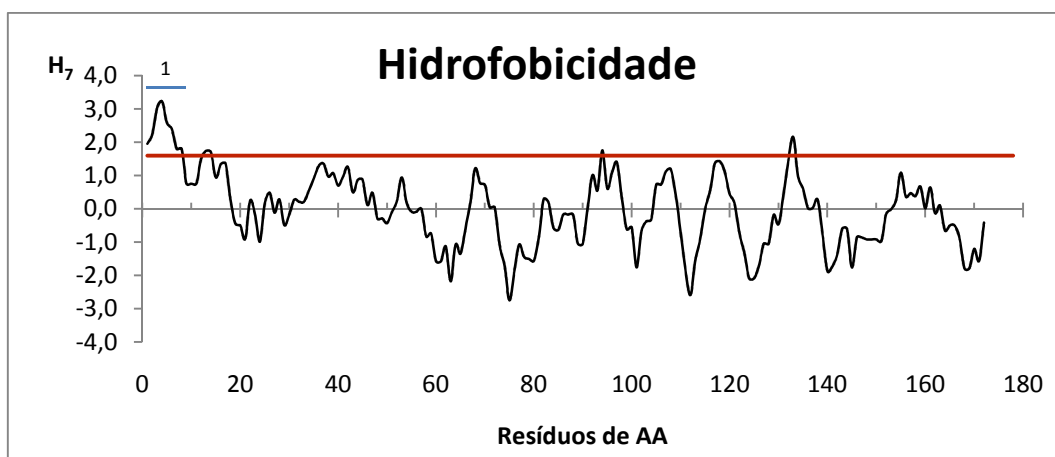
Observa-se que a  $\beta$ -caseína possui cargas reais positivas entre os dois intervalos: 93-114 e 145-184 (Gráfico 84). Existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores. Era de esperar altas cargas no N-terminal e C-terminal (Caessens, 1999).

Observa-se igualmente que as cargas absolutas da  $\beta$ -caseína mostram que ela tem um comportamento polar positivo ao longo de praticamente toda a cadeia polipeptídica, particularmente em duas zonas: 1-65 e 79-149. No entanto, observa-se um valor de cargas absolutas de zero, no intervalo 66-78. Se a  $\beta$ -caseína for colocada numa solução com corrente eléctrica, ela deslocar-se-á para o cátodo (pólo negativo).

É de notar igualmente, que em algumas zonas as cargas reais e as cargas absolutas confundem-se. É o que acontece nos intervalos: 63-78 e 147-171 (Gráfico 84).

4.2.3.4.  $\beta$ -lactoglobulina

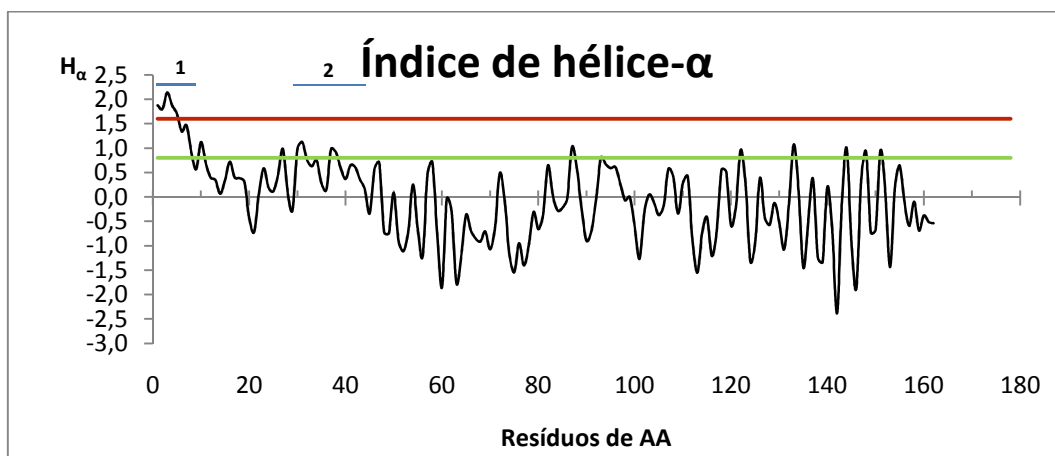
A hidrofobicidade da  $\beta$ -lactoglobulina pode ser apresentada da seguinte forma.



**Gráfico 85:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a  $\beta$ -lactoglobulina (178 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).

A  $\beta$ -lactoglobulina mostra ter um comportamento mais hidrofóbico que as proteínas do leite observadas anteriormente (Gráfico 85). Verifica-se que ela tem um pico hidrofóbico (1), visto que ultrapassa o valor de 1,6 e possui mais de 19 aminoácidos necessários para ser considerado um pico hidrofóbico.

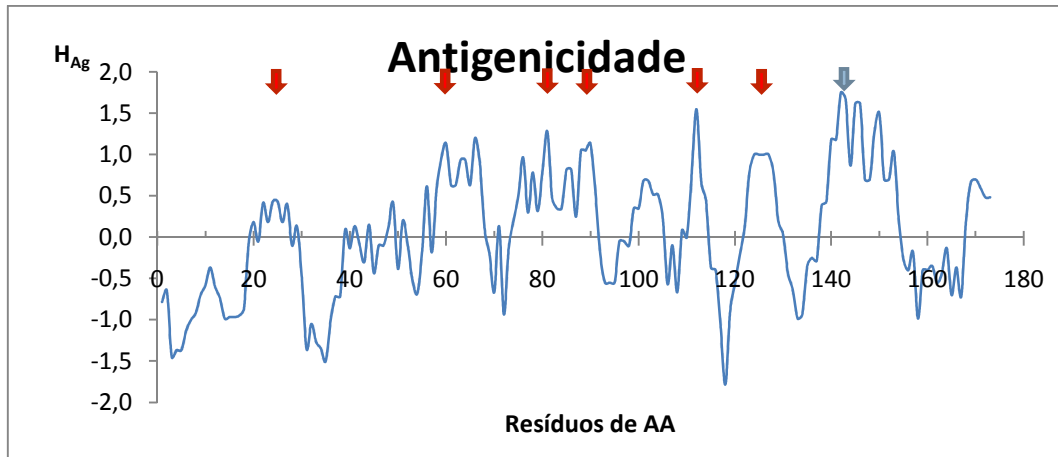
No gráfico 86, observa-se o resultado dos cálculos para o índice de hélice- $\alpha$ .



**Gráfico 86:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a  $\beta$ -lactoglobulina (178 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

Relativamente ao índice de hélice- $\alpha$ , observa-se um pico hidrofóbico no início da proteína visto que ultrapassa o valor de 1,6 e possui mais de 19 aminoácidos. No pico 2, os valores de hidrofobicidade situam-se entre o valor 0,8 (linha verde) e 1,6 (linha vermelha). Estes dois picos podem revelar a presença de hélice- $\alpha$  anfífilas que se situam geralmente no interior da proteína.

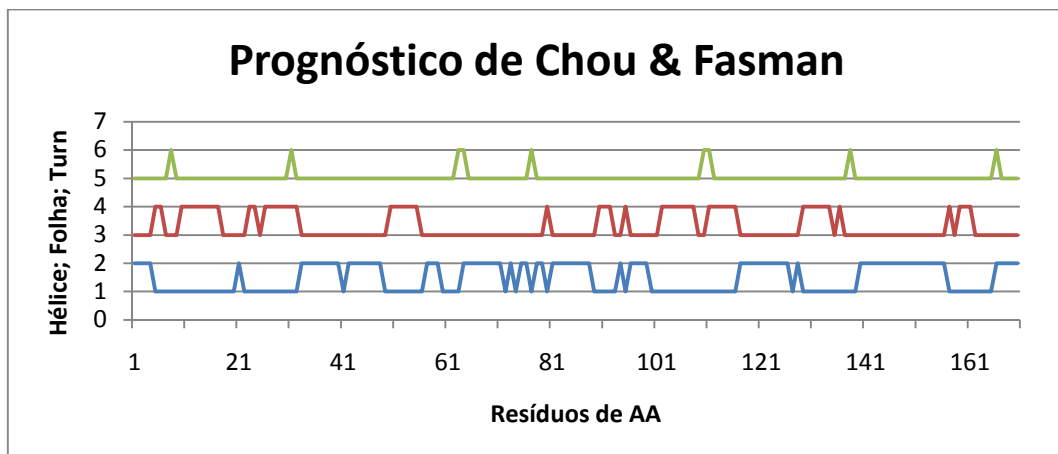
A antigenicidade da  $\beta$ -lactoglobulina pode ser observada através do gráfico 87.



**Gráfico 87:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a  $\beta$ -lactoglobulina, usando a escala de Hopp & Wood (azul). Zonas antigénicas conhecidas (setas vermelhas) e zona antigénica possível (seta azul).

A previsão antigénica revela seis picos (setas vermelhas) que marcam a presença de seis antigénios (Gráfico 87). Esta presença era bastante previsível visto que ao longo da cadeia polipeptídica da  $\beta$ -lactoglobulina, existem epítomos principais para a IgE humana. Esses epítomos são formados pelos resíduos de aminoácidos: 41-60, 102-124, os intermediários: 25-40 e com menor influencia: 84-91, 125-135 e 78-83 (Jedrychowski & Wichers, 2009). Para além disso, a  $\beta$ -lactoglobulina tem a capacidade de se ligar aos lípidos, incluindo: o retinol, a  $\beta$ -caroteno, os ácidos gordos saturados e insaturados e aos hidrocarbonetos alifáticos (Ehn et al., 2004). A maioria destes picos é corroborada pelos dados do gráfico 86. Para além disso, existe um pico que não se inclui nos epítomos descritos na literatura mas que revela uma certa probabilidade de se encontrar uma ligação antigénica nessa zona (seta azul).

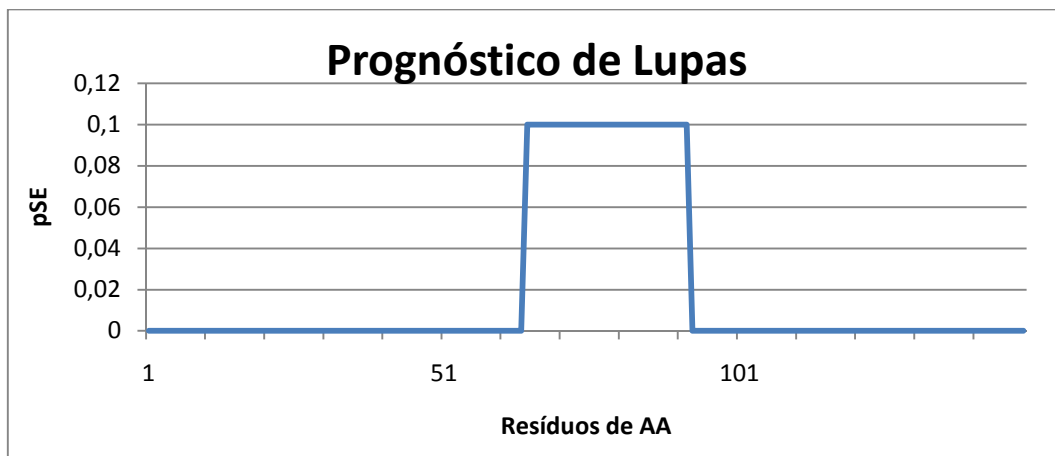
O gráfico 88 demonstra a aplicação do método de Chou & Fasman.



**Gráfico 88:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na  $\beta$ -lactoglobulina. Cada pico corresponde à presença de um possível elemento.

Observa-se que existe uma grande probabilidade de encontrar hélices- $\alpha$  e folhas- $\beta$  ao longo da cadeia polipeptídica da  $\beta$ -lactoglobulina (Gráfico 88). No entanto, a probabilidade de encontrar voltas é substancialmente inferior. Nenhum deste facto pode ser confirmado à luz de dados teóricos. Todavia, chama-se à atenção de que o método de Chou & Fasman não foi muito conclusivo na detecção das folhas- $\beta$  relativamente à BSA.

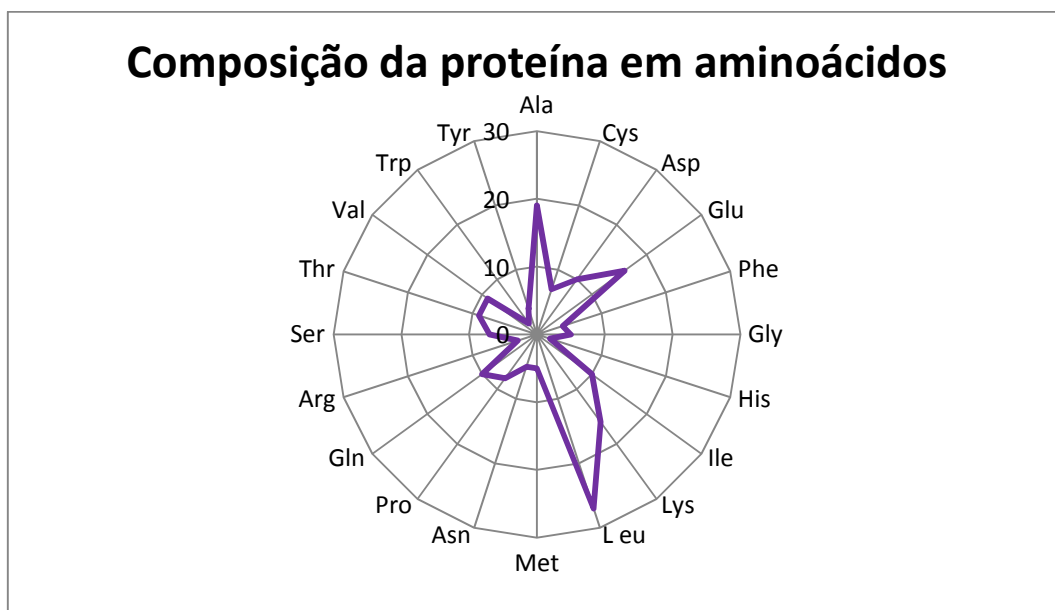
O prognóstico de Lupas encontra-se no gráfico 89.



**Gráfico 89:** Representação gráfica do prognóstico de Lupas para as super-hélice- $\alpha$  na  $\beta$ -lactoglobulina. A cada pico corresponder a possível presença dessa estrutura.

Observa-se que existe alguma probabilidade de encontrar uma super-hélice- $\alpha$  na  $\beta$ -lactoglobulina (Gráfico 89). O único pico encontra-se entre os resíduos 65-92.

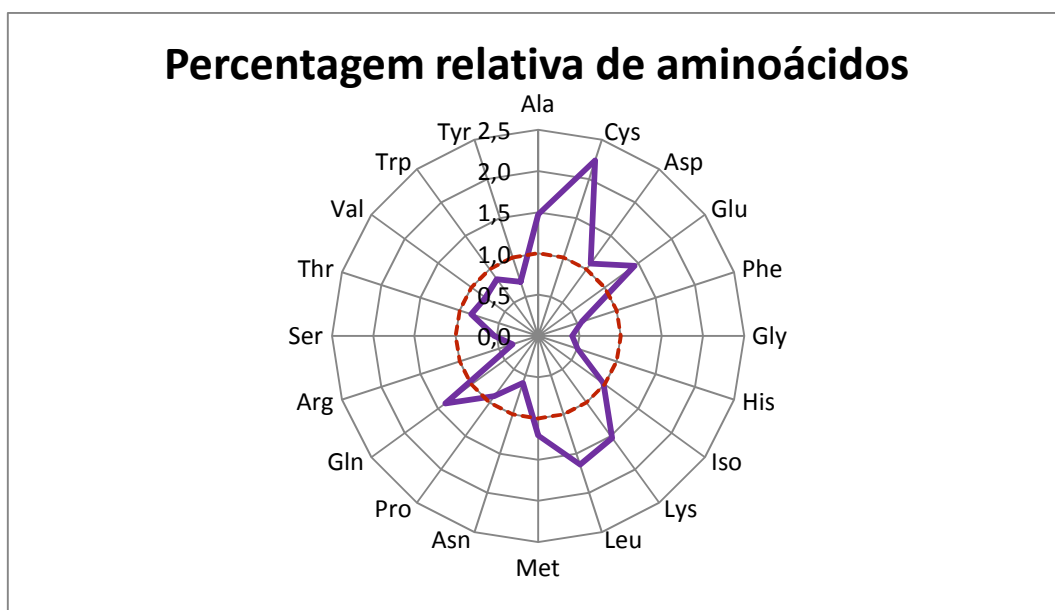
Para facilitar a observação da composição global da  $\beta$ -lactoglobulina, representaram-se os aminoácidos e respectiva quantidade através de um gráfico em radar (Gráfico 90).



**Gráfico 90:** Representação gráfica da quantidade real de cada aminoácido na  $\beta$ -lactoglobulina (178 resíduos de aminoácido).

A  $\beta$ -lactoglobulina é constituída pelos vinte principais aminoácidos (Gráfico 90). Sendo assim, a  $\beta$ -lactoglobulina é constituída maioritariamente por (ordem decrescente): leucina, alanina, glutamato, lisina, aspartato, isoleucina e glutamina. Tem uma quantidade média de: treonina, valina, prolina, cisteína, serina, glicina, metionina, asparagina, fenilalanina, tirosina, arginina, histidina e triptofano. Visto que a  $\beta$ -lactoglobulina é rica em leucina, existe uma forte suspeita de que ela tenha pelo menos uma hélice- $\alpha$  na sua estrutura (Berg et al., 2002).

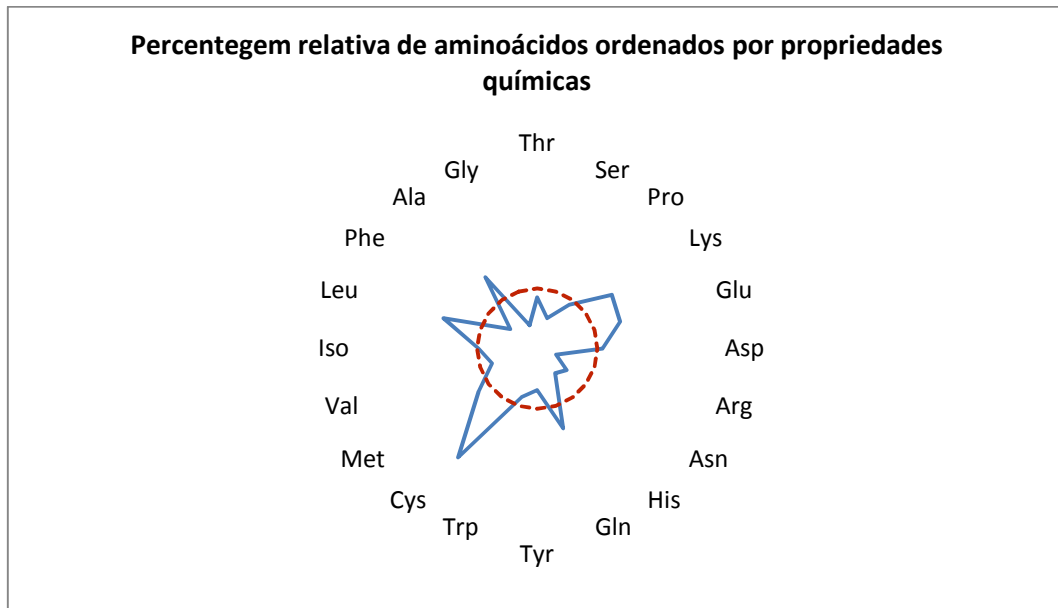
Através de um gráfico em radar é possível comparar a percentagem de aminoácidos da  $\beta$ -lactoglobulina relativamente à GenPep#97 ( Gráfico 91).



**Gráfico 91:** Representação gráfica da percentagem de cada aminoácido na  $\beta$ -lactoglobulina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína. Isto facilita a observação da composição relativa desta proteína.

Observa-se que a  $\beta$ -lactoglobulina possui oito aminoácidos com uma percentagem relativa acima da média, representada a tracejado vermelho, tendo sido calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína (Gráfico 91). Desses oito, os mais evidentes são a cisteína, seguida da leucina, lisina, alanina, glutamato, glutamina, metionina e aspartato. Os restantes aminoácidos têm um valor abaixo da média.

Uma das formas de verificar o comportamento da  $\beta$ -lactoglobulina, é dispor a percentagem relativa de aminoácidos de modo a que os aminoácidos hidrofóbicos fiquem à esquerda e os carregados fiquem à direita (Gráfico 92).



**Gráfico 92:** Representação gráfica da percentagem relativa dos aminoácidos da  $\beta$ -lactoglobulina, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

Ao contrário das outras três proteínas do leite vistas até aqui, pode dizer-se que a  $\beta$ -lactoglobulina parece ser menos solúvel (Gráfico 92). Apesar de haver quatro tipos de aminoácidos carregados acima da média e outros quatro hidrofóbicos, os valores numéricos dos aminoácidos hidrofóbicos são superiores. Entre os aminoácidos carregados temos: lisina, glutamato, glutamina, e aspartato. Entre os aminoácidos hidrofóbicos temos: cisteína, leucina, alanina e metionina.

As cargas reais e absolutas da  $\beta$ -lactoglobulina encontram-se representadas no gráfico 93.

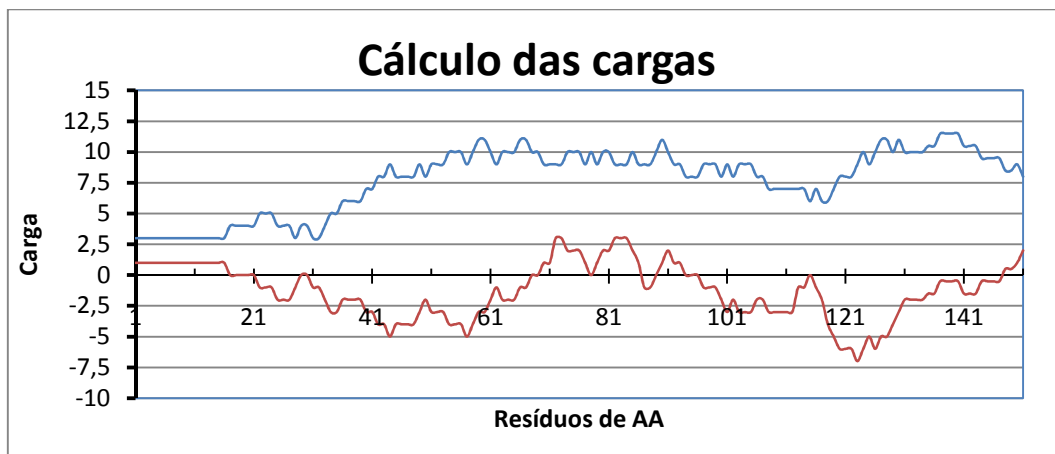


Gráfico 93: Representação gráfica das cargas da  $\beta$ -lactoglobulina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

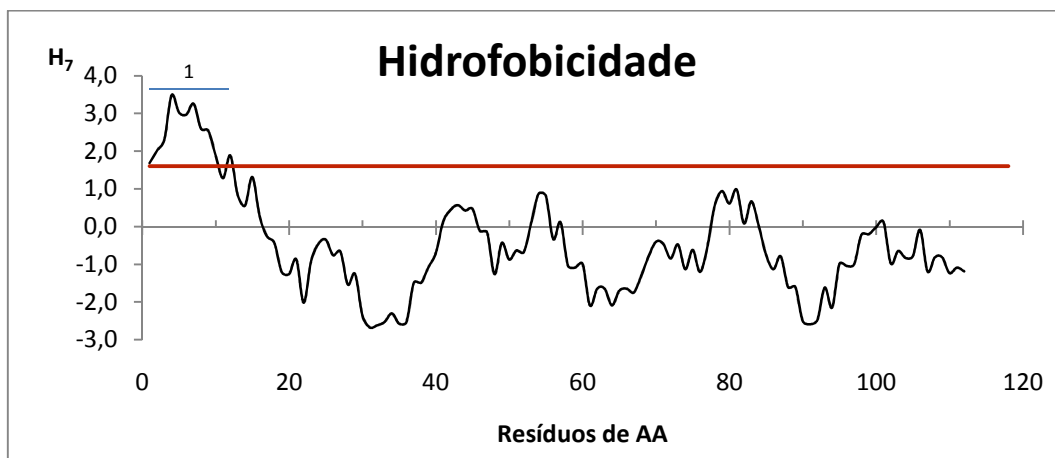
Observa-se que a  $\beta$ -lactoglobulina possui poucos picos com cargas reais positivas (Gráfico 93). No entanto, destacam-se dois intervalos 70-77 e 79-86. Existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores. Nomeadamente, o intervalo 1-16, onde a carga é mantida sempre a valor 1.

Observa-se igualmente que as cargas absolutas da  $\beta$ -lactoglobulina mostram ter um comportamento polar positivo ao longo de toda a cadeia polipeptídica, em particular a partir do resíduo 22 onde elas têm um valor acrescido. Tal como as restantes proteínas do leite, se a  $\beta$ -lactoglobulina for colocada numa solução com uma corrente eléctrica, ela deslocar-se-á para o cátodo (pólo negativo).

#### 4.2.4. Proteínas do organismo

##### 4.2.4.1. $\beta$ -2-microglobulina

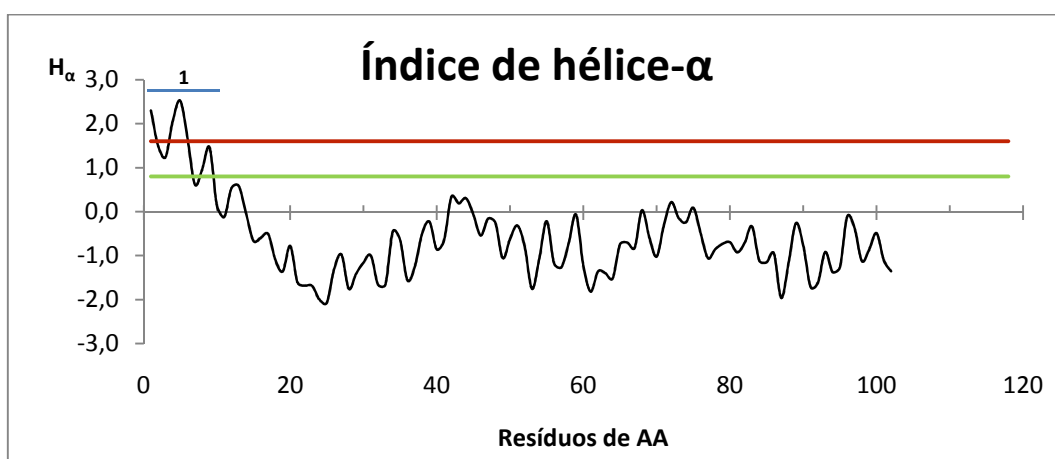
A hidrofobicidade  $H_7$  para a  $\beta$ -2-microglobulina pode ser visualizada no gráfico 94.



**Gráfico 94:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a  $\beta$ -2-microglobulina (118 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).

Observa-se que a  $\beta$ -2-microglobulina tem um comportamento hidrofílico ao longo da sua sequência de aminoácidos (Gráfico 94), excepto no início, (1), onde a sua hidrofobicidade atinge valores muito acima do valor referenciado de 1,6. Os restantes picos observados não têm nem um número suficiente de aminoácidos nem uma hidrofobicidade relevante para serem levados em conta. Este comportamento era totalmente previsível, visto que a  $\beta$ -2-microglobulina ( $\beta$ 2m), é a cadeia lateral solúvel do MHC (Main Histocompatibility Complex) classe I (Quesada et al., 1998).

No gráfico 95, observa-se o resultado dos cálculos para o índice de hélice- $\alpha$ .



**Gráfico 95:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a  $\beta$ -2-microglobulina (118 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

A presença de um pico no início do índice de hélice- $\alpha$  (Gráfico 95; pico 1) pode revelar a existência de uma hélice- $\alpha$  interna da proteína. Ao longo do resto da cadeia polipeptídica, a  $\beta$ -2-microglobulina mostra ter um comportamento hidrofílico.

A antigenicidade da  $\beta$ -2-microglobulina pode ser representada através do gráfico 96.

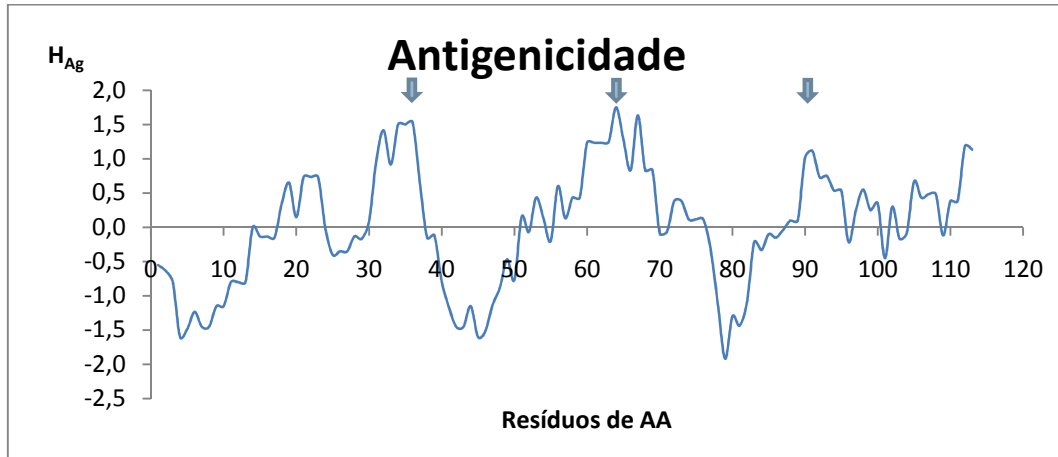


Gráfico 96: Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a  $\beta$ -2-microglobulina, usando a escala de Hopp & Wood (azul). Zonas antigénicas possíveis (setas azuis).

Observa-se que no gráfico 96 estão presentes três picos susceptíveis de indicar a presença de zonas de ligação antigénicas.

O algoritmo de Chou & Fasman está representado no gráfico 97.

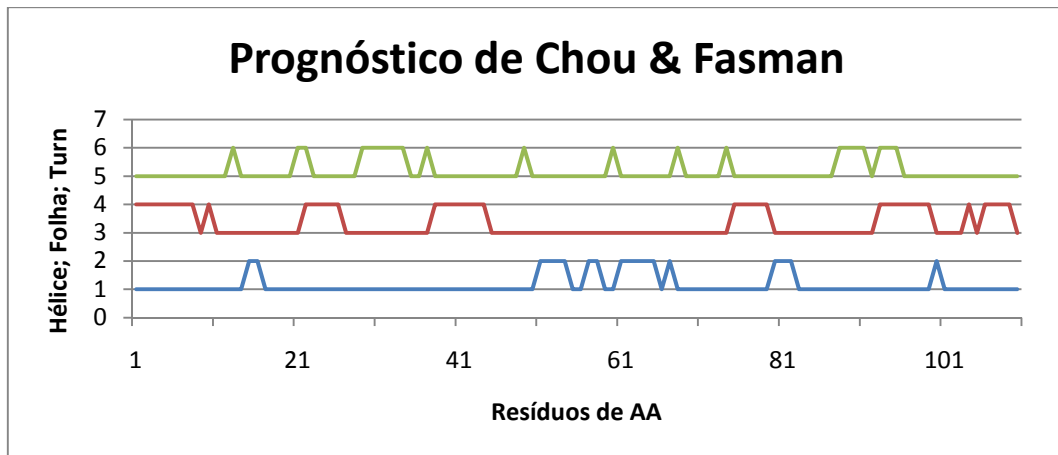


Gráfico 97: Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na  $\beta$ -2-microglobulina. Cada pico corresponde à presença de um possível elemento.

Observa-se que existe alguma probabilidade de encontrar cinco hélices- $\alpha$ , seis folhas- $\beta$  e quatro voltas ao longo da cadeia polipeptídica da  $\beta$ -2-microglobulina (Gráfico 97). Estes valores não puderam ser comparados com valores teóricos.

Depois de efectuados todos os cálculos, o algoritmo de Lupas encontram-se no gráfico 98.

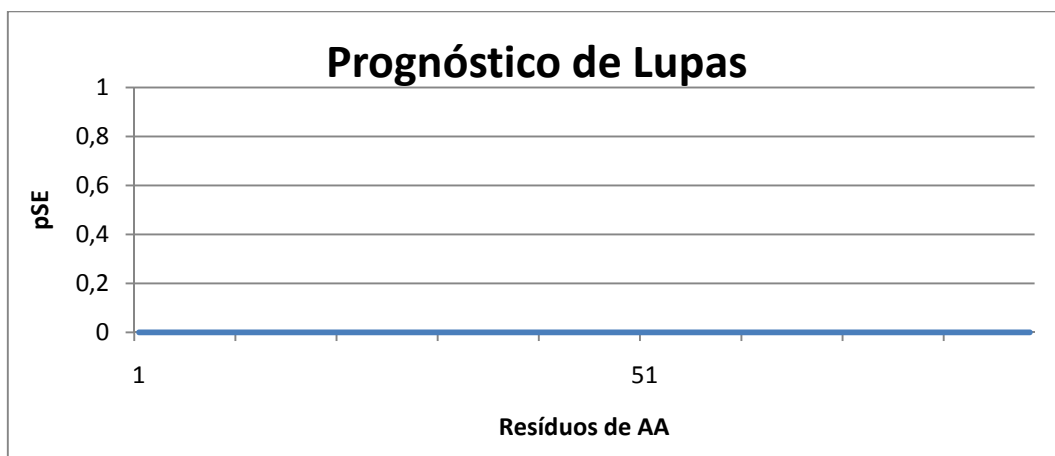


Gráfico 98: Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na  $\beta$ -2-microglobulina.

Observa-se que a probabilidade de encontrar uma super-hélice- $\alpha$  na  $\beta$ -2-microglobulina através do método de Lupas é nula (Gráfico 98).

O gráfico 99 ajuda-nos a visualizar a composição das proteínas em aminoácidos.

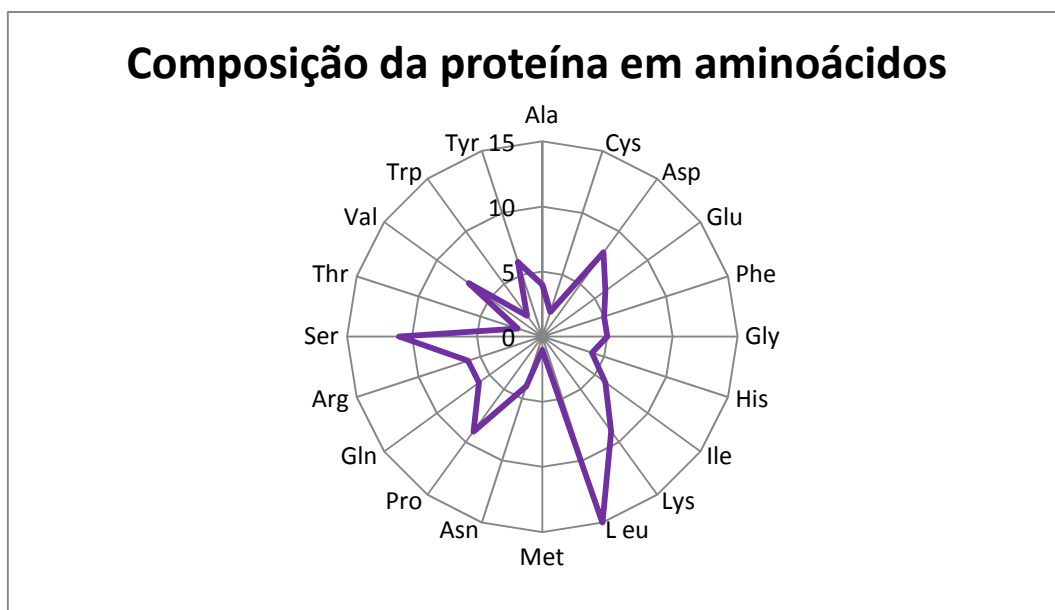
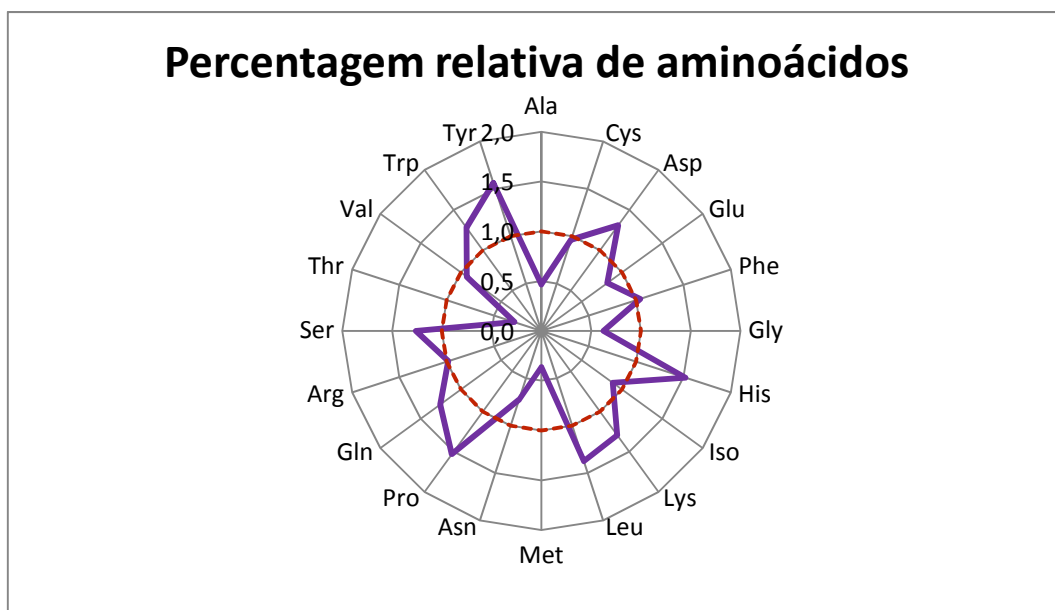


Gráfico 99: Representação gráfica da quantidade real de cada aminoácido na  $\beta$ -2-microglobulina (118 resíduos de aminoácido).

A  $\beta$ -2-microglobulina é constituída pelos vinte principais aminoácidos (Gráfico 99). Sendo assim, a  $\beta$ -2-microglobulina é constituída maioritariamente por (ordem decrescente): leucina e serina. Tem uma quantidade relativa de: lisina, prolina, aspartato, valina, glutamato, isoleucina, glutamina, arginina, tirosina, fenilalanina e glicina. Possui um valor mínimo de: alanina, histidina, asparagina, cisteína, treonina, triptofano e metionina. Tendo em conta a grande quantidade de leucina, é fortemente provável que a  $\beta$ -2-microglobulina

tenha uma ou mais hélices- $\alpha$  na sua estrutura (Berg et al., 2002). No entanto, a grande quantidade de serina pode condicionar a formação desse mesmo tipo de estrutura (Berg et al., 2002). A quantidade média da valina e isoleucina favorece a formação de folhas- $\beta$  mas pode ser perturbada pela prolina (Berg et al., 2002) Essa prolina, por sua vez, iria favorecer a formação de voltas (Berg et al., 2002).

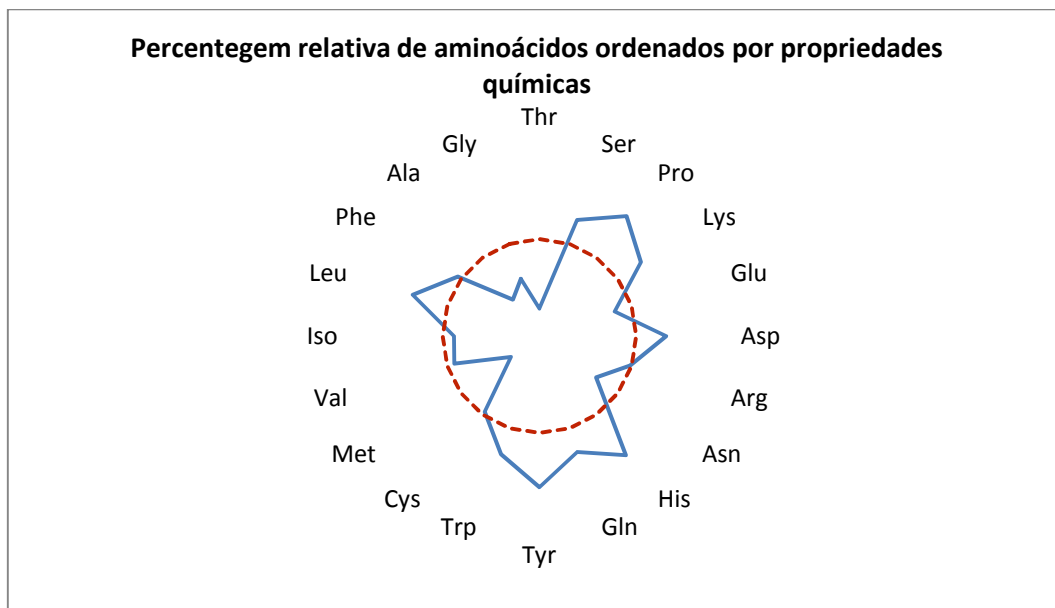
O gráfico 100 mostra a composição relativa em aminoácidos da  $\beta$ -2-microglobulina em relação à GenPep#97.



**Gráfico 100:** Representação gráfica da percentagem de cada aminoácido na  $\beta$ -2-microglobulina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na  $\beta$ -2-microglobulina. Isto facilita a observação da composição relativa desta proteína.

Observa-se que a  $\beta$ -2-microglobulina possui dez aminoácidos com uma percentagem relativa acima da média, representada a tracejado vermelho, tendo sido calculada a partir das percentagens relativas dos 20 aminoácidos presentes na  $\beta$ -2-microglobulina (Gráfico 100). Desses dez, o mais evidente é a tirosina, seguida da prolina, histidina, leucina, aspartato, lisina, triptofano, serina, glutamina e fenilalanina. Os restantes aminoácidos têm um valor abaixo da média.

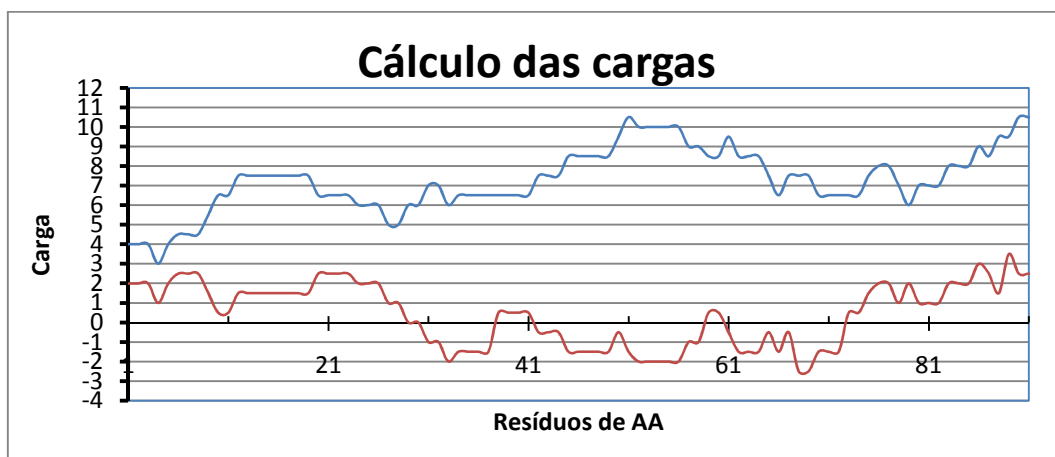
A distribuição dos aminoácidos segundo as suas propriedades químicas, revela algumas características da  $\beta$ -2-microglobulina.



**Gráfico 101:** Representação gráfica da percentagem relativa dos aminoácidos da  $\beta$ -2-microglobulina, mediante as suas propriedades químicas. Os aminoácidos mais carregados situam-se à direita e os mais hidrofóbicos situam-se à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na  $\beta$ -2-microglobulina.

Relativamente aos aminoácidos acima da média (tracejado vermelho), observa-se que os aminoácidos carregados estão em maior número e em percentagens relativas superiores aos aminoácidos hidrofóbicos que estão em número inferior (Gráfico 101). Isto evidencia o comportamento hidrofílico da  $\beta$ -2-microglobulina.

As cargas da  $\beta$ -2-microglobulina podem ser observadas através do gráfico 102.



**Gráfico 102:** Representação gráfica das cargas da  $\beta$ -2-microglobulina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

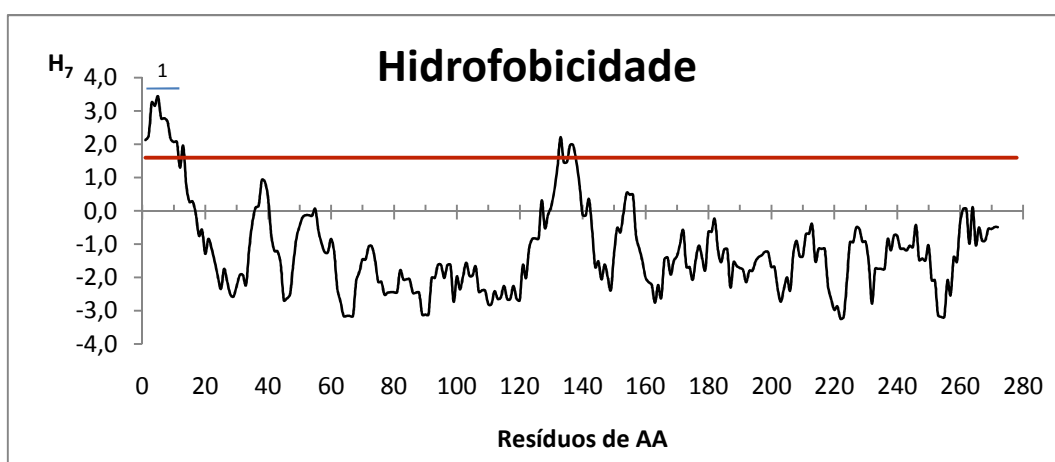
Destaca-se que a  $\beta$ -2-microglobulina possui duas zonas principais com cargas reais positivas. Essas zonas são definidas pelos intervalos 1-28 e 73-91. Existem outros intervalos

onde as cargas reais também são positivas mas são muito mais curtos e de cargas inferiores (Gráfico 102).

Observa-se igualmente que as cargas absolutas da  $\beta$ -2-microglobulina, mostram que ela tem um comportamento polar positivo ao longo de toda a cadeia polipeptídica. Isto significa que se ela for colocada numa solução com uma corrente eléctrica, ela deslocar-se-á para o cátodo.

#### 4.2.4.2. Osteopontina

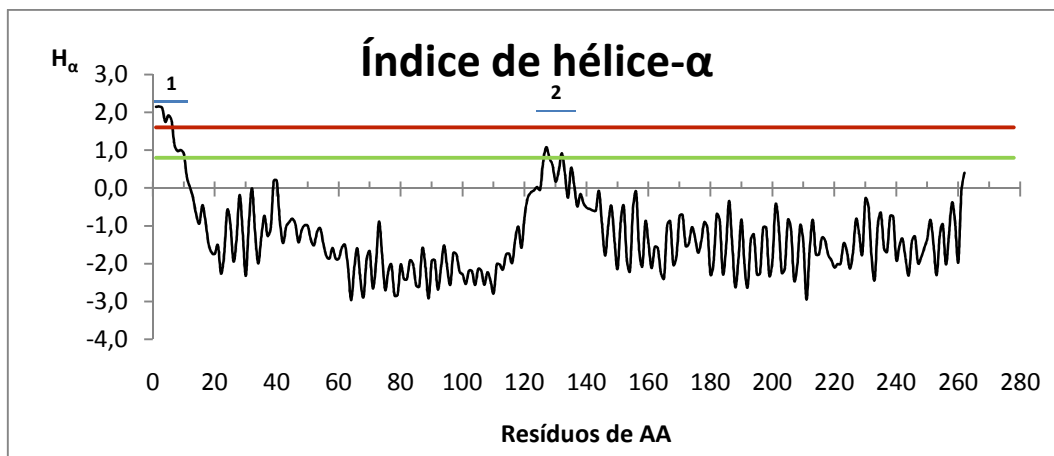
No gráfico 103 encontra-se o resultado do estudo de hidrofobicidade ( $H_7$ ) para a osteopontina.



**Gráfico 103:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a osteopontina (278 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).

A osteopontina apresenta um pico no início da sua cadeia polipeptídica (Gráfico 103, pico 1) e outro no intervalo 130-140. Se no primeiro pico se torna evidente a existência de uma zona hidrofóbica devido ao facto deste ultrapassar o nível de referência 1,6 o segundo pico não é muito relevante, visto que não possui uma quantidade mínima de 19 resíduos de aminoácidos para ser considerado um pico hidrofóbico. Isto apesar de também ele ultrapassar o valor de referência de 1,6 (linha vermelha). À luz dos dados teóricos a osteopontina é geralmente considerada hidrofílica (Rodrigues *et al.*, 2007; Wang e Denhardt, 2008), o que está de acordo com o gráfico apresentado.

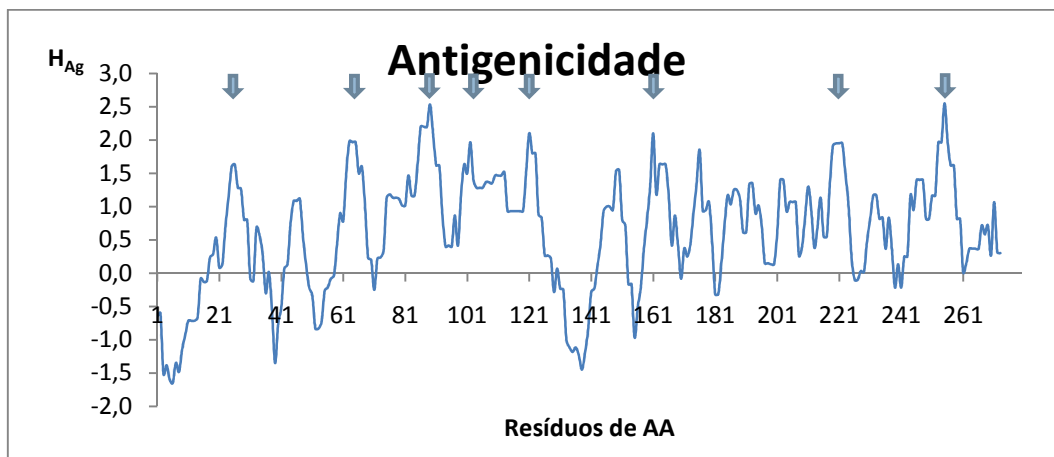
As hélices- $\alpha$  podem ser detectadas através do índice de hélices- $\alpha$  (Gráfico 104).



**Gráfico 104:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a osteopontina (278 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

No índice de hélices- $\alpha$  (Gráfico 104), observam-se dois picos. No pico 1 observa-se nitidamente um comportamento hidrofóbico uma vez que, ultrapassa o valor referenciado de 1,6 (linha vermelha). O pico dois manifesta um comportamento anfifílico, visto que, tem um valor superior a 0,8 mas inferior a 1,6, o que poderá indicar a presença de uma hélice- $\alpha$  anfifílica. Estes tipos de hélices- $\alpha$  são particulares devido ao facto de terem na sua composição aminoácidos hidrofóbicos e hidrofílicos (Murray *et al.*, 2002).

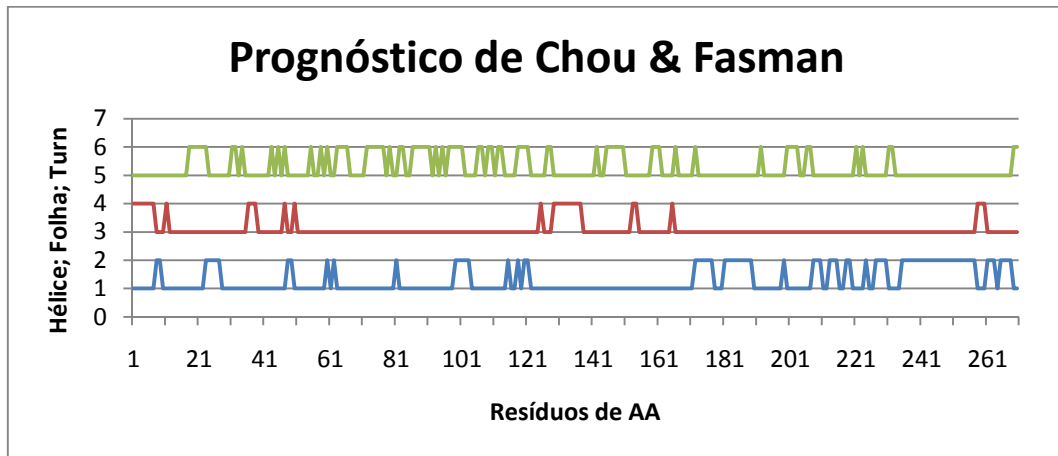
A antigenicidade da osteopontina pode ser representada graficamente do seguinte modo.



**Gráfico 105:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a osteopontina, usando a escala de Hopp & Wood. Zonas antigénicas possíveis (setas azuis).

Observa-se oito picos que prevêm a existência de oito zonas antigénicas na osteopontina, (Gráfico 105). Tendo em conta as inúmeras funções que a osteopontina desempenha no organismo, nomeadamente a ligação a integrinas, a presença destas oito zonas de ligação é totalmente possível (Higashikawa *et al.*, 2007; Rodrigues *et al.*, 2007; Wang e Denhardt, 2008).

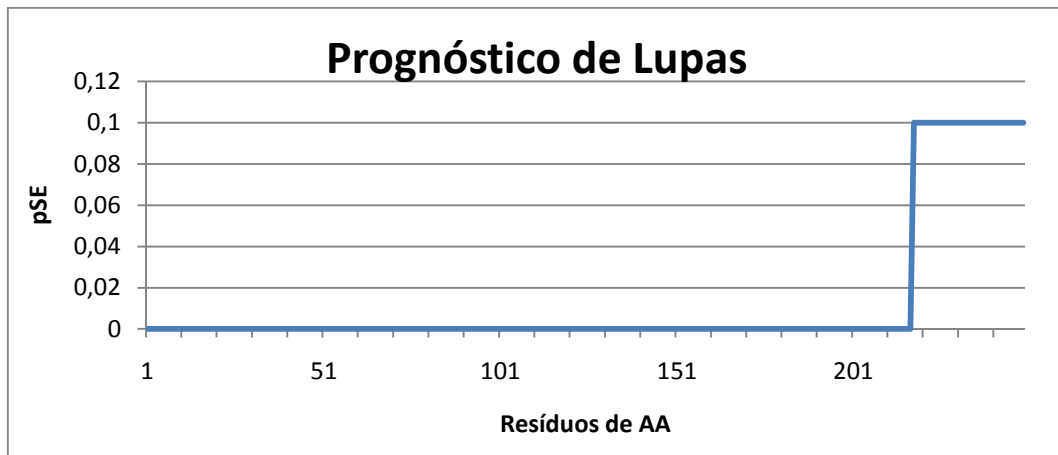
O gráfico 106 revela as previsões do método de Chou & Fasman para os três elementos principais da estrutura secundária.



**Gráfico 106:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na osteopontina. Cada pico corresponde à presença de um possível elemento.

Observa-se que existe uma grande probabilidade de encontrar hélices- $\alpha$  e voltas ao longo da cadeia polipeptídica da osteopontina, (Gráfico 106). No entanto, existe uma fraca probabilidade de encontrar folhas- $\beta$ . Nenhum destes resultados pôde ser confirmados à luz de dados teóricos visto que nada foi encontrado referente a este tipo de estruturas para esta proteína.

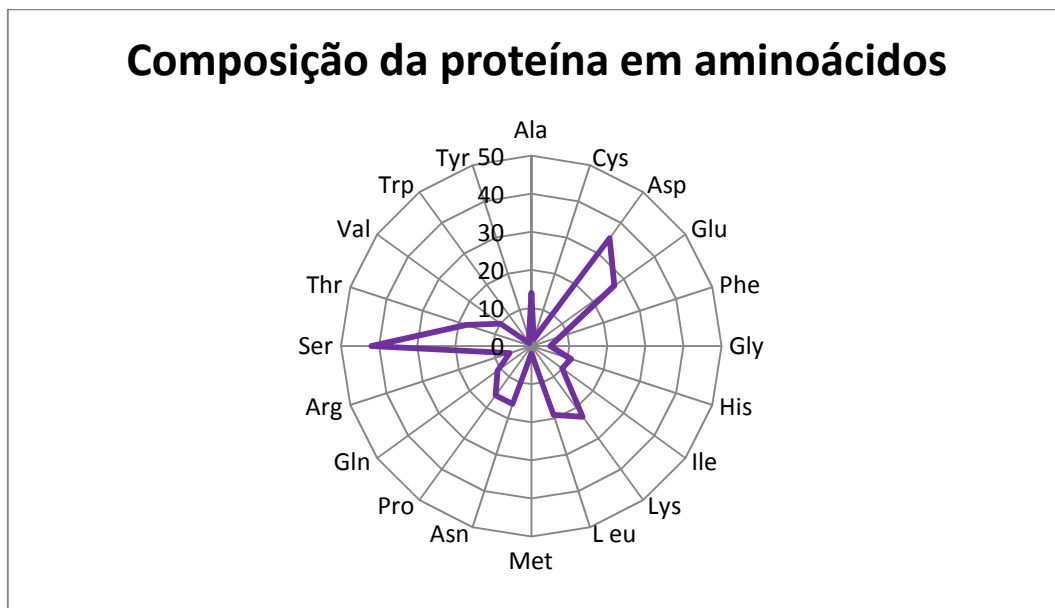
O prognóstico de Lupas é apresentado no gráfico 107.



**Gráfico 107:** Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na osteopontina.

Observa-se que existe alguma probabilidade de encontrar uma super-hélice- $\alpha$  na osteopontina (Gráfico 107). O único pico encontra-se a partir dos resíduos 218.

A composição em aminoácidos da osteopontina pode ser apresentada num gráfico em radar (Gráfico 108).

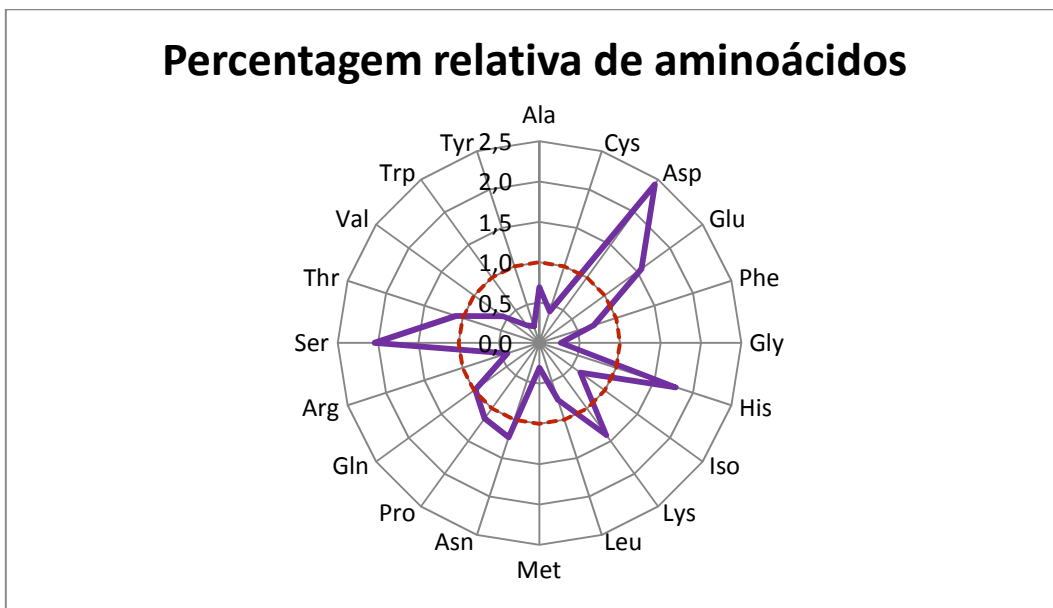


**Gráfico 108:** Representação gráfica da quantidade real de cada aminoácido na osteopontina (278 resíduos de aminoácido).

A osteopontina é constituída pelos vinte principais aminoácidos (Gráfico 108). Sendo assim, a osteopontina é constituída maioritariamente por (ordem decrescente) serina e aspartato. Tem uma quantidade média de glutamato, lisina, leucina, treonina, asparagina, prolina, alanina, histidina, glutamina, valina e isoleucina. Possui um valor mínimo de fenilalanina, arginina e glicina. A osteopontina possui igualmente alguns aminoácidos residuais tais como cisteína, tirosina, metionina e triptofano. Visto que a osteopontina é relativamente rica em glutamato, existe uma forte suspeita de que ela tenha pelo menos uma hélice- $\alpha$  na sua estrutura. Os valores elevados de serina e aspartato podem influenciar a formação de hélices- $\alpha$  caso existam. O valor mínimo de valina isoleucina e treonina dão poucas hipóteses a possível presença de folhas- $\beta$ . A presença de voltas é igualmente presumida devido à existência de valores mínimos de asparagina e prolina (Berg et al., 2002).

É de notar que esta proteína possui uma quantidade de aspartato inferior (35 resíduos) relativamente ao indicado por Rodrigues et al., (2007) (48 resíduos).

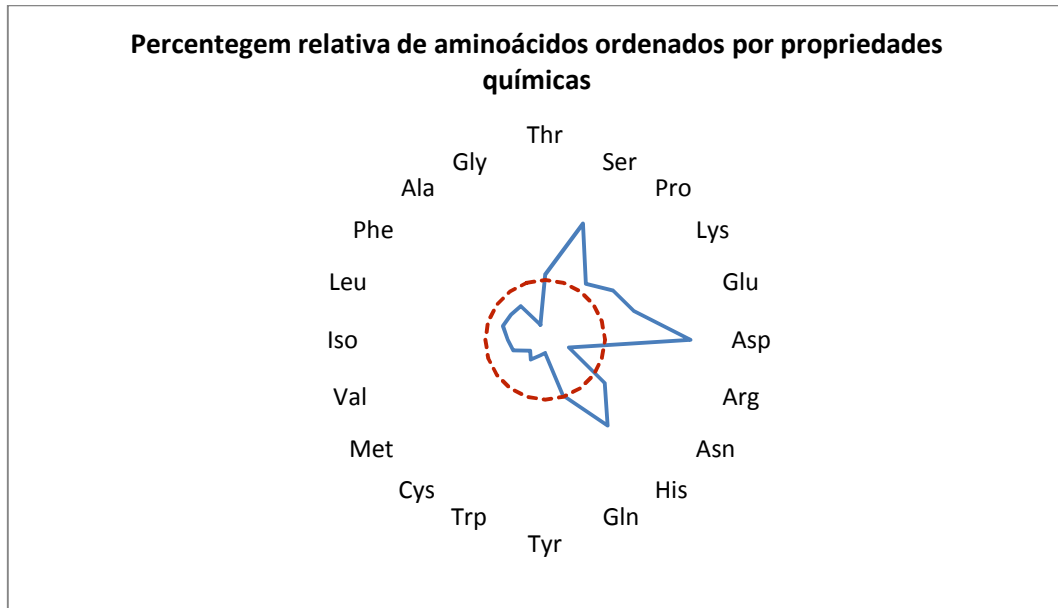
O gráfico 109 mostra a relatividade entre a percentagem de aminoácidos da osteopontina e da GenPep#97.



**Gráfico 109:** Representação gráfica da percentagem de cada aminoácido na osteopontina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na osteopontina.

Observa-se que a osteopontina possui oito aminoácidos com uma percentagem relativa acima da média, calculada a partir das percentagens relativas dos 20 aminoácidos presentes na osteopontina e representada a tracejado vermelho (Gráfico 109). Desses oito, o mais evidente é o aspartato, seguida da serina, histidina, glutamato, lisina, asparagina, prolina e treonina. Os restantes aminoácidos têm um valor abaixo da média.

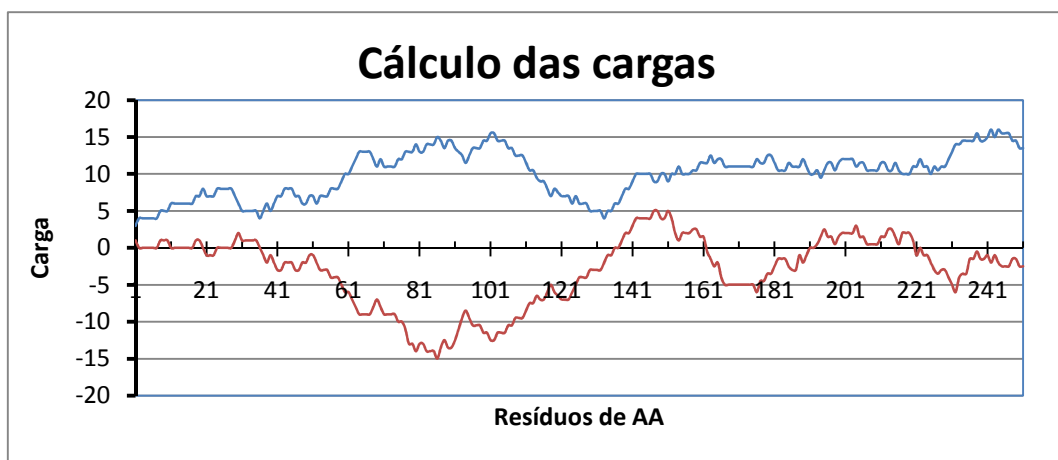
A distribuição dos aminoácidos segundo as suas propriedades químicas, revela algumas características da osteopontina, e está representada no gráfico 110.



**Gráfico 110:** Representação gráfica da percentagem relativa dos aminoácidos da osteopontina, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão situados à direita e os mais hidrofóbicos estão situados à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na osteopontina.

Observa-se que todos os aminoácidos com valores superiores à média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína) são aminoácidos carregados (Gráfico 110). Isto indica claramente que a osteopontina é extremamente hidrossolúvel (Wang e Denhardt, 2008). Os aminoácidos que contribuem para esse efeito são o aspartato, serina, histidina, glutamato, lisina, asparagina e prolina. A treonina também tem valores acima da média mas o seu contributo no comportamento global da osteopontina é pouco relevante.

As cargas da osteopontina podem ser observadas através do gráfico 111.



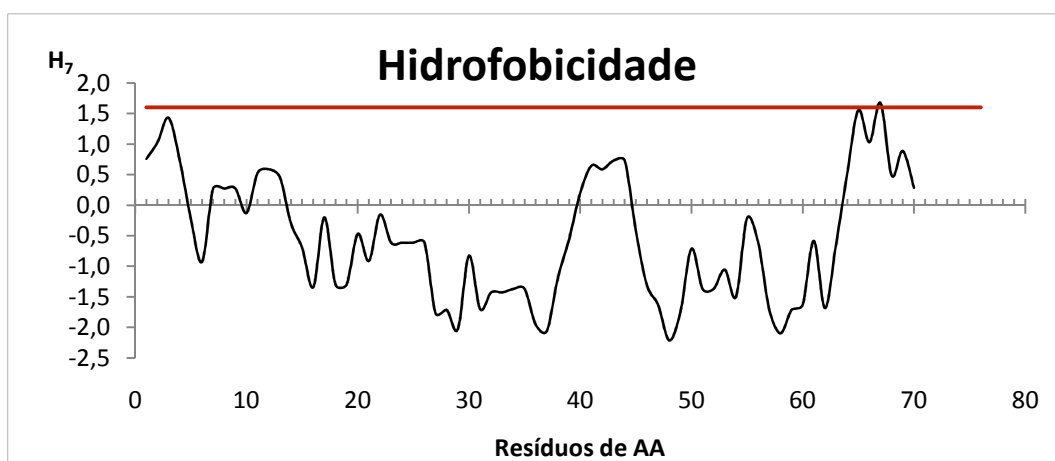
**Gráfico 111:** Representação gráfica das cargas da osteopontina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

Observa-se que a osteopontina possui poucas cargas reais positivas. No entanto, é de salientar um pico em que os valores da carga são mais elevados. Esse pico encontra-se no intervalo 138-161 (Gráfico 111). Existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores. Um deles é o intervalo 193-220.

Observa-se igualmente que as cargas absolutas da osteopontina mostram que ela tem um comportamento polar positivo ao longo de toda a cadeia de aminoácidos. Destacam-se no entanto dois picos onde os valores são mais importantes, definidos pelos intervalos 60-120 e 230-251. Isto significa que se a osteopontina for colocada numa solução com uma corrente eléctrica contínua, ela deslocar-se-ia para o cátodo.

#### 4.2.4.3. Ubiquitina

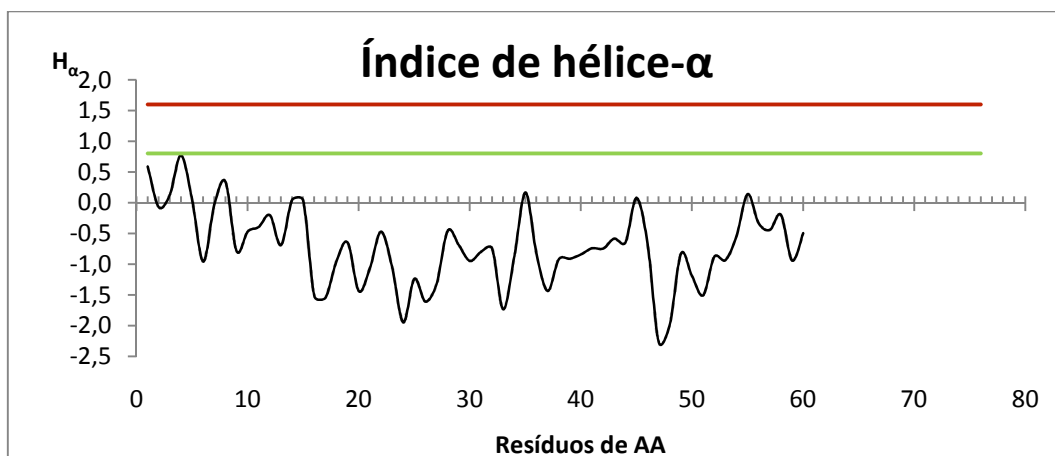
A ubiquitina é uma pequena proteína que também desempenha muitas funções no organismo, e os seus níveis de hidrofobicidade estão representados no gráfico 112, segundo a escala de Kyte & Doolittle.



**Gráfico 112:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a ubiquitina (76 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho).

Observa-se que a ubiquitina (Gráfico 112) mostra ter um comportamento bastante hidrofílico, isto apesar de, no final da sua cadeia existir um pico que ultrapassa o valor de hidrofobicidade de 1,6. No entanto, este pico não é considerado, visto que não possui um mínimo de 19 aminoácidos (Han e Tashjian, 2003).

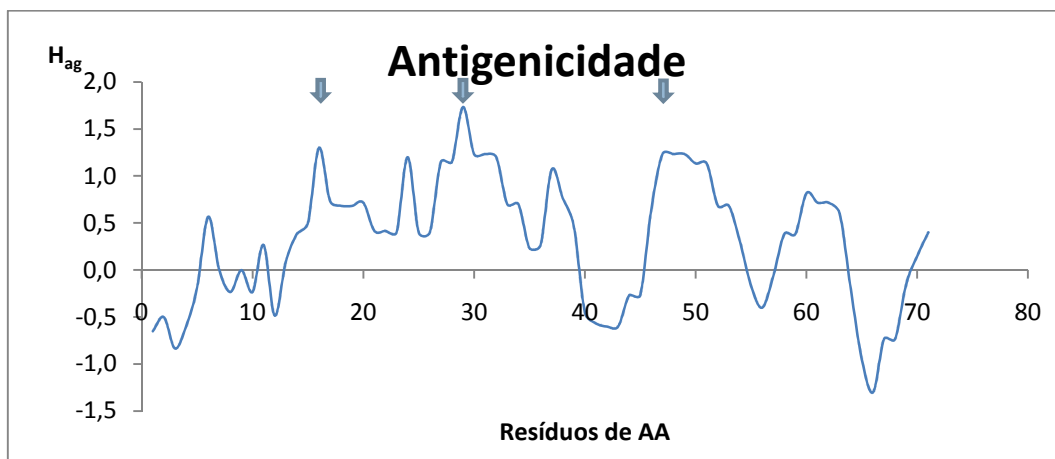
No gráfico 113 observa-se o resultado dos cálculos para o índice de hélice- $\alpha$  da ubiquitina.



**Gráfico 113:** Representação gráfica do cálculo do índice de hélice- $\alpha$  para a ubiquitina (76 resíduos de aminoácido), usando a escala de Kyte & Doolittle (preto). Linha de referência de 1,6 (vermelho) e linha de referência de 0,8 (verde claro).

No índice de hélices- $\alpha$ , continua a observar-se a tendência hidrofílica da ubiquitina (Gráfico 113), visto que todos os valores são inferiores a 0,8. Ao observar-se o gráfico, conclui-se que existe a possibilidade de a ubiquitina não possuir hélices- $\alpha$  anfílicas na sua estrutura.

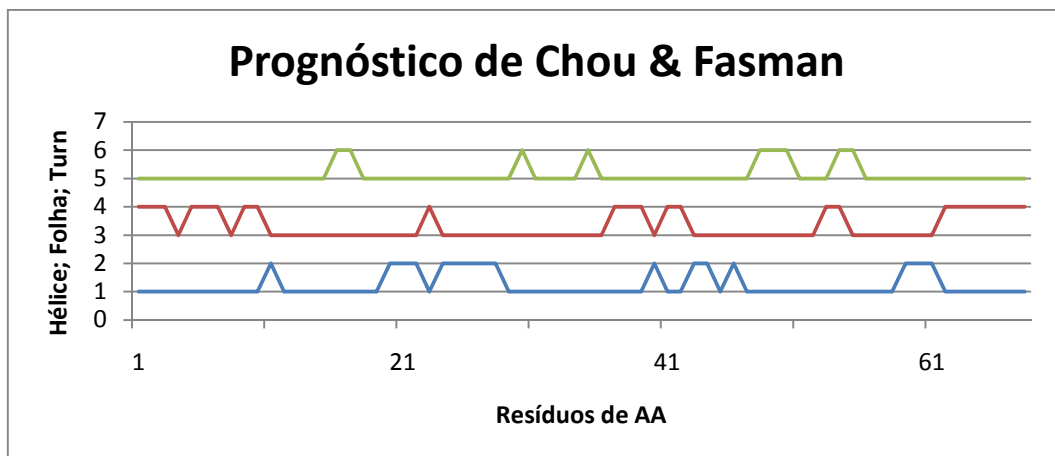
A antigenicidade da ubiquitina pode ser representada graficamente do seguinte modo.



**Gráfico 114:** Representação gráfica do cálculo da antigenicidade  $H_{Ag}$  para a ubiquitina, usando a escala de Hopp & Wood (azul). Zonas antigénicas possíveis (setas azuis).

Pela observação do gráfico 114, constata-se que a ubiquitina apresenta apenas três picos antigénicos relevantes, nos resíduos 15, 28 e 47.

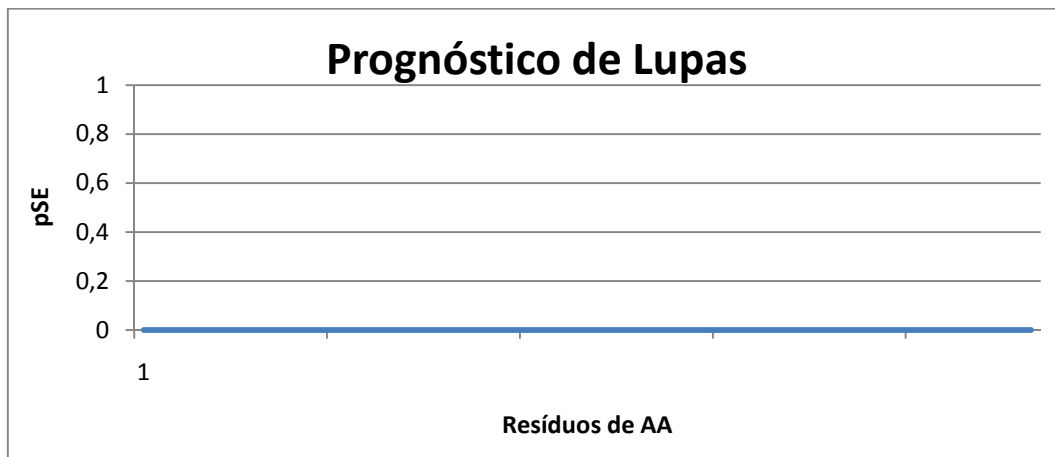
Através do método de Chou & Fasman podem detectar-se as seguintes estruturas.



**Gráfico 115:** Representação gráfica do prognóstico de Chou & Fasman para os três elementos principais da estrutura secundária: hélice- $\alpha$  (azul); folha- $\beta$  (vermelho) e volta (verde), na ubiquitina. Cada pico corresponde à presença de um possível elemento.

Observa-se que existe alguma probabilidade de encontrar quatro hélices- $\alpha$ , sete folhas- $\beta$  e três voltas ao longo da cadeia polipeptídica da ubiquitina (Gráfico 115). Contudo, conclui-se que a estimativa do número de folhas- $\beta$ , recorrendo ao método de Chou & Fasman, pode não ser o mais adequado, visto que para a BSA foram detectadas várias folhas- $\beta$ , quando esta não as possui.

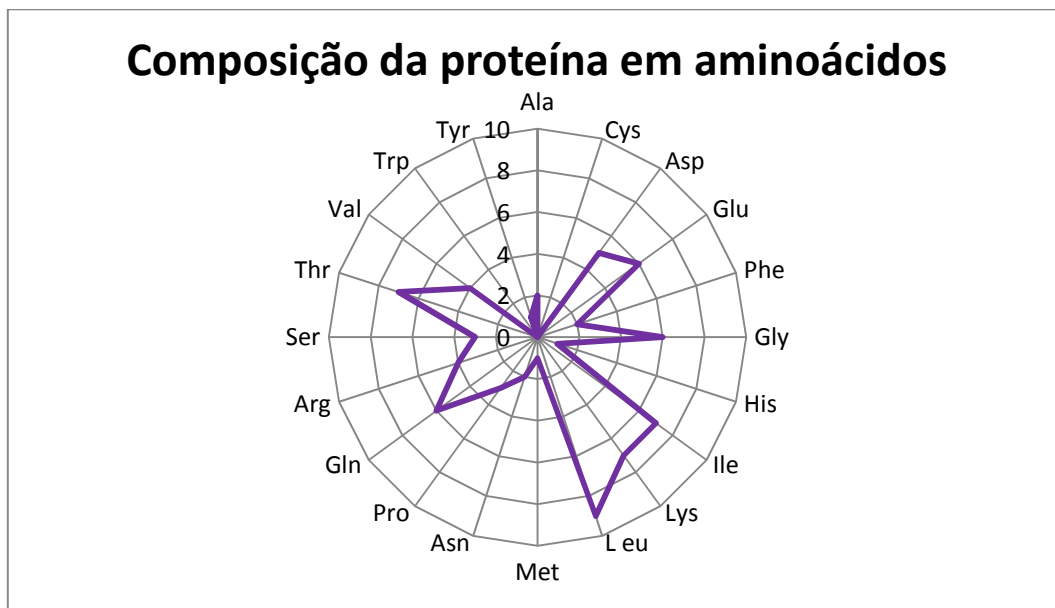
O prognóstico de Lupas é apresentado no gráfico 116.



**Gráfico 116:** Representação gráfica do prognóstico de Lupas para as super-hélices- $\alpha$  na ubiquitina.

Observa-se que a probabilidade de haver super-hélices- $\alpha$  na ubiquitina é nula (Gráfico 116).

A composição em aminoácidos da ubiquitina pode ser apresentada num gráfico em radar (Gráfico 117).

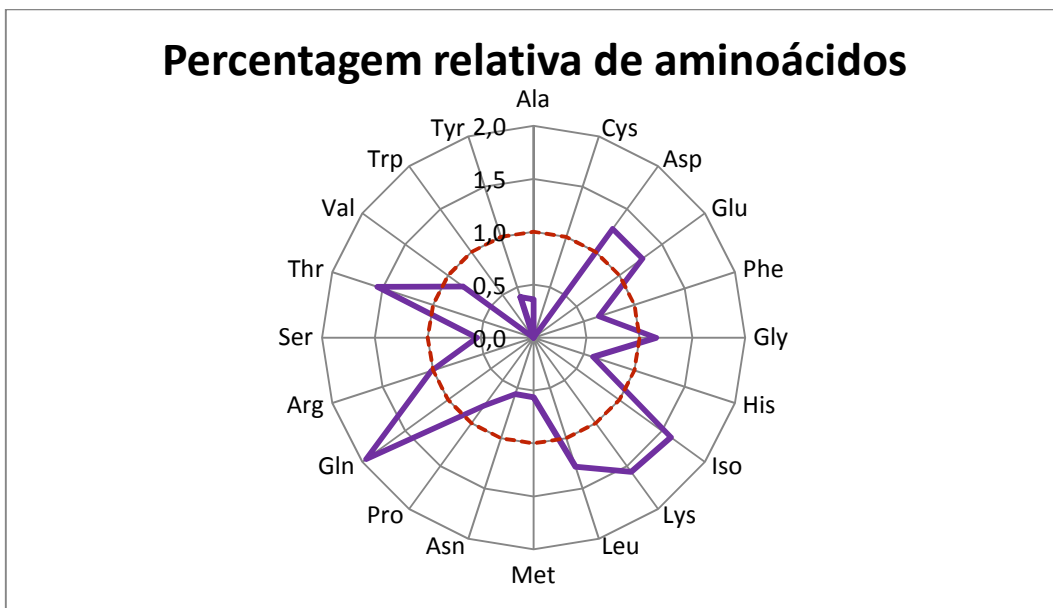


**Gráfico 117:** Representação gráfica da quantidade real de cada aminoácido na ubiquitina (76 resíduos de aminoácido).

A ubiquitina é constituída pelos principais aminoácidos excepto a cisteína e o triptofano (Gráfico 117). Sendo assim, a ubiquitina é constituída maioritariamente por leucina. Tem uma quantidade média de isoleucina, lisina, treonina, glutamato, glicina, glutamina, aspartato, arginina e valina. Possui um valor mínimo de prolina, serina, alanina, fenilalanina, asparagina, histidina, metionina e tirosina.

Tendo em conta que a ubiquitina é rica em leucina, existe uma forte probabilidade para que ela tenha pelo menos uma hélice- $\alpha$  na sua estrutura. Os valores médios de isoleucina, treonina e valina revelam também a possível presença de folhas- $\beta$ . A presença de voltas é igualmente possível devido à existência de valores médios de glicina (Berg et al., 2002).

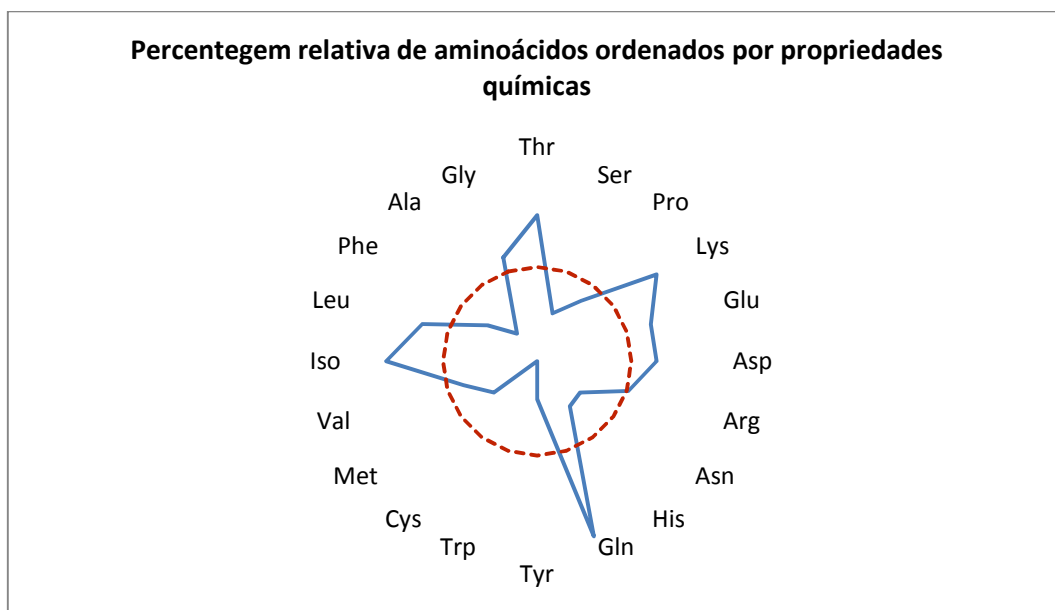
O gráfico 118 mostra a relatividade entre a percentagem de aminoácidos da ubiquitina e da GenPep#97.



**Gráfico 118:** Representação gráfica da percentagem de cada aminoácido na ubiquitina, relativamente à GenPep#97. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

Observa-se que a ubiquitina possui nove aminoácidos com um valor acima da média, sendo a média representada a tracejado vermelho e calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína (Gráfico 118). Desses nove, o mais evidente é a glutamina, seguida da isoleucina, lisina, treonina, leucina, aspartato, glutamato, glicina e arginina. Os restantes aminoácidos têm um valor abaixo da média.

A distribuição dos aminoácidos segundo as suas propriedades químicas, revela algumas características da ubiquitina.



**Gráfico 119:** Representação gráfica da percentagem relativa dos aminoácidos da ubiquitina, mediante as suas propriedades químicas. Os aminoácidos mais carregados estão à direita e os mais hidrofóbicos estão à esquerda. O círculo a tracejado vermelho representa uma média calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína.

Observa-se que, a maioria dos aminoácidos com valores acima da média (calculada a partir das percentagens relativas dos 20 aminoácidos presentes na proteína) são hidrofílicos (Gráfico 119). À luz destes dados a ubiquitina mostra ser uma proteína bastante solúvel em solventes polares. Em termos comparativos, ela possui cinco aminoácidos carregados contra três hidrofóbicos.

Os aminoácidos que contribuem para a sua solubilidade são a glutamina, lisina, aspartato, glutamato e arginina. Os aminoácidos que contribuem para a sua hidrofobicidade são a isoleucina, leucina e glicina. A treonina também tem valores acima da média mas não contribui significativamente no comportamento global da ubiquitina.

As cargas da ubiquitina podem ser observadas através do gráfico 120.

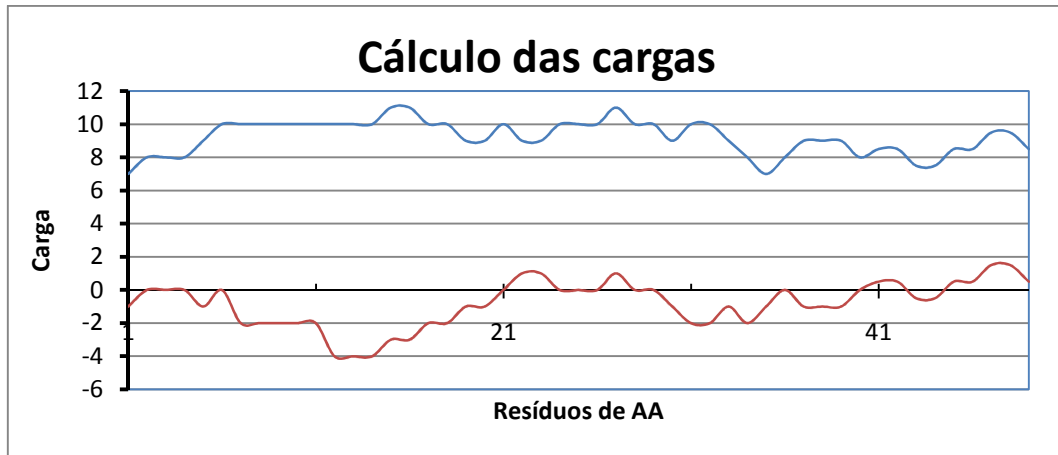


Gráfico 120: Representação gráfica das cargas da ubiquitina. Tanto as cargas reais (castanho) como as cargas absolutas (azul), são calculadas num intervalo de 28 aminoácidos.

Observa-se que a ubiquitina possui poucas cargas reais positivas mas também é preciso ter em conta que a ubiquitina é constituída por apenas 76 resíduos de aminoácidos (Garzón *et al.*, 2007; Chew *et al.*, 2005). Destaca-se apenas o intervalo no final desta proteína (45-49), que no entanto representa quatro resíduos de aminoácidos (Gráfico 120). Existem outros intervalos onde as cargas reais também são positivas mas são mais curtos e de cargas inferiores.

De um modo totalmente oposto, observa-se que as cargas absolutas da ubiquitina, mostram que ela tem um comportamento polar positivo ao longo de toda a cadeia polipeptídica e com valores muito superiores às cargas reais. Como em todas as outras proteínas estudadas até aqui, isto significa que se a ubiquitina for colocada numa solução com uma corrente eléctrica contínua, ela deslocar-se-á para o cátodo.

## 5. CONCLUSÃO

---

Este estudo teve como objectivo verificar se a partir das propriedades físico-químicas dos aminoácidos da estrutura primária das proteínas, era possível prever comportamentos hidrofóbicos e alergénicos das diferentes proteínas estudadas. Para além disso, também se pretendeu prever os diversos elementos da estrutura secundária (hélices- $\alpha$ ; folhas- $\beta$  e voltas) utilizando diferentes métodos baseados em probabilidades.

O facto de ainda hoje ser bastante difícil afirmar com toda a certeza de que uma determinada proteína possui este ou aquele tipo de organização tridimensional, faz desta abordagem um passo importante na compreensão e elaboração de novos métodos de previsão estrutural.

Para concretizar os objectivos, o trabalho subdividiu-se em cinco abordagens diferentes: a utilização de escalas de hidrofobicidade para determinar o comportamento hidrofóbico ( $H_7$ ) das proteínas, a possível presença de hélices- $\alpha$  anfífilas ( $H_\alpha$ ) e as zonas antigénicas ( $H_{Ag}$ ); a utilização do método de Chou & Fasman para determinar os três tipos de estruturas (hélices- $\alpha$ ; folhas- $\beta$  e voltas); a utilização do método de Lupas *et al.* (Filby, 1998) para determinar a presença de super-hélices- $\alpha$ ; a determinação da composição real em aminoácidos de cada proteína, a percentagem de cada aminoácido em relação à BSA e a utilização dessa mesma percentagem para averiguar se a proteína é hidrofóbica ou carregada (Gráficos em radar); e por fim, a determinação das cargas reais e absolutas de cada proteína.

Testaram-se as causas de hidrofobicidade de Kyte & Doolittle; Hopp & Woods; Sweet; Welling; Eisenberg; Black com a BSA. Optou-se pela utilização da escala de Kyte & Doolittle por ser a que permitia uma maior discriminação nos resultados.

No que diz respeito aos métodos das escalas hidrofóbicas ( $H_7$ ;  $H_\alpha$  e  $H_{Ag}$ ), nem todas as proteínas tiveram um comportamento ao qual se esperava.

No caso do  $H_7$ , as proteínas do leite mostraram ter um comportamento mais hidrossolúvel ao longo das suas cadeias do que as proteínas da clara do ovo o que normalmente deveria ser ao contrário. Apesar do critério de um mínimo de 19 aminoácidos num pico, para que este seja considerado como sendo hidrofóbico, nem sempre foi evidente distinguir os verdadeiros picos hidrofóbicos dos restantes.

No Índice de hélices- $\alpha$  ( $H_\alpha$ ), observou-se vários picos hidrofóbicos e anfífilos nas diversas proteínas. Em doze proteínas, oito apresentaram um pico inicial hidrofóbico. Exceptuando os simples valores de referencia de hidrofobicidade de 1,6 e de hidrofilicidade de 0,8 a ausência de outro tipo de critérios tornou difícil a distinção de picos anfífilos que provassem a presença de hélices- $\alpha$  do mesmo género.

Relativamente à antigenicidade das proteínas em estudo ( $H_{Ag}$ ), a total ausência de critérios teóricos que definem a escolha dos picos antigénicos, dificultou a análise deste tipo de gráfico. No entanto, em algumas das proteínas, como por exemplo: a  $\alpha$ -caseína-s1,  $\alpha$ -caseína-s2,  $\beta$ -caseína e  $\beta$ -lactoglobulina, os picos encontrados ao nível do gráfico corroboram com as zonas antigénicas descritas na literatura para essas proteínas. Contudo os picos obtidos permitem conhecer os aminoácidos presentes responsáveis podendo ajudar a “desenhar” anticorpos específicos.

No que diz respeito ao método de Chou & Fasman (Filby, 1998), todas as proteínas manifestaram a presença de inúmeras hélices- $\alpha$ , folhas- $\beta$  e voltas nas suas estruturas secundárias. No entanto, este método mostrou ser inadequado na previsão de folhas- $\beta$  pelo menos em alguns casos.. A comprovar este facto, tem-se o caso da proteína de referência (BSA), que não possui folhas- $\beta$  e no entanto o método de Chou & Fasman dá um prognóstico de várias folhas- $\beta$ . Em relação às hélices- $\alpha$  e às voltas, a previsão destes elementos aproximou-se mais dos dados teóricos mas nunca com valores exactos.

No método de Lupas *et al.* (Filby, 1998), apenas oito das doze proteínas apresentam super-hélices- $\alpha$ . Das quatro proteínas que não possuem super-hélices- $\alpha$ , duas são da clara do ovo (lisozima C e ovomucóide) e as outras duas são do organismo ( $\beta$ -2-microglobulina e ubiquitina).

A comparação da percentagem de cada aminoácido em cada proteína, relativamente à percentagem dos mesmos aminoácidos da BSA, revela também ela um gráfico em radar característico a cada proteína. Para além disso, é possível visualizar qual é a percentagem de aminoácidos que está acima da média, média essa que foi calculada a partir do estudo de 10000 proteínas.

Relativamente aos gráficos das cargas, pode dizer-se que revelou zonas polares em todas as proteínas, susceptíveis de anunciar zonas de maior interacção com o solvente. Para além disso, todas as proteínas em estudo revelaram ter cargas absolutas com valores relativamente elevados excepto na  $\beta$ -caseína onde os valores mínimos atingem o zero.

Apesar de os métodos utilizados mostrarem algumas incoerências de resultados com determinadas proteínas, eles revelaram ser instrumentos bastante úteis para uma primeira abordagem na previsão do comportamento e da presença de elementos da estrutura secundária.

Com base na integração dos resultados teóricos obtidos para cada proteína, conjuntamente com outros resultados teóricos e práticos obtidos por outros métodos, será possível no futuro conseguir conhecer a estrutura terciária das proteínas, cujo interesse e importância no âmbito da ciência da vida, é crescente.

## 6. FUTURAS INVESTIGAÇÕES

---

Todos os trabalhos científicos comportam os seus erros e dúvidas que ficaram por esclarecer. Determinados processos poderiam ter sido feitos de uma outra maneira ou de um modo mais aprofundado. Este trabalho não é excepção. A razão para o sucedido deve-se ao facto de que este tipo de trabalho faz parte da génese de um ramo da ciência que se esforça a representar tridimensionalmente as proteínas. Sendo assim, alguns processos, métodos de cálculos e referências deveriam ser reconsiderados.

Por exemplo, poderiam usar-se outras escalas de hidrofobicidade para além daquelas que foram utilizadas. Isto apesar da escala de Kyte & Doolittle ser habitualmente utilizada neste tipo de análise.

Outra sugestão, seria utilizar o método de Chou & Fasman com várias proteínas bem conhecidas estruturalmente e verificar qual o grau de precisão desse método na detecção dos três elementos diferentes: hélice- $\alpha$ , folha- $\beta$  e volta.

A mesma sugestão poderia ser aplicada ao método de Lupas.

## 7. BIBLIOGRAFIA

---

- Alleoni A.C.C. “Albumen Protein and Functional Properties of Gelation and Foaming”, *Scientia Agricola*, Vol.63(3):291-298 (2006).
- Bayard C., Lottspeich F. “Bioanalytical characterization of proteins”, *Journal of chromatography B*, Vol. 756: 113-122 (2001).
- Belitz H.-D, Grosch W. e Schieberle P., *Food Chemistry*, 4th revised and extended edition, Springer, Berlin, Germany (2009).
- Berg J.M., Tymoczko J.L. e Stryer L., *Biochemistry*, fifth edition, W.H. Freeman, New York, USA (2002).
- Bernard H., Meisel H., Creminon C. e Wal J.M. “Post translational phosphorylation affects the IgE binding capacity of caseins”, *FEBS Letters*, Vol. 467(2-3): 239-244 (2000).
- Black S.D. e Mould D.R. “Amino acid scale: Hydrophobicity of physiological L-alpha amino acids”, *Analytical Biochemistry*, Vol. 193: 72-82 (1991).
- Bujnicki J.M., *Practical Bioinformatics, Volume 15 de Nucleic Acids and Molecular Biology*, Springer, Warsaw, Poland (2008).
- Caessens P.W.J.R., De Jongh H.H.J., Norde W. e Gruppen H. “The adsorption-induced secondary structure of B-casein and of distinct parts of its sequence in relation to foam and emulsion properties”, *Biochimica et Biophysica Acta*, Vol. 1430: 73-83 (1999).
- Carter D.C. e Ho J.X. “Structure of serum albumin”, em: *Advances in protein chemistry Volume 45 de Lipoproteins, apolipoproteins, and lipases*, ed Verne N. Schumaker, Academic Press, California, USA (1994).
- Chang R. “Chapter 26 - Polímeros orgânicos, sintéticos e naturais”, em: *Química*, Fifth Edition, McGraw-Hill: 1098 (1995).
- Chew G.H., Galloway L.C., McIntyre N.R., Schroder L.A., Richards K.M., Miller S.A., Wright D.W. e Merkler D.J. “Ubiquitin and ubiquitin-derived peptides as substrates for peptidylglycine  $\alpha$ -amidating monooxygenase”, *FEBS Letters*, Vol. 579: 4678-4684 (2005).
- Chou P.Y. e Fasman G.D., “Prediction of the secondary structure of proteins from their amino acid sequence”, *Adv Enzymol Relat Areas Mol Biol*, Vol. 47:45-148 (1978).
- Coutouly G., Klein E., Barbieri E. e Kriat M., *Travaux diriges de biochimie, biologie moleculaire et bio-informatique*, 3<sup>e</sup> Édition, DOIN (2006).
- Ehn B.M., Ekstrand B., Begtsson U. e Ahlstedt S. “Modification of IgE binding during heat processing of the cow’s milk allergen beta-lactoglobulin”, *Journal of Agricultural and Food Chemistry*, Vol. 52(5): 1398-1403 (2004).
- Eisenberg D., Schwarz E., Komarony M. e Wall R. “Amino acid scale: Normalized consensus hydrophobicity scale”, *Journal of Molecular Biology*, Vol. 179: 125-142 (1984).
- Ent F. e Löwe J. “Crystal structure of the ubiquitin-like protein YukD from *Bacillus subtilis*”, *FEBS Letters*, Vol. 579: 3837-3841 (2005).

- Feinstein G. "Purification of trypsin by affinity chromatography on ovomucoid-sepharose resin", *FEBS Letters*, Vol. 7(4): 353-355 (1970).
- Filby G., *Spreadsheets in Science and Engineering*, Springer, New York, USA (1998).
- Garret R.H. e Grisham C.M., *Biochemistry*, Fourth edition, Cengage Learning, Boston, USA (2008).
- Garzón M., Eifler K., Faust A., Scheel H., Hofmann K., Konez C., Yephremob A. e Bachmair A. "PRT6/At5g02310 encodes an *Arabidopsis* ubiquitin ligase of the N-end rule pathway with arginine specificity and is not CER3 locus", *FEBS Letters*, Vol. 581: 3189-3196 (2007).
- Goldsby R.A., Kindt T.J. e B.A. Osbourne, *Kuby Immunology*, Fourth edition, W.H. Freeman & Company, New York, USA (2000).
- Hamdani S., Joly D., Carpentier R. e Tajmir-Riahi H.A. "The effect of methylamine on the solution structures of human and bovine serum albumins", *Journal of molecular structure*, Vol. 936: 80-86 (2009).
- Han B. e Tashjian A.H. "Chapter 6 - User-Friendly and Versatile Software for Analysis of Protein Hydrophobicity", em: *Biocomputing: Computer Tools for Biologists*, ed Stuart M. Brown, Eaton Publishing, New York, USA (2003).
- He X.M. e Carter D.C. "Atomic structure and chemistry of human serum albumin", *Nature*, Vol. 358: 209-214 (1992).
- Higashikawa F., Eboshida A. e Yokosaki Y. "Enhanced biological activity of polymeric osteopontin", *FEBS Letters*, Vol. 581: 2697-2701 (2007).
- Hoffmann-Sommergruber K. e Mills E.N. "Food allergen protein families and their structural characteristics and application in component-resolved diagnosis: new data from the EuroPrevall project", *Analytical and Bioanalytical Chemistry*, Vol. 395(1): 25-35 (2009).
- Holen E. e Elsayed S. "Characterization of four major allergens of hen-egg white by IEF/SDS-PAGE combined with electrophoretic transfer and IgE-immunoautoradiography", *International Archives of Allergy and Immunology*, Vol. 91: 136-141 (1990).
- Hong Y.-H., Ahn H.-C., Lim J., Kim H.-M., Ji H.-Y., Lee S., Kim J.-H., Park E.Y., Song H.K. e Lee B.-J. "Identification of a novel ubiquitin binding site of STAM1 VHS domain by NMR spectroscopy", *FEBS Letters*, Vol. 583: 287-292 (2009).
- Hoop T.P. e Woods K.R. "Prediction of protein antigenic determinants from amino acid sequences", *Proceedings of the National Academy of Sciences*, Vol. 78: 3824-3828 (1981).
- Hopp T.P. e Woods K.R. "Amino acid scale: Hydrophilicity", *Proceedings of the National Academy of Sciences*, Vol. 78: 3824-3828 (1981).
- Huopalahti R., López-Fandiño R., Anton M. e Schade R., *Bioactive egg Compounds*, Springer, Berlin, Germany (2007).
- Jedrychowski L. e Wichers H.J., *Chemical and biological properties of food allergens*, First Edition, CRC Press, Taylor & Francis Group, Boca Raton, FL, USA (2009).
- Kyte J. e Doolittle R.F. "Amino acid scale: Hydropathicity", *Journal of Molecular Biology*, Vol. 157: 105-132 (1982).

- Lide D.R., *CRC Handbook of Chemistry and Physics*, 90<sup>th</sup> edition, CRC Press, Taylor & Francis Group, Boca Raton, FL, USA (2009).
- Matsuda T., Watanabe K. e Sato Y. “Secondary structure of reduced ovomucoid and renaturation of reduced ovomucoid and its reduced fragments A(1-130) and B(131-186)”, *FEBS Letters*, Vol. 124(2): 185-188 (1981).
- Mine Y., *Egg bioscience and biotechnology*, John Wiley & Sons, Inc., New Jersey, USA (2008).
- Mine Y., Sasaki E. e Zhang J.W. “Reduction of antigenicity and allergenicity of genetically modified egg white allergene, ovomucoid third domain”, *Biochemical and Biophysical Research Communications*, Vol. 302: 133-137 (2003).
- Mosmann T.R., Cherwinski H., Bond M.W., Giedlin M.A., Coffman R.L., “Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins”, *J Immunol*, Vol. 136(7): 2348-2357 (1986).
- Mosmann T.R. e Coffman R.L., “TH1 and TH2 cells: Different patterns of lymphokine secretion lead to different functional properties”, *Annu Rev Immunol*, Vol. 7: 145-173 (1989).
- Mount W.D. “Chapter 10 - Protein Classification and Structure Prediction”, em: *Bioinformatics Sequence and Genome Analysis*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York: 409-444 (2004).
- Murray C., *Precis de Biochimie de Harper*, Traduction de la 23<sup>e</sup> édition américaine, De Boeck, Bruxelles (2002).
- Nelson D.L. e Cox M.M., *Lehninger Principles of Biochemistry*, Fourth Edition, W.H. Freeman, New York, USA (2004).
- Quesada J.M., Alonso J., Gonzalez J., Muñoz R., Jans I., Martiu A. e Boillon R. “Serum B-2 microglobulin is a marker of high bone remodelling in elderly women”, *Mechanisms of Ageing and Development*, Vol. 102: 293-298 (1998).
- Rao M.J.K. e Argos P. “Amino acid scale: Membrane buried helix parameter”, *Biochimica et Biophysica Acta*, Vol. 869: 197-214 (1986).
- Rodrigues L.R., Teixeira J.A., Schmitt F.L., Paulsson M. e Lindmark-Månsson H. “The Role of Osteopontin in Tumor Progression and Metastasis in Breast Cancer”, *Cancer Epidemiology, Biomarkers & Prevention*, Vol. 16(6): 1087-1097 (2007).
- Spuergin P., Walter M., Schiltz E., Deichmann K., Forest J. e Mueller H. “Allergenicity of alpha-casein from cow, sheep and goat”, *Allergy*, Vol. 52(3): 293-298 (1997).
- Stewart A.J., Blindauer C.A., Berezenko S., Sleep D. e Sadler P.J. “Interdomain zinc site on human albumin”, *PNAS*, Vol. 100(7): 3701-3706 (2003).
- Sweet R.M. e Eisenberg D. “Amino acid scale: Optimized matching hydrophobicity (OMH)”, *Journal of Molecular Biology*, Vol. 171: 479-488 (1983).
- Untersmayr E. e Jensen-Jarolim E., “Mechanisms of type I food allergy”, *Pharmacol Ther*, Vol. 112(3): 787-798 (2006).
- Voet D. e Voet J.G. *Biochemistry*, Third Edition, John Wiley & Sons, Inc., New Jersey, USA (2004).

Wang K.X. e Denhardt D.T. "Osteopontin: Role in immune regulation and stress responses", *Cytokine & Growth Factor Reviews*, Vol. 19: 333-345 (2008).

Welling G.W., Weijer W.J., Van der Zee R. e Welling-Wester S. "Amino acid scale: Antigenicity value X 10", *FEBS Letters*, Vol. 188: 215-218 (1985).

Whitford D. "Chapter 5 - The structure and Function of Membrane Proteins", em: *Proteins - Structure and Funtion*, John Wiley & Sons, Inc., New York: 11-12 (2005).

## 8. CIBERGRAFIA

---

---

<http://www.ncbi.nlm.nih.gov> (28/06/2010)

<http://foodallergens.ifr.ac.uk/>

<http://www.uniprot.org>

## 9. ANEXO

### Proteína de referência

Albumina Bovina

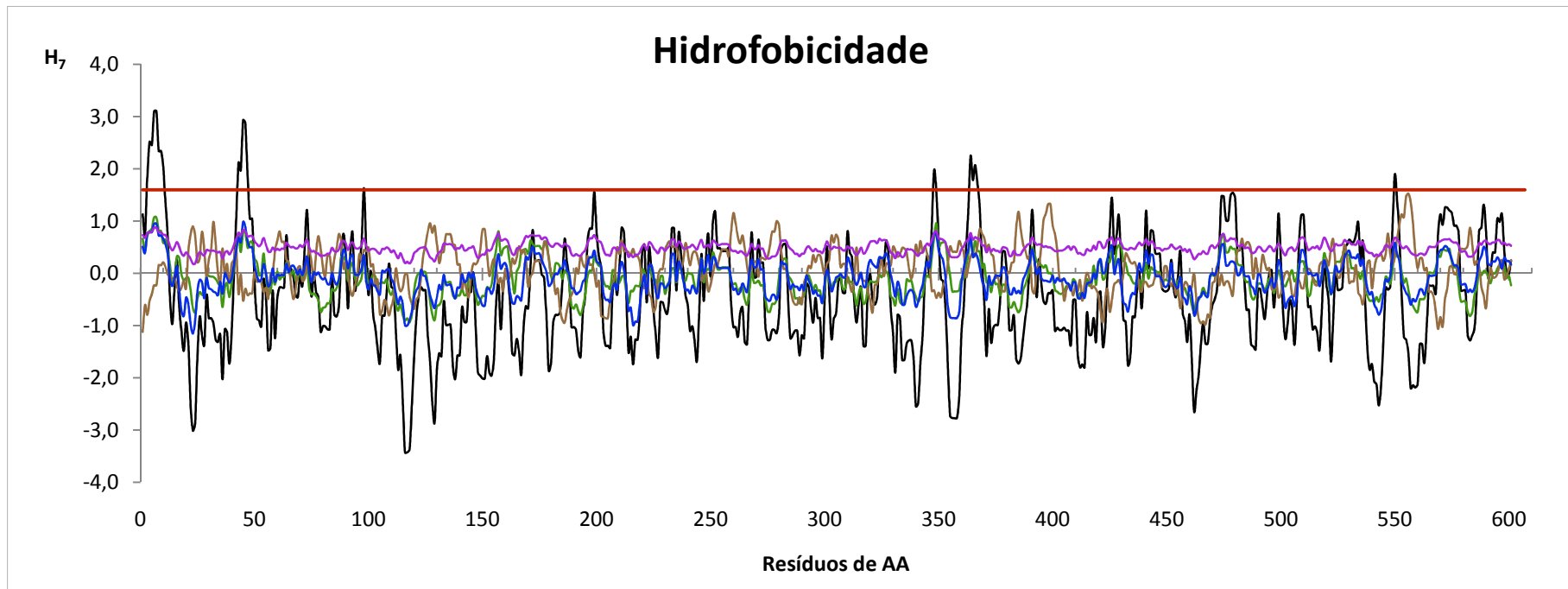
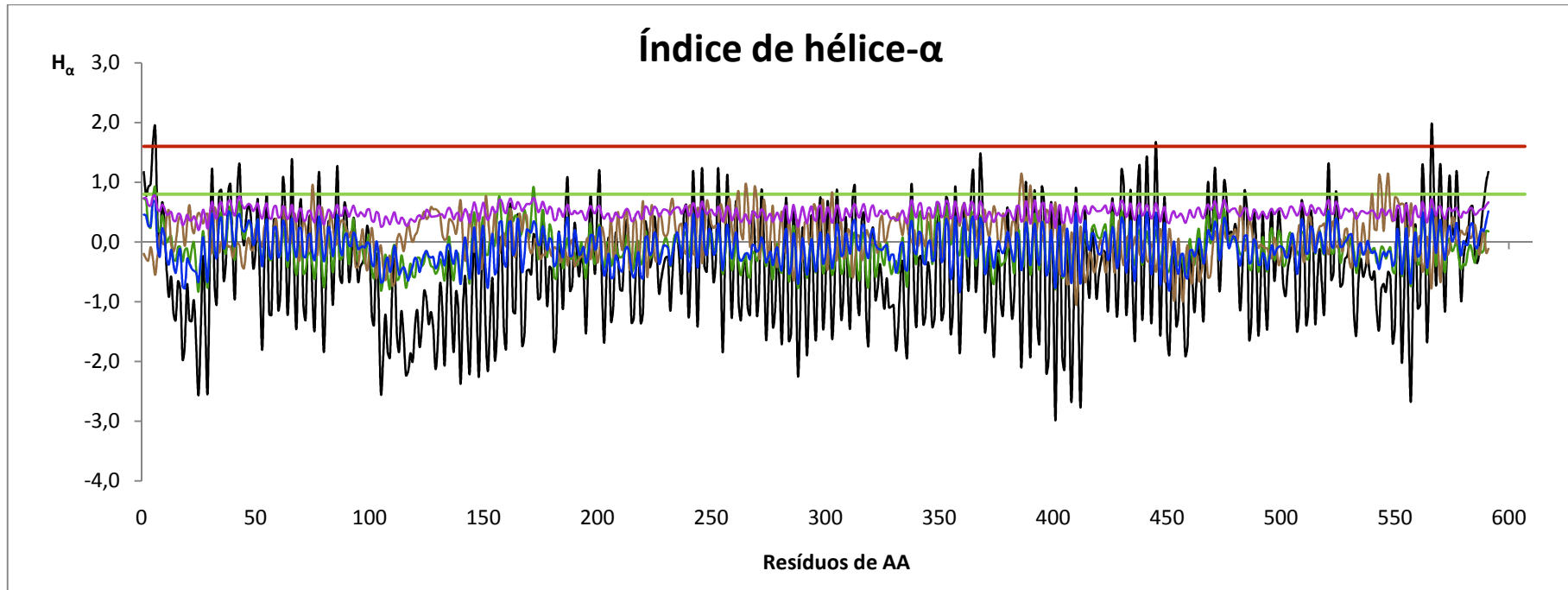


Gráfico 121: Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a BSA (607 resíduos de aminoácidos). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).



**Gráfico 122:** Representação gráfica do índice de hélice- $H_{\alpha}$  para a BSA (607 resíduos de aminoácidos). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

## Proteínas do ovo

### Lisozima C

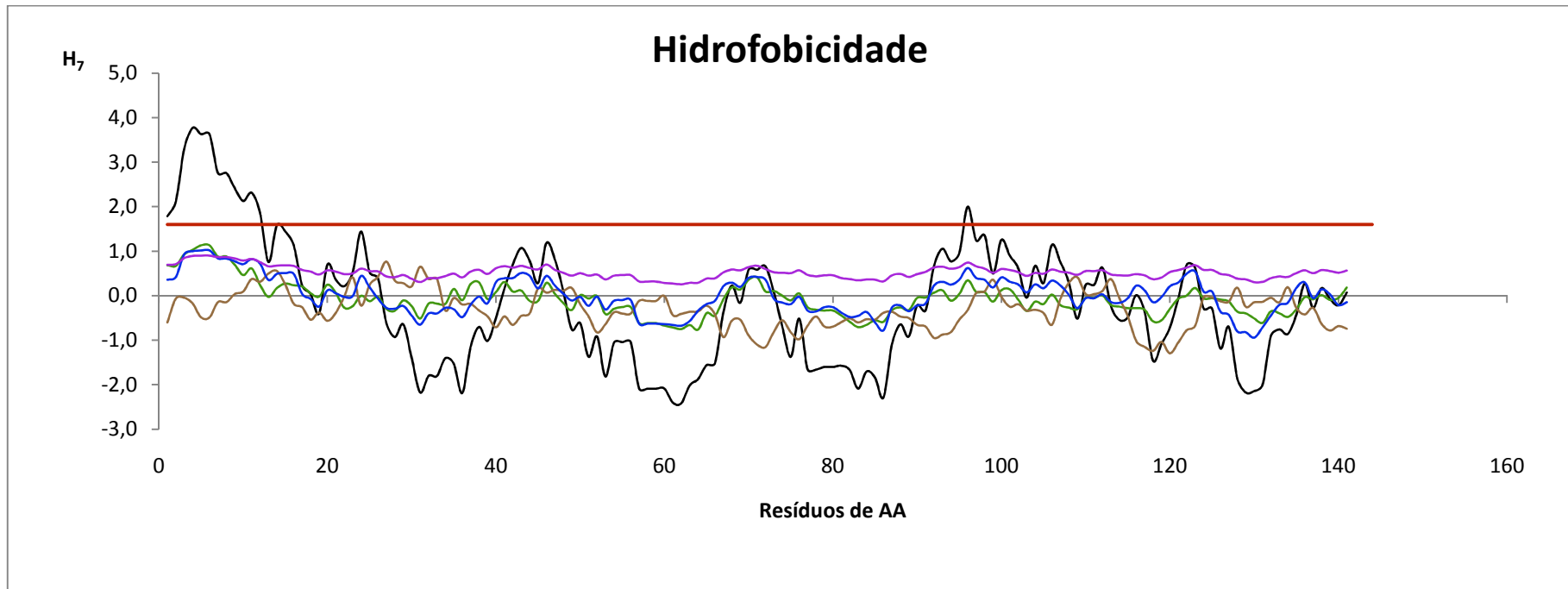
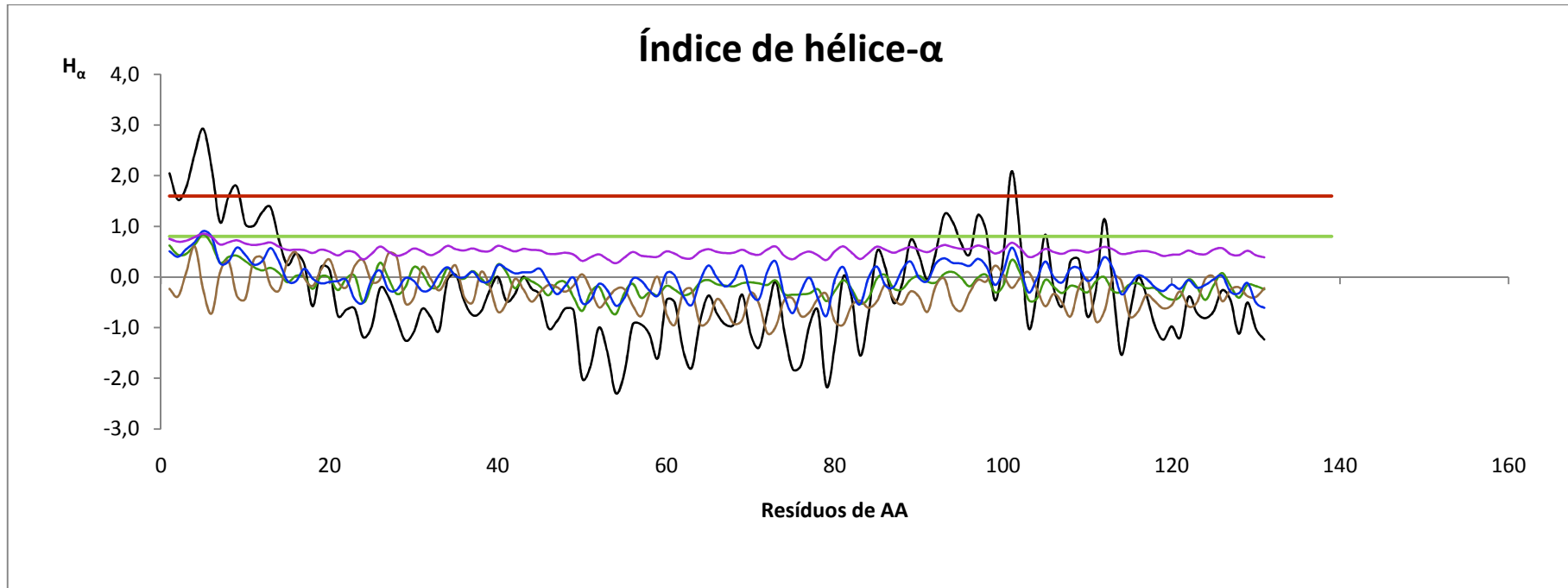
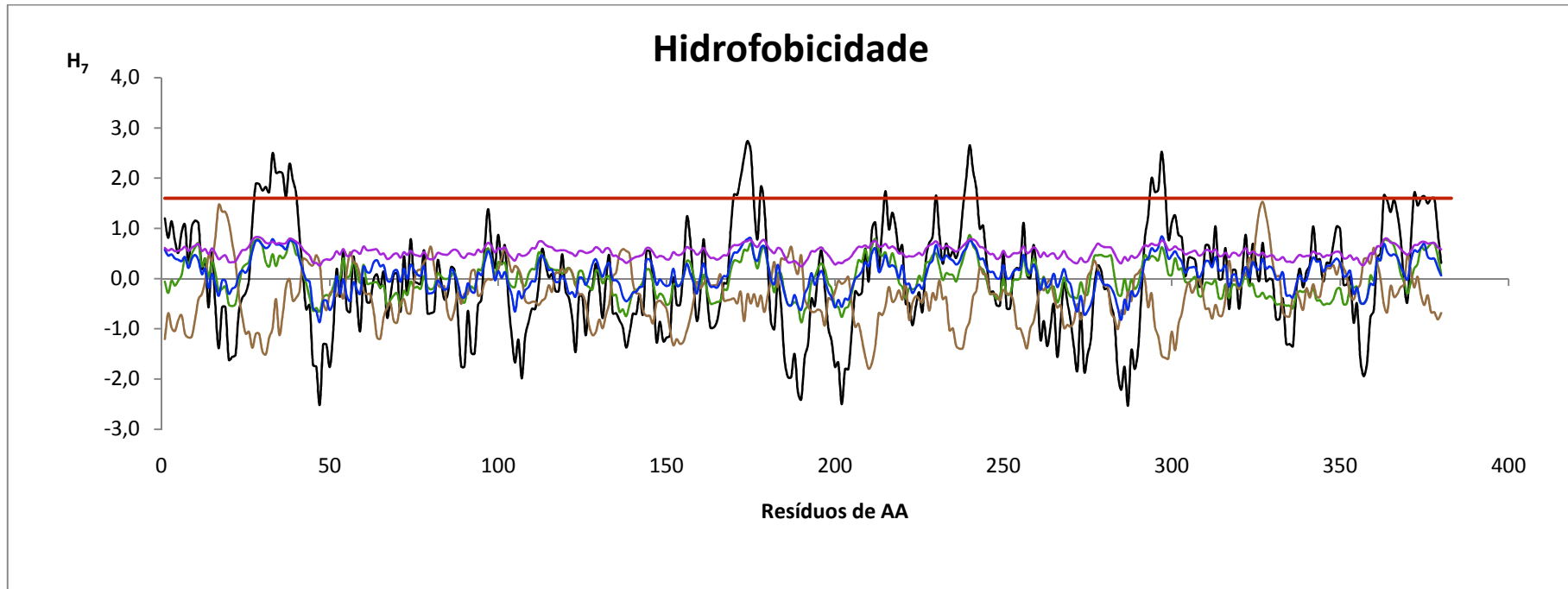


Gráfico 123: Representação gráfica do cálculo da hidrofobicidade H7 para a lisozima C (147 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

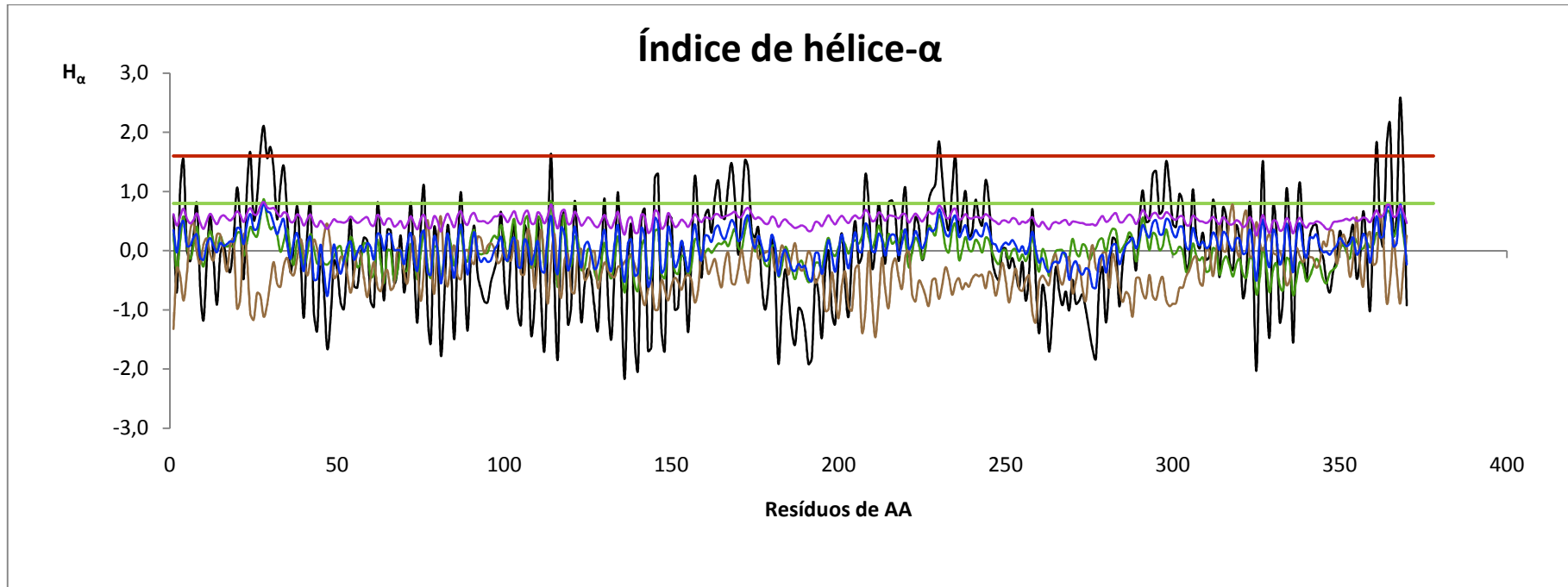


**Gráfico 124:** Representação gráfica do índice de hélice- $\alpha$   $H_\alpha$  para a lisozima C (147 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

### Ovalbumina

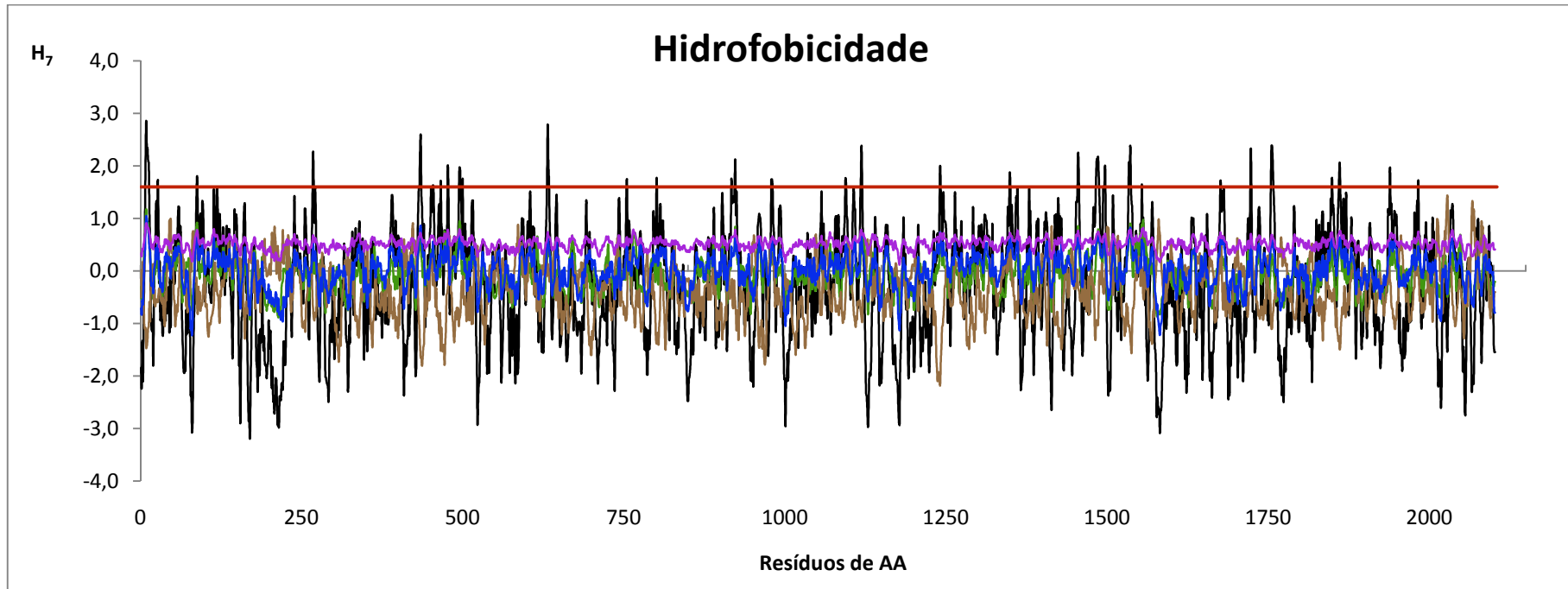


**Gráfico 125:** Representação gráfica do cálculo da hidrofobicidade H<sub>7</sub> para a ovalbumina (386 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

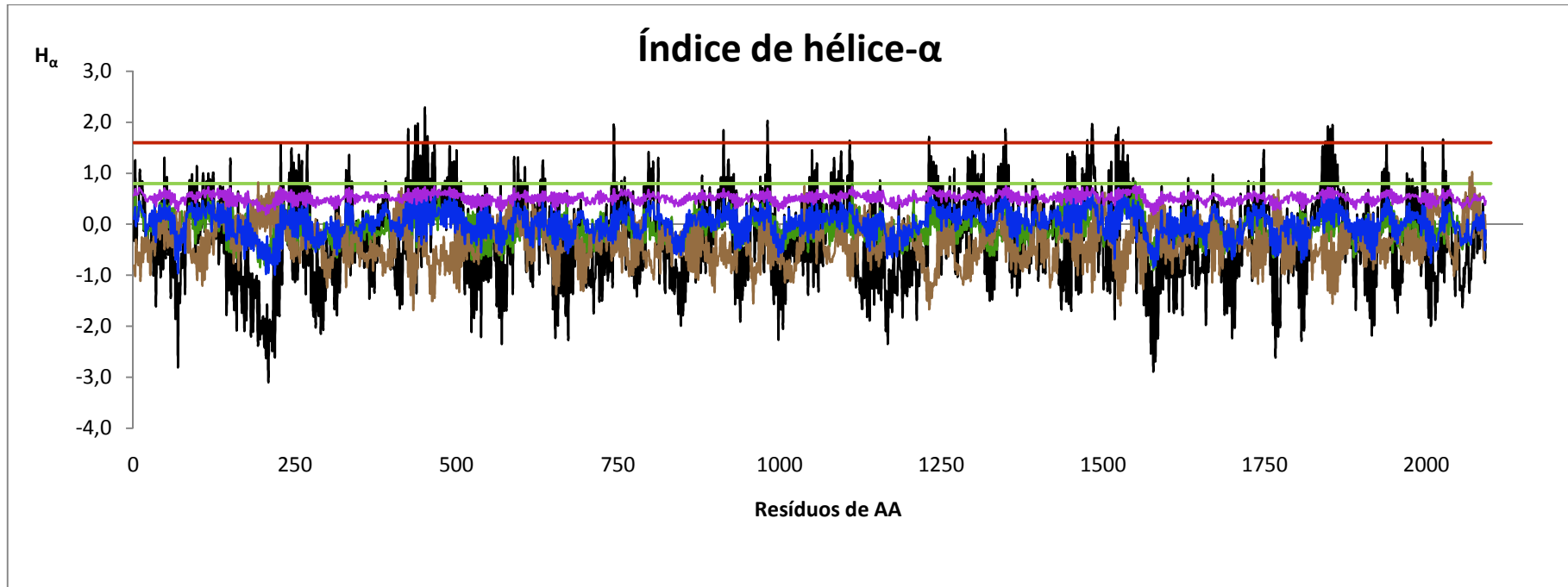


**Gráfico126:** Representação gráfica do índice de hélice- $H_{\alpha}$  para a ovalbumina (386 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

Ovomucina

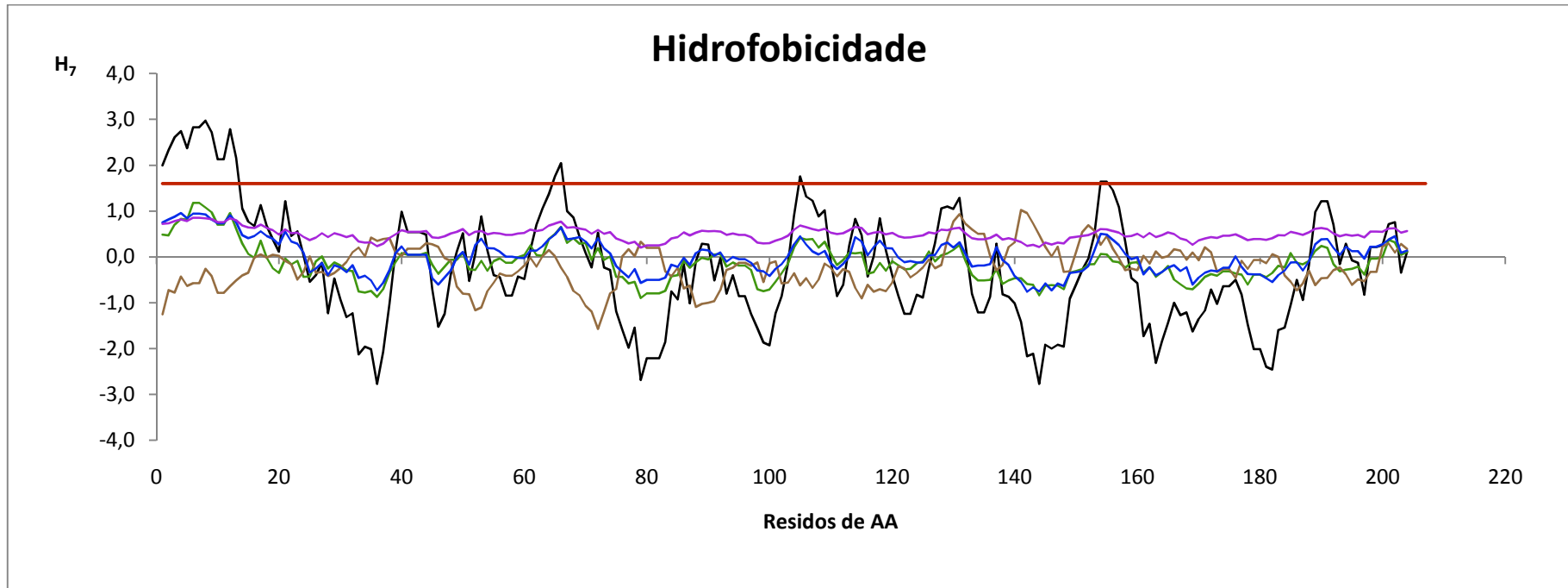


**Gráfico 127:** Representação gráfica do cálculo da hidrofobicidade  $H_7$  para a ovomucina (2108 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

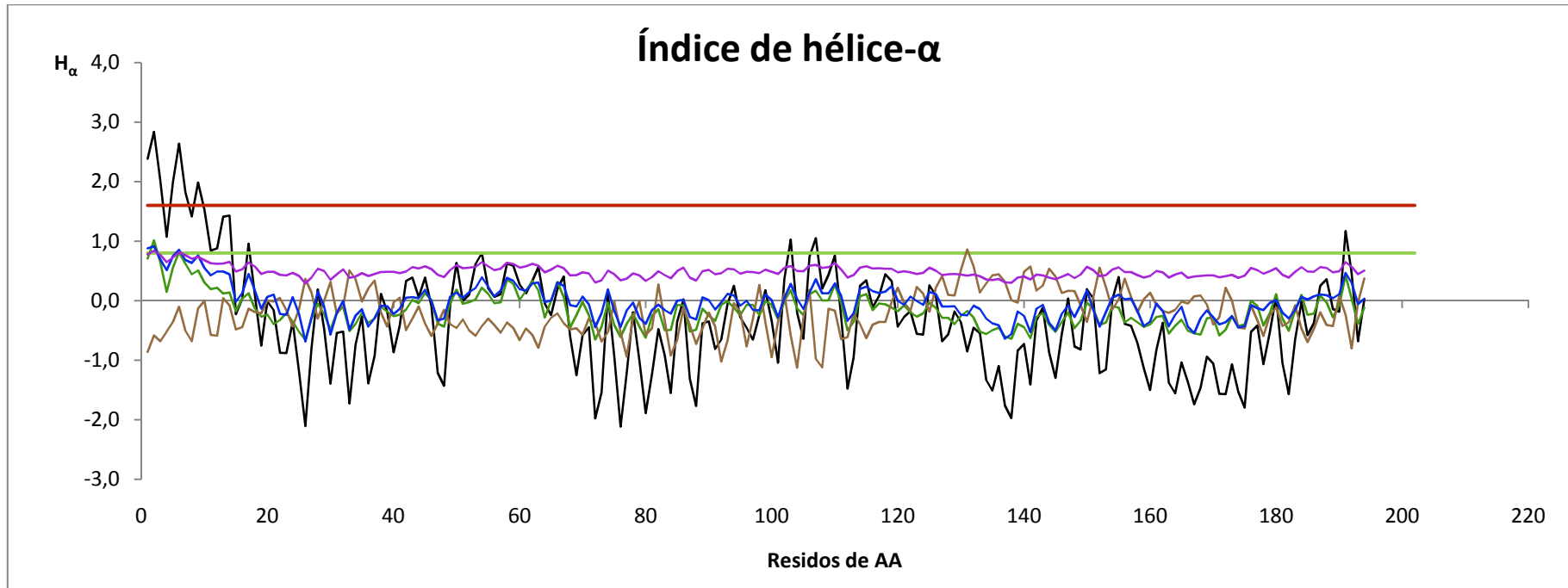


**Gráfico 128:** Representação gráfica do índice de hélice- $\alpha$   $H_{\alpha}$  para a ovomucina (2108 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

Ovomucóide



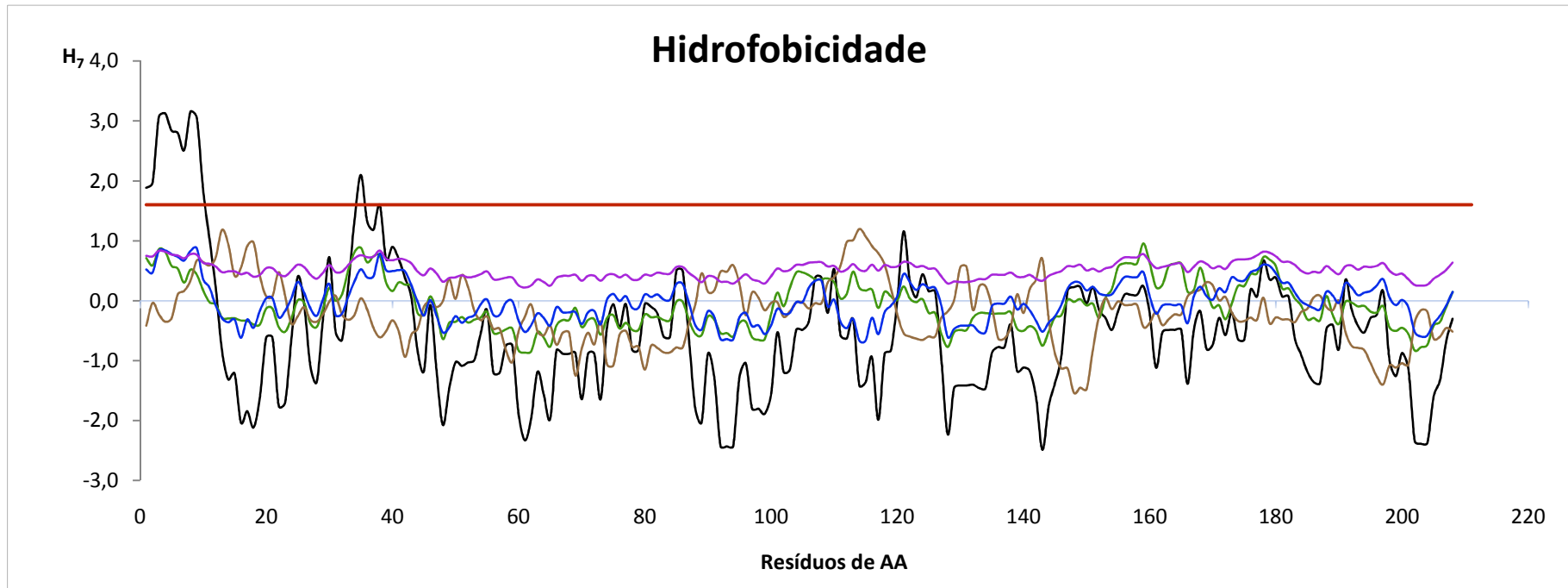
**Gráfico 129:** Representação gráfica do cálculo da hidrofobicidade H7 para a ovomucóide (210 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).



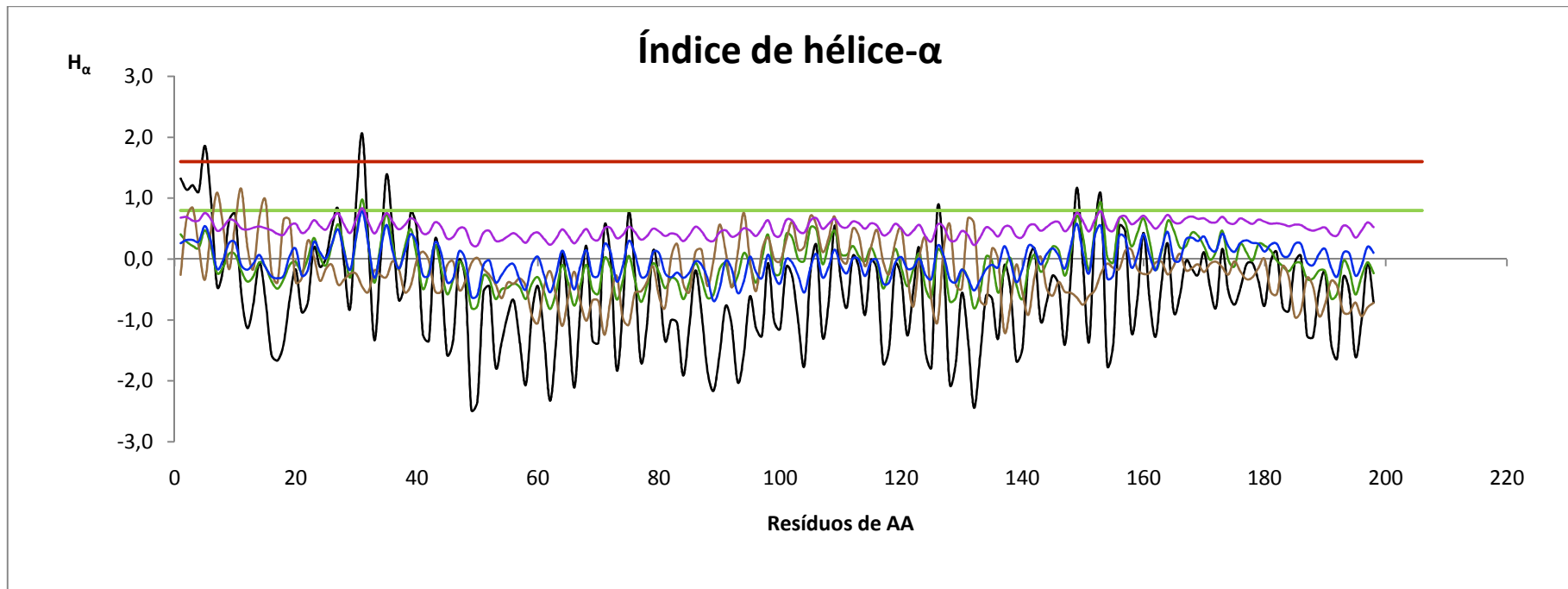
**Gráfico 130:** Representação gráfica do índice de hélice- $\alpha$   $H_{\alpha}$  para a ovomucóide (210 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

## Proteínas do leite

### A-caseína s1



**Gráfico 131:** Representação gráfica do cálculo da hidrofobicidade H7 para a  $\alpha$ -caseína s1 (214 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).



**Gráfico 132:** Representação gráfica do índice de hélice- $H_{\alpha}$  para a  $\alpha$ -caseína s1 (214 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

A-caseína s2

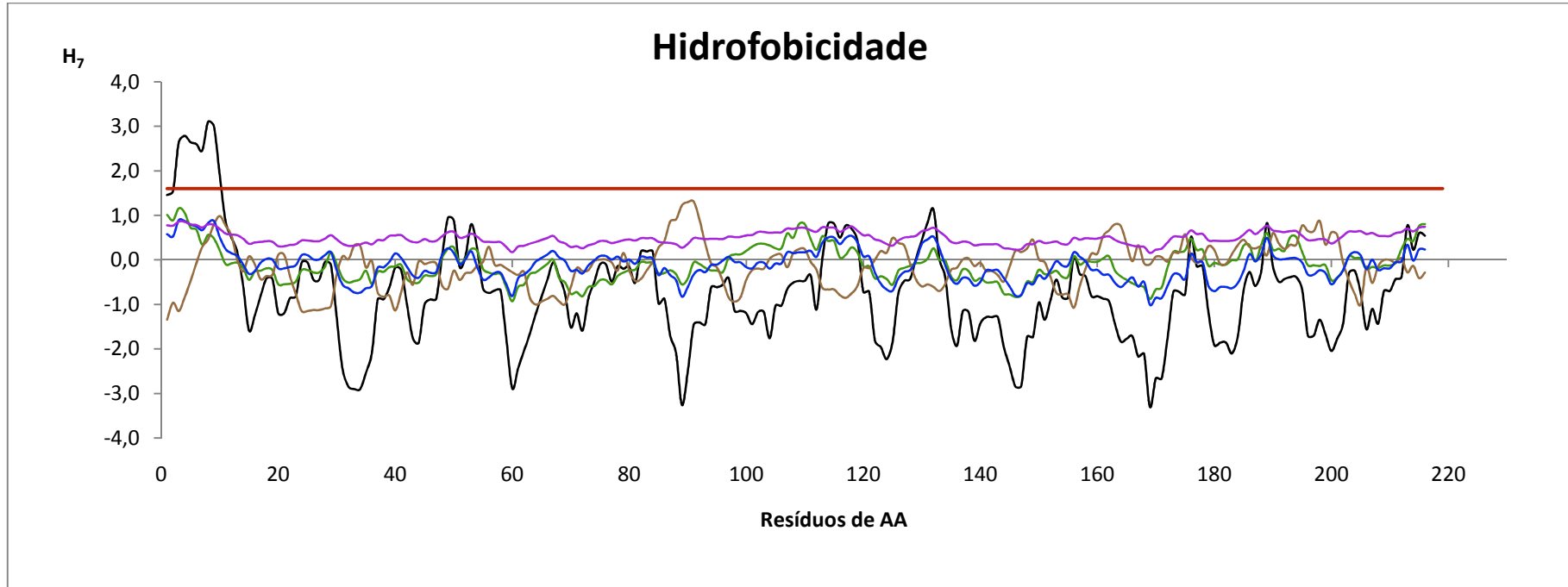
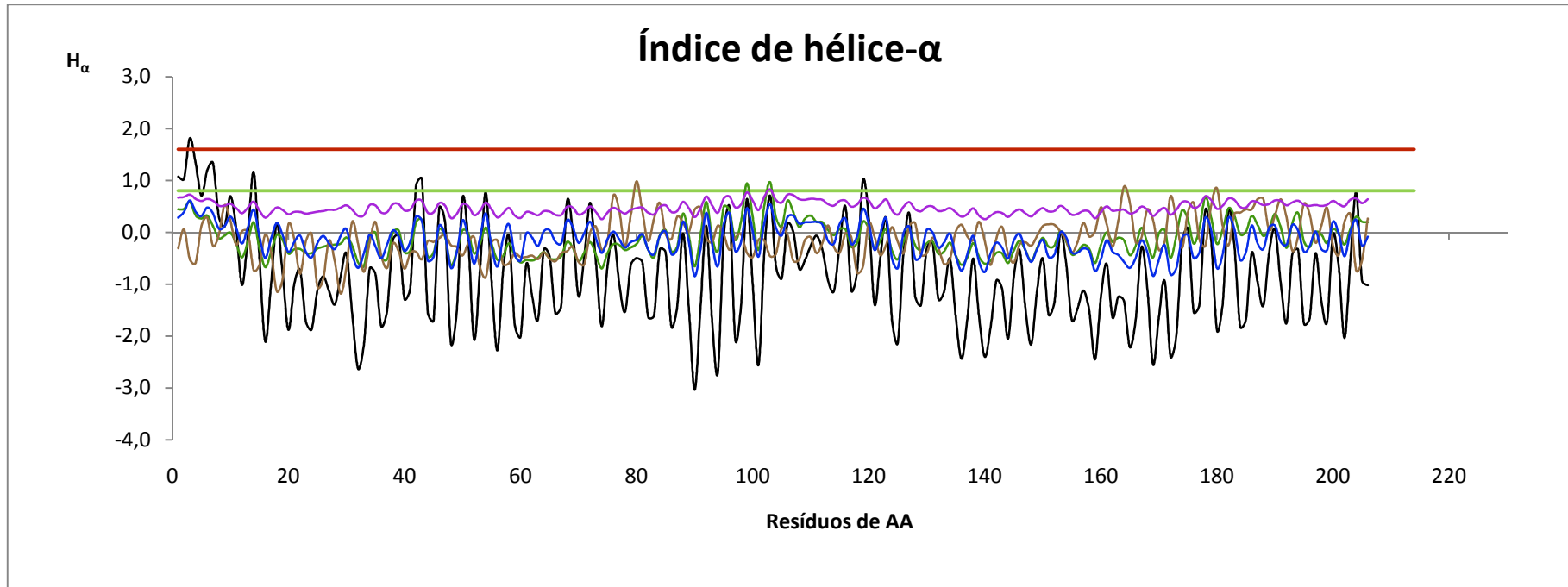


Gráfico 133: Representação gráfica do cálculo da hidrofobicidade H7 para a α-caseína s2 (222 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).



**Gráfico 134:** Representação gráfica do índice de hélice- $H_{\alpha}$  para a  $\alpha$ -caseína s2 (222 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

B-caseína

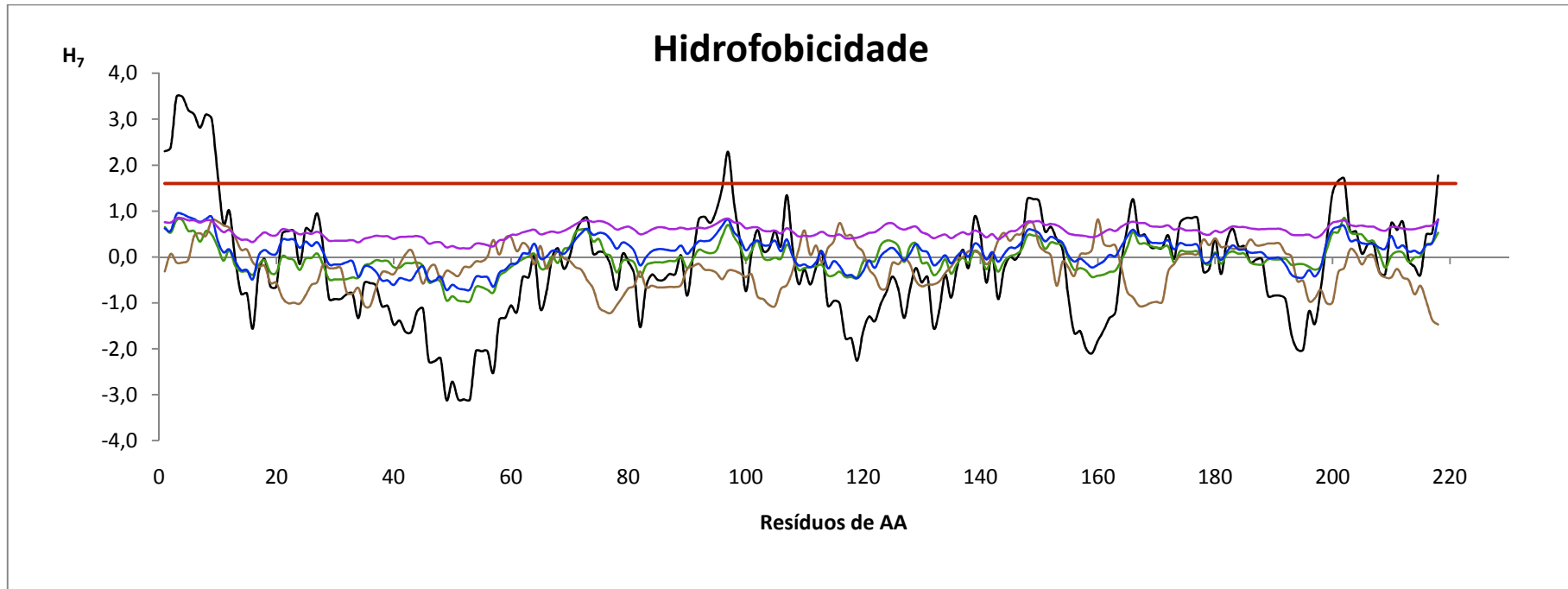
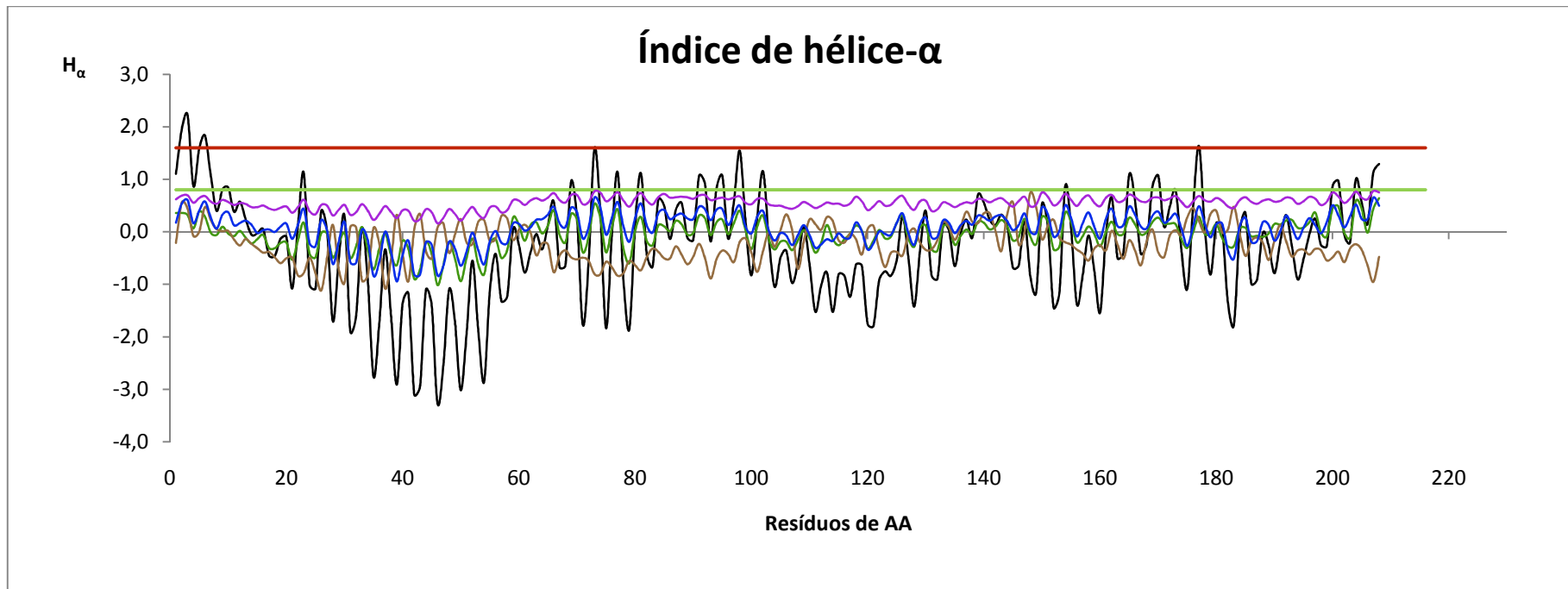
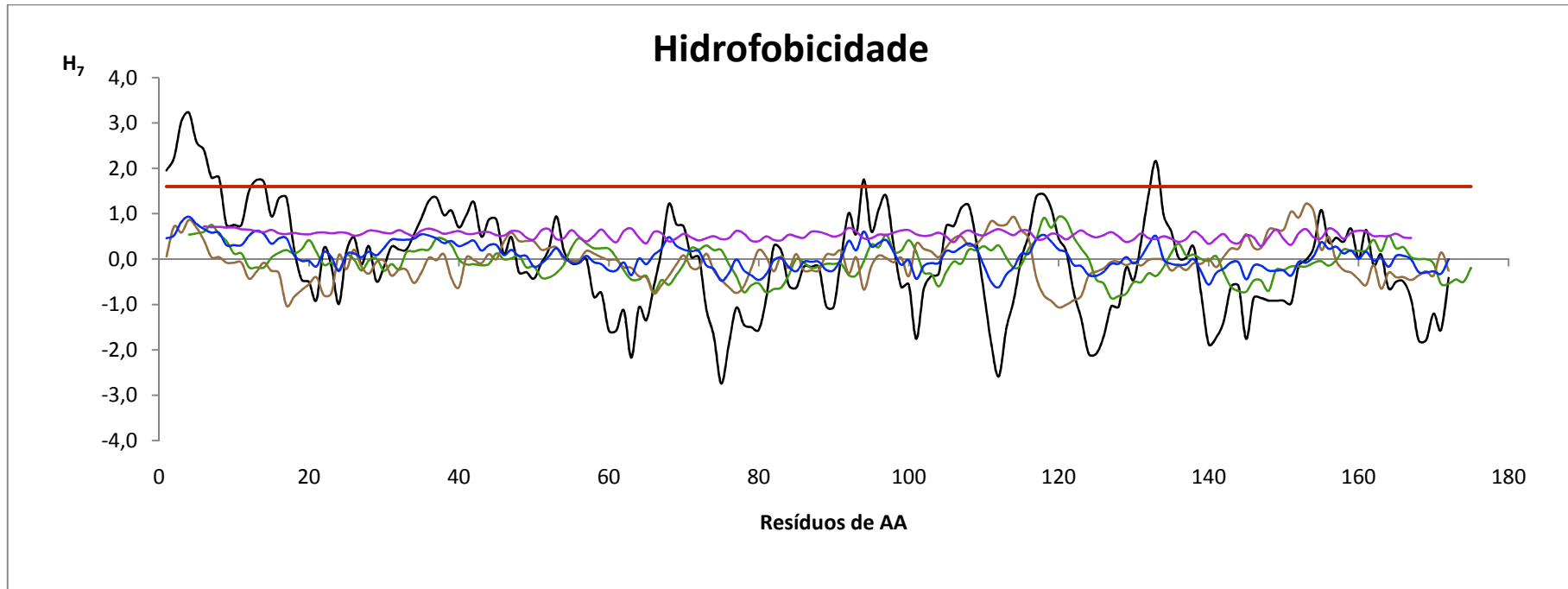


Gráfico 135: Representação gráfica do cálculo da hidrofobicidade H7 para a  $\beta$ -caseína (224 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

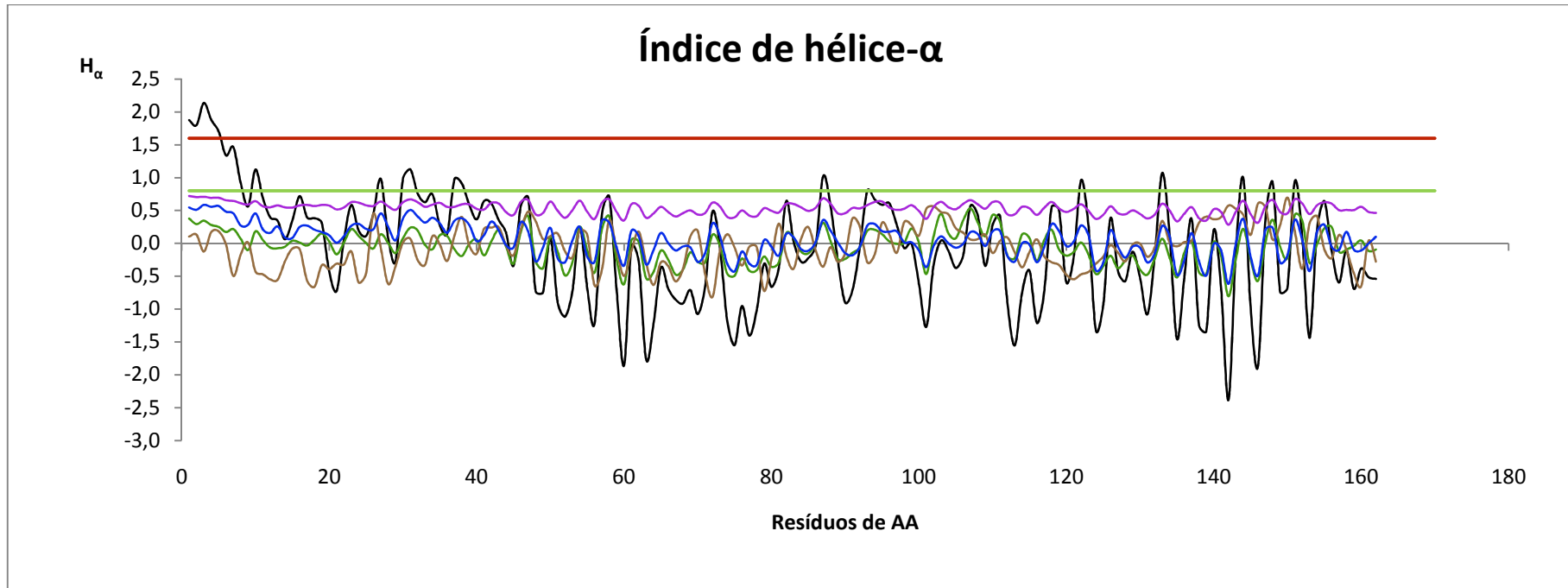


**Gráfico 136:** Representação gráfica do índice de hélice- $\alpha$   $H_{\alpha}$  para a  $\beta$ -caseína (224 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

### B-lactoglobulina



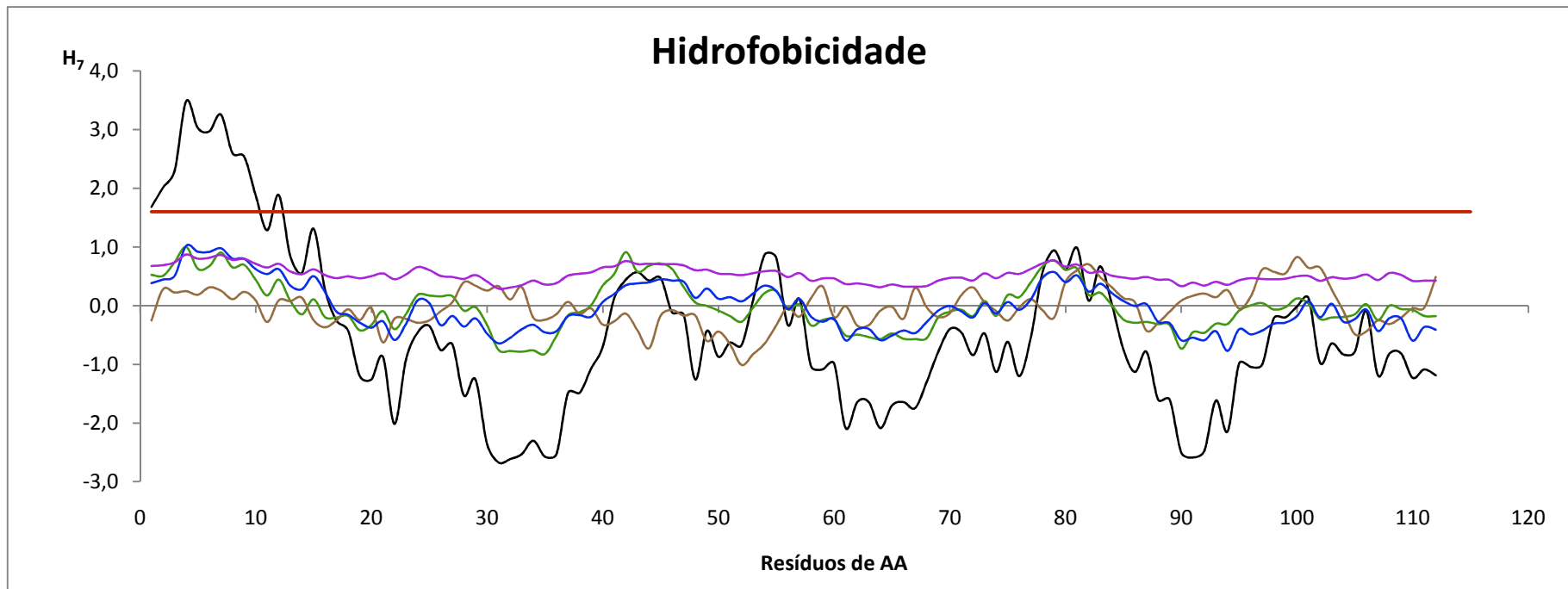
**Gráfico 137:** Representação gráfica do cálculo da hidrofobicidade H7 para a B-lactoglobulina (178 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).



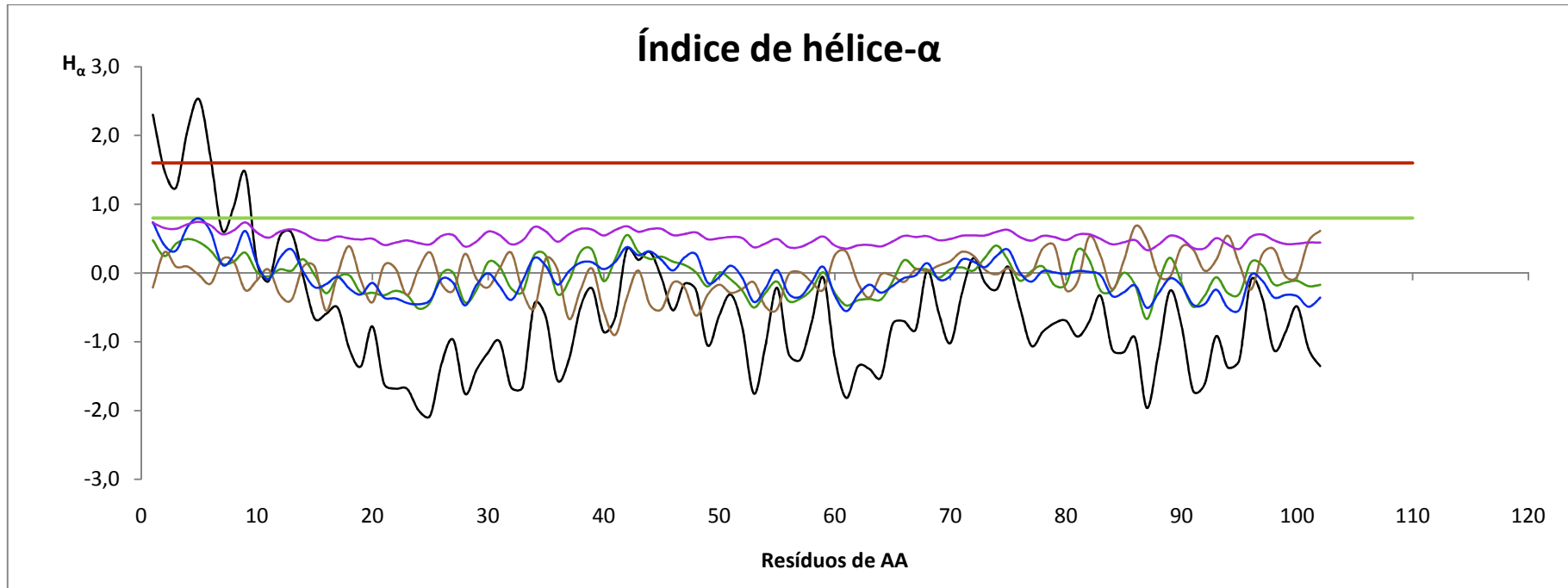
**Gráfico 138:** Representação gráfica do índice de hélice- $\alpha$   $H_{\alpha}$  para a  $\beta$ -lactoglobulina (178 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

## Proteínas do organismo

### B-2-microglobulina

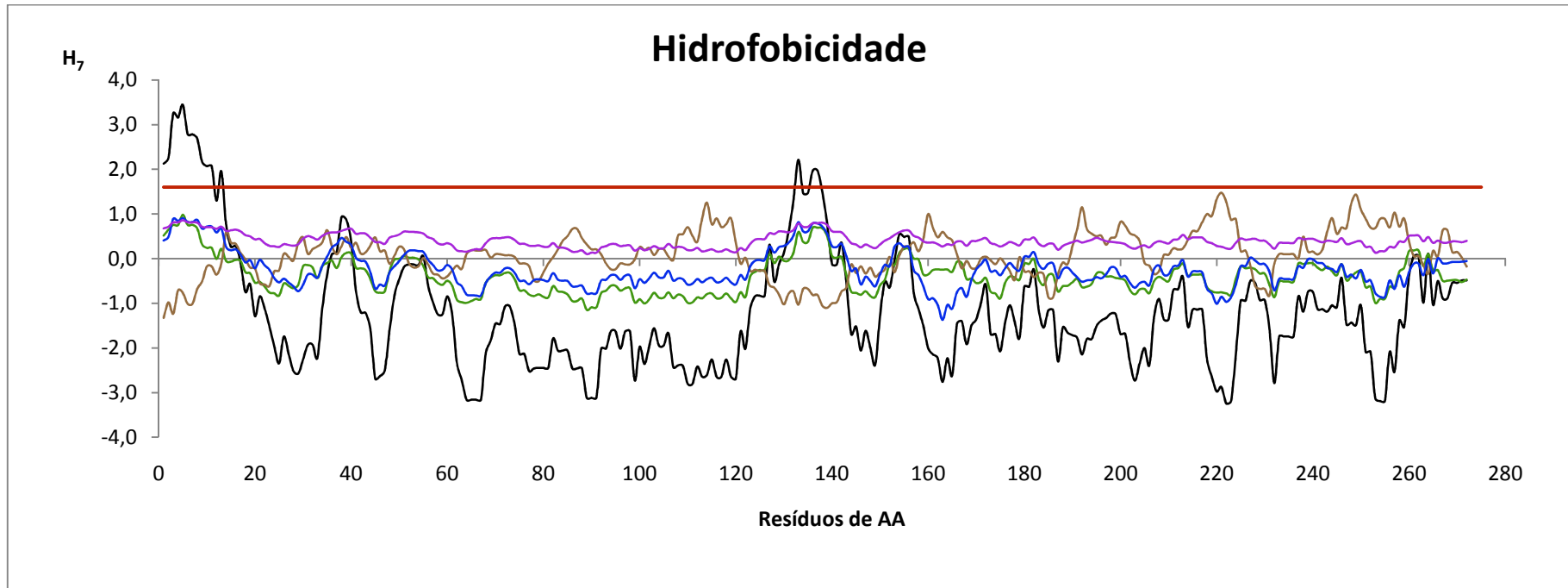


**Gráfico 139:** Representação gráfica do cálculo da hidrofbicidade H7 para a B-2-microglobulina (118 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

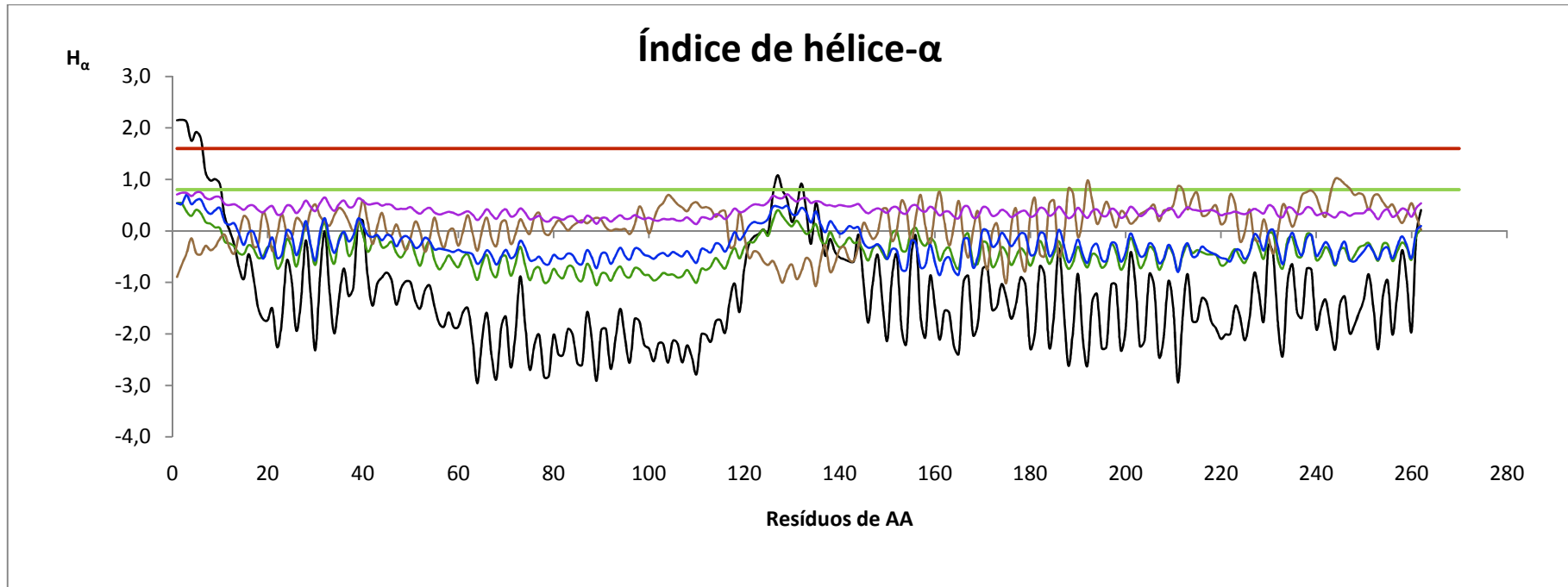


**Gráfico 140:** Representação gráfica do índice de hélice- $\alpha$   $H_{\alpha}$  para a  $\beta$ -2-microglobulina (118 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

### Osteopontina

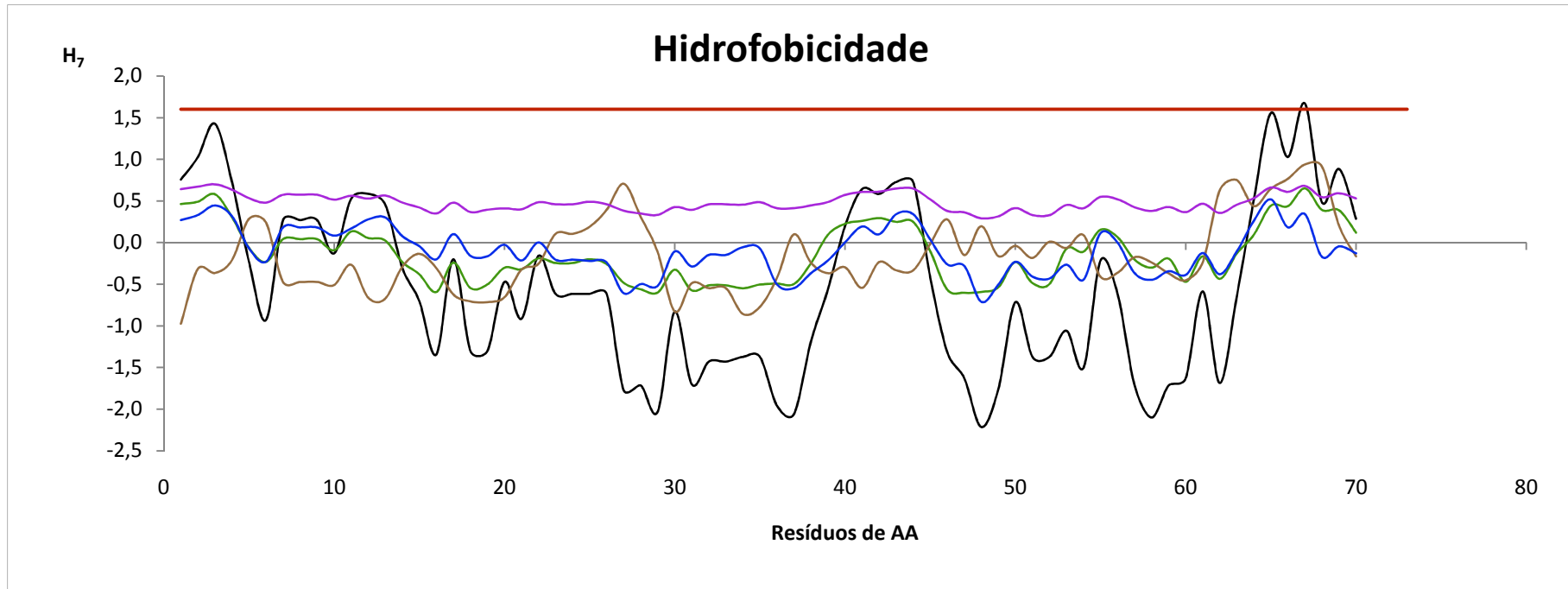


**Gráfico 141:** Representação gráfica do cálculo da hidrofobicidade H7 para a osteopontina (278 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

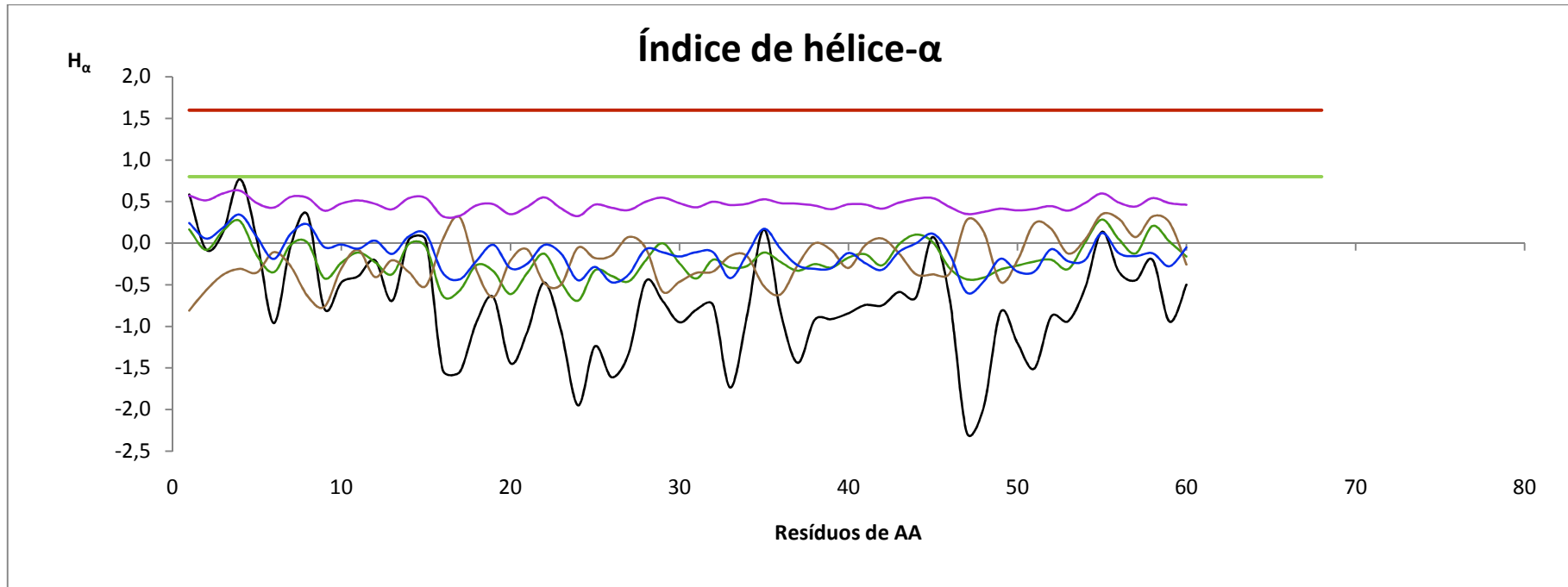


**Gráfico 142:** Representação gráfica do índice de hélice- $\alpha$   $H_{\alpha}$  para a osteopontina (278 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).

### Ubiquitina



**Gráfico 143:** Representação gráfica do cálculo da hidrofobicidade H<sub>7</sub> para a ubiquitina (76 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).



**Gráfico 144:** Representação gráfica do índice de hélice- $\alpha$   $H_{\alpha}$  para a ubiquitina (76 resíduos de aminoácido). No gráfico estão representadas as seguintes escalas: Kyte & Doolittle (Preto), Sweet (verde), Welling (castanho), Eisenberg (azul) e Black (roxo).