



(19) 中華民國智慧財產局

(12) 發明說明書公開本

(11) 公開編號：TW 202000238 A

(43) 公開日：中華民國 109 (2020) 年 01 月 01 日

(21) 申請案號：108106562

(22) 申請日：中華民國 108 (2019) 年 02 月 26 日

(51) Int. Cl. : *A61K47/65 (2017.01)*

A61K47/68 (2017.01)

C12N15/09 (2006.01)

C07K14/00 (2006.01)

(30) 優先權：2018/02/26 美國

62/634,985

(71) 申請人：比利時商艾伯林克斯公司 (比利時) ABLYNX NV (BE)

比利時

(72) 發明人：迪 巴班德瑞 維諾尼卡 DE BRABANDERE, VERONIQUE (BE)；布里吉 安 BRIGE, ANN (BE)；史丹森斯 派屈克 STANSSENS, PATRICK (BE)；迪 巴克 皮耶特 真 DE BOCK, PIETER-JAN (BE)；魅爾奇爾斯 湯姆 MERCHERS, TOM (BE)；迪 佛戈羅利斯 安東尼 DE FOUGEROLLES, ANTONIN (CA)

(74) 代理人：陳長文

申請實體審查：無 申請專利範圍項數：25 項 圖式數：12 共 59 頁

(54) 名稱

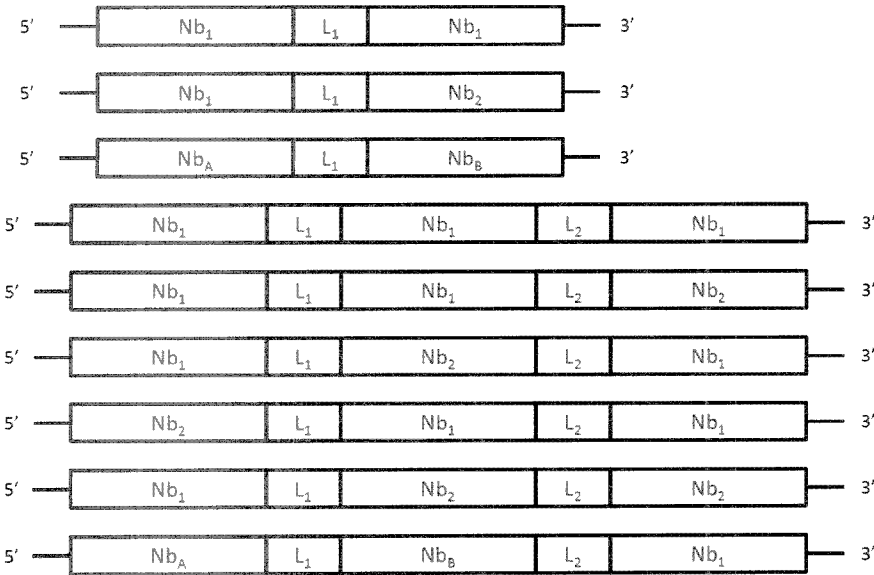
編碼肽連接子之經改良核苷酸序列

(57) 摘要

本發明提供編碼甘胺酸絲胺酸連接子且使用過量 GGA、GGG 及 GGT/GGU 密碼子以編碼該等甘胺酸殘基之經改良核苷酸序列及核酸。本發明進一步關於編碼包含甘胺酸絲胺酸連接子之(融合)蛋白質及多肽之核苷酸序列及核酸，該等核苷酸序列及核酸包含本發明此類經改良之核苷酸序列及核酸。

The invention provides improved nucleotide sequences and nucleic acids that encode glycine serine linkers and that use an excess of GGA, GGG, and GGT/GGU codons to encode the glycine residues. The invention further relates to nucleotide sequences and nucleic acids that encode (fusion) proteins and polypeptides comprising glycine serine linkers, which nucleotide sequences and nucleic acids comprise such improved nucleotide sequences and nucleic acids of the invention.

指定代表圖：



【圖1】

【發明說明書】

【中文發明名稱】

編碼肽連接子之經改良核苷酸序列

【英文發明名稱】

IMPROVED NUCLEOTIDE SEQUENCES ENCODING PEPTIDE LINKERS

【技術領域】

【0001】 本發明係關於編碼肽連接子之經改良核苷酸序列及核酸。

【0002】 本發明亦係關於編碼含有肽連接子之(融合)蛋白質及多肽之核苷酸序列及核酸，該等核苷酸序列及核酸含有此類編碼肽連接子之經改良核苷酸序列及核酸。

【0003】 本發明亦關於用於表現/產生含有肽連接子之(融合)蛋白質及多肽之方法，其涉及使用此類編碼肽連接子之經改良核苷酸序列及核酸。

【0004】 本發明之其他態樣、實施例、用途及優勢將自本文中之進一步描述而變得清楚。

【先前技術】

【0005】 使用肽連接子以連接兩個或更多個蛋白質、肽、肽部分、結合域或結合單元在此項技術中已熟知。一種常用類別之肽連接子被稱為「Gly-Ser」或「GS」連接子。此等為基本上由甘胺酸(G)及絲胺酸(S)殘基組成之連接子，且通常包含諸如GGGGS基元之肽基元之一或多個重複序列(例如，具有式(Gly-Gly-Gly-Gly-Ser)_n，其中n可為1、2、3、4、5、6、7或更大)。此類GS連接子之一些常用實例為15GS連接子(n = 3)及

35GS連接子($n = 7$)。舉例而言，參考Chen等人, *Adv. Drug Deliv. Rev.* 2013年10月15日; 65(10): 1357-1369；及Klein等人, *Protein Eng. Des. Sel.* (2014) 27 (10): 325-330。

【0006】 多肽及包含此類GS連接子之(融合)蛋白質通常係藉由適當地表現包含兩個或更多個編碼待連接之相關肽部分之核苷酸序列的遺傳構築體來產生，其中編碼肽部分之此等核苷酸序列經一或多個編碼一或多個GS連接子之核苷酸序列適合且可操作地連接，以使得在適合的宿主細胞或宿主生物體中適合的表現之後，視情況在適用於分離及/或純化之步驟之後獲得所需融合蛋白質或多肽。此類基因構築體之一些較佳但非限制性實例(使用Nanobodies作為待連接之肽之代表性實例，參見表III之圖例)示意性地展顯示於圖1中，其中NB₁、NB₂、NB_A、NB_B等指示編碼待連接之肽部分的核苷酸序列，且L₁、L₂、L₃等指示編碼適合的GS連接子之核苷酸序列。此類基因構築體可為DNA或RNA，且可例如呈適合載體，諸如表現載體之形式。此全部在蛋白質工程改造之技術中為熟知的；例如參考標準手冊，諸如本文中所提及之Sambrook等人及Ausubel等人。

【0007】 通常亦已知，歸因於基因密碼之簡併，在編碼GS連接子之核苷酸序列中，四種不同密碼子中之各者可用於編碼甘胺酸殘基，即GGU (或GGT)、GGC、GGA及/或GGG (類似地已知GS連接子中之絲胺酸殘基可由UCU (或TCT)、UCC (或TCC)、UCA (或TCA)、UCG (或TCG)、AGU (或AGT)及/或AGC密碼子編碼)。

【發明內容】

【0008】 現已發現，可藉由使用過量GGA及GGG密碼子以編碼GS連接子中之甘胺酸殘基(亦即與GGT/GGU及/或GGC密碼子之量相比)來提

供編碼GS連接子之經改良核苷酸序列。

【0009】 已進一步發現，可藉由使用過量GGA、GGG及GGT/GGU密碼子以編碼GS連接子中之甘胺酸殘基(亦即與GGC密碼子之量相比)來提供編碼GS連接子之經改良核苷酸序列。

【0010】 因此，在第一態樣中，本發明係關於一種編碼GS連接子之核苷酸序列及/或核酸(如進一步定義於本文中)，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGA、GGG或GGT/GGU。

【0011】 在此態樣中，本發明亦係關於一種編碼GS連接子之核苷酸序列及/或核酸(如進一步定義於本文中)，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGA或GGG。

【0012】 在此態樣中，本發明亦係關於一種編碼GS連接子之核苷酸序列及/或核酸(如進一步定義於本文中)，其中少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGC。

【0013】 在另一態樣中，本發明係關於一種編碼肽連接子之核苷酸序列及/或核酸，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGA、GGG或GGT/GGU。

【0014】 在此態樣中，本發明亦係關於一種編碼肽連接子之核苷酸序列及/或核酸(如進一步定義於本文中)，其中由該核苷酸序列或核酸編碼之肽連接子包含甘胺酸及絲胺酸殘基或基本上由甘胺酸及絲胺酸殘基組

成，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGA或GGG。

【0015】 在此態樣中，本發明亦係關於一種編碼肽連接子之核苷酸序列及/或核酸，其中少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGC。

【圖式簡單說明】

【0016】 圖1示意性地顯示含有連接子之奈米抗體構築體之一些非限制性實例；

【0017】 圖2示意性地顯示實例1中用以說明本發明之四價奈米抗體構築體。圖2亦顯示T10肽在此構築體中之定位；

【0018】 圖3顯示肽T10之胺基酸序列(SEQ ID NO:10)及密碼子使用(SEQ ID NO:11)。在該序列中，其中觀測到與天冬胺酸錯誤配對(misincorporation)之胺基酸殘基及密碼子以**粗體/下劃線**指示(應注意，對於以斜體/下劃線指示之殘基/密碼子，可能已預期到錯誤配對，但未觀測到)。

【0019】 圖4顯示奈米抗體構築體A中之35個GS連接子之胺基酸序列(SEQ ID NO:12)及編碼序列(SEQ ID NO:13至15)。容易發生天冬胺酸錯誤配對的甘胺酸之特異性密碼子(GGT及GGC)以**粗體/下劃線**指示。絲胺酸之密碼子標註於小帽中。

【0020】 圖5顯示在UV 254 nm (紅色(下部)跡線)及UV 280 nm (藍色(上部)跡線)下記錄之源15S柱(GE Healthcare Life Sciences)上之經純

化奈米抗體構築體A及pH梯度(綠色跡線，CX-1 pH梯度緩衝液A (pH 5.6) 及B (pH 10.2)、熱科學)之陽離子交換層析圖。pH記錄顯示於灰色跡線中。預峰為奈米抗體構築體A之酸性變體。合併溶離份14、15、16及17以用於酸性變體之後續表徵，且溶離份18用於主峰之表徵；

【0021】 圖6顯示自經純化奈米抗體構築體A之陽離子交換分級分離收集之酸性變體(上片)及主峰(下片)獲得之Max-ent解卷積質譜。酸性溶離份中所量測之最重要的質量為59689.4 Da，其比在pH-IEX主峰溶離份(59630.9 Da，參見下片)中所量測之奈米抗體構築體A之質量高58道爾頓；

【0022】 圖7列出由Asp-N消化物(在天冬胺酸之N端處裂解之內切蛋白酶)產生之胰蛋白酶肽T10之肽片段(SEQ ID NO:16至33)。各裂解位點對應於與天冬胺酸交換之甘胺酸；

【0023】 圖8顯示在(a)奈米抗體構築體A之GS連接子中三個位點(C1、C2及C3)之Gly與Asp錯誤配對的相對程度；(b)在藉由pH-IEX消耗之具有Asp錯誤配對之變體之後的奈米抗體構築體A；(c)其中100%之GGC密碼子序列經GGG、GGA或GGT密碼子序列置換之奈米抗體構築體A；

【0024】 圖9顯示產生以研究價數及連接子長度對如實例3中所描述之Gly與Asp錯誤配對之影響的十個構築體；

【0025】 圖10顯示9GS連接子中兩個位點(C1及C2)之Gly與Asp錯誤配對之相對程度；(A)二價構築體、(B)三價構築體、(C)四價構築體；

【0026】 圖11顯示20GS連接子中五個位點(C1、C2、C3、C4及C5)之Gly與Asp錯誤配對之相對含量；(A)二價構築體、(B)三價構築體、(C)

四價構築體；

【0027】 圖12顯示35GS連接子中九個位點(C1至C9)之Gly與Asp錯誤配對之相對程度；(A)二價構築體、(B)三價構築體及(C)四價構築體、(D)無GGC密碼子之四價構築體。

【實施方式】

【0028】 如本文中進一步描述，由該等核苷酸序列或核酸編碼之肽連接子將一般包含至少5個胺基酸殘基及高達50個或更多個胺基酸殘基(但實際上將通常包含10個與40個之間的胺基酸殘基，諸如約15個胺基酸殘基至約35個胺基酸殘基)。此外，如本文中進一步描述，由該等核苷酸序列或核酸編碼之肽連接子與絲胺酸殘基之數目相比將通常含有過量甘胺酸殘基，例如各絲胺酸殘基之3個與6個之間的甘胺酸殘基。此外，由該等核苷酸序列或核酸編碼之肽連接子通常將含有序列基元之一或多個(諸如兩個或更多個)重複序列。

【0029】 在另一態樣中，本發明係關於一種編碼肽連接子之核苷酸序列及/或核酸(如本文中進一步描述)，其中由該核苷酸序列或核酸編碼之肽連接子包含序列基元GGGGS (SEQ ID NO:1)之一或多個(諸如兩個或更多個)重複序列或基本上由其組成，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGA、GGG或GGT/GGU。

【0030】 在此態樣中，本發明亦係關於一種編碼肽連接子之核苷酸序列及/或核酸(如本文中進一步描述)，其中由該核苷酸序列或核酸編碼之肽連接子包含序列基元GGGGS (SEQ ID NO:1)之一或多個(諸如兩個或更多個)重複序列或基本上由其組成，其中超過70%、較佳超過85%、更佳

超過90%，諸如超過95%及高達99%及更多(包括100%)編碼該肽連接子中之甘胺酸殘基的密碼子為GGA或GGG。

【0031】 在此態樣中，本發明亦係關於一種編碼肽連接子之核苷酸序列及/或核酸(如本文中進一步描述)，其中由該核苷酸序列或核酸編碼之肽連接子包含序列基元GGGGS (SEQ ID NO:1)之一或多個(諸如兩個或更多個)重複序列或基本上由其組成，其中少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGC。

【0032】 舉例而言，在本發明之此態樣中，由該核苷酸序列或核酸編碼之肽連接子可包含序列基元GGGGS之2、3、4、5、6、7、8、9或10個重複序列或基本上由其組成。

【0033】 在另一態樣中，本發明係關於一種編碼肽連接子之核苷酸序列及/或核酸(如本文中進一步描述)，其中由該核苷酸序列或核酸編碼之肽連接子具有式(Gly-Gly-Gly-Gly-Ser)_n (其中n可為1、2、3、4、5、6、7或更大)，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGA、GGG或GGT/GGU。

【0034】 在此態樣中，本發明亦係關於一種編碼肽連接子之核苷酸序列及/或核酸(如本文中進一步描述)，其中由該核苷酸序列或核酸編碼之肽連接子具有式(Gly-Gly-Gly-Gly-Ser)_n (其中n可為1、2、3、4、5、6、7或更大)，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGA或GGG。

【0035】 在此態樣中，本發明亦係關於一種編碼肽連接子之核苷酸序列及/或核酸(如本文中進一步描述)，其中由該核苷酸序列或核酸編碼之肽連接子具有式(Gly-Gly-Gly-Gly-Ser)_n (其中n可為1、2、3、4、5、6、7或更大)，其中少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGC。

【0036】 舉例而言，在本發明之此態樣中，由該核苷酸序列或核酸編碼之肽連接子可包含序列基元GGGGS之2、3、4、5、6、7、8、9或10個重複序列或基本上由其組成。

【0037】 在另一態樣中，本發明係關於以下通式之核苷酸序列及/或核酸

$$(A_x-B_p-A_y-B_q)_n,$$

其中：

- A表示編碼甘胺酸殘基之密碼子，其可獨立地為(選自)GGU (或GGT)、GGC、GGA及/或GGG密碼子；且
- B表示編碼絲胺酸殘基之密碼子，其可獨立地為(選自)UCU (或TCT)、UCC (或TCC)、UCA (或TCA)、UCG (或TCG)、AGU (或AGT)及/或AGC密碼子；
- x為0至10 (且較佳0至5)之整數，且y為0至10 (且較佳0至5)之整數，使得(x+y)之總和在1與10之間，且較佳在3、4、5、6、7或8之間；
- p為0或1，且q為0或1，使得(p+q)之總和為2或1且較佳為1；
- n為1至10之整數(亦即，使得核苷酸序列及/或核酸包含基元(A_x-B_p-A_y-B_q)之n個重複序列，其中A、B、p、q、x及y如本文中所描述)；

- 在基元($A_x-B_p-A_y-B_q$)之各重複序列中，各A、B、p、q、x及y可獨立地如本文中所描述(但根據一較佳態樣，在基元($A_x-B_p-A_y-B_q$)之各重複序列中，各A、B、p、q、x及y係相同的)；

其限制條件為超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGA、GGG或GGT/GGU；

其限制條件為超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGA或GGG；及/或

其限制條件為少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGC。

【0038】 在另一態樣中，本發明係關於以下通式之核苷酸序列及/或核酸

$(A_x-B)_n$ ，

其中：

- A表示編碼甘胺酸殘基之密碼子，其可獨立地為(選自)GGU (或GGT)、GGC、GGA及/或GGG密碼子；且

- B表示編碼絲胺酸殘基之密碼子，其可獨立地為(選自)UCU (或TCT)、UCC (或TCC)、UCA (或TCA)、UCG (或TCG)、AGU (或AGT)及/或AGC密碼子；

- x為1至10之整數，且較佳為3、4、5、6、7或8；

- n為1至10之整數(亦即，使得核苷酸序列及/或核酸包含基元(A_x-B))

之 n 個重複序列，其中各A、B及 x 如本文中所描述)；

- 在基元(A_x-B)之各重複序列中，各A、B及 x 可獨立地如本文中所描述(但根據一較佳態樣，在基元(A_x-B)之各重複序列中，各A、B及 x 係相同的)；

其限制條件為超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGA、GGG或GGT/GGU；

其限制條件為超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGA或GGG；及/或

其限制條件為少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGC。

【0039】 在另一態樣中，本發明係關於表I中所示之化學式中的一者的核苷酸序列及/或核酸，其中：

- A表示編碼甘胺酸殘基之密碼子，其可獨立地為(選自)GGU (或GGT)、GGC、GGA及/或GGG密碼子；且

- B表示編碼絲胺酸殘基之密碼子，其可獨立地為(選自)UCU (或TCT)、UCC (或TCC)、UCA (或TCA)、UCG (或TCG)、AGU (或AGT)及/或AGC密碼子；

其限制條件為超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGA、GGG或GGT/GGU；

其限制條件為超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGA或GGG；及/或

其限制條件為少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼甘胺酸殘基的密碼子(如由表I之化學式中之A表示)為GGC。

【0040】 一般而言，編碼Gly-Ser連接子且其中該等GS連接子中之甘胺酸殘基顯著地或排他性地由GGA、GGG或GGT/GGU密碼子編碼之本文中所描述的核苷酸序列及核酸在本文中亦被稱為「本發明之GS連接子-編碼序列」。一般而言，編碼Gly-Ser連接子且其中該等GS連接子中之甘胺酸殘基顯著地或排他性地由GGA或GGG密碼子編碼之本文中所描述的核苷酸序列及核酸在本文中亦被稱為「本發明之GS連接子-編碼序列」。一般而言，編碼Gly-Ser連接子且其中該等GS連接子中之甘胺酸殘基中之幾乎無一者或非任一者由GGC密碼子編碼的本文中所描述之核苷酸序列及核酸在本文中亦被稱為「本發明之GS連接子-編碼序列」。

【0041】 在本發明之一個較佳但非限制性態樣中，超過95%及高達99%或更多(且包括100%)之在本發明之GS連接子-編碼序列中編碼甘胺酸殘基的密碼子為GGA、GGG或GGT/GGU。

【0042】 在本發明之一個較佳但非限制性態樣中，大於95%及高達99%或更多(且包括100%)之在本發明之GS連接子-編碼序列中編碼甘胺酸殘基的密碼子為GGA或GGG。

【0043】 在本發明之一個較佳但非限制性態樣中，少於5%及高達少於1%或更低(且包括0%)之在本發明GS連接子-編碼序列中編碼甘胺酸殘

基的密碼子為GGC。表II給出本發明之GS連接子-編碼序列之一些代表性但非限制性實例。本發明之GS連接子-編碼序列之其他實例將基於本文中之揭示內容為熟習此項技術者所清楚。

表I：

[illegible]

表II：

[illegible]

		AGCGGAGGGGGAGGGAGCGGAGGGGGAGGGAGCGGAGGGGGA GGGAGCGGAGGGGGAGGGAGC
9	35GS	GGAGGAGGAGGAUCUGGGGGGGGGGGGUCCGGAGGAGGAGGA UCAGGGGGGGGGGGGUCGGGAGGAGGAGGAAGUGGGGGGGGG GGGAGCGGAGGAGGAGGAUCU

【0044】 不限於任何特定解釋、假設或機制，推斷使用此類核苷酸序列(亦即，相比於使用編碼含有較高量/比例之GGU及/或GGC密碼子之GS連接子的核苷酸序列；或相比於使用編碼含有較高量/比例之GGC密碼子之GS連接子的核苷酸序列)，會降低天冬胺酸殘基被錯誤地包括於所需GS連接子(而非預期甘胺酸殘基)中之風險及/或減少在合適宿主或宿主生物體中表現時被錯誤地包括於所需GS連接子中之天冬胺酸殘基的量。

【0045】 因此，當用於表現及/或產生融合蛋白質或多肽時，本發明亦減少在表現之產物中獲得之污染物(亦即，含有具有一或多個天冬胺酸殘基而非預期甘胺酸殘基之GS連接子的污染物)的量，且亦減少與所需GS連接子中不期望存在之天冬胺酸殘基相關的有害作用(諸如不期望的異構化形成異天冬胺酸)，以及增加對蛋白降解之易感性。

【0046】 因此在另一態樣中，本發明係關於一種編碼(融合)蛋白質或融合多肽之核苷酸序列及/或核酸，其中由該核苷酸序列及/或核酸編碼之融合蛋白質或多肽包含經一或多個GS連接子適合地連接之兩個或更多個肽部分，其中一或多個GS連接子由一或多個本發明之GS連接子-編碼序列編碼(亦即，藉由其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGG、GGG或GGT/GGU的核苷酸序列或核酸)。

【0047】 在此態樣中，本發明亦係關於一種編碼(融合)蛋白質或融合多肽之核苷酸序列及/或核酸，其中由該核苷酸序列及/或核酸編碼之融合蛋白質或多肽包含經由一或多個GS連接子適當連接之兩個或更多個肽

部分，其中一或多個GS連接子由一或多個本發明之GS連接子-編碼序列編碼(亦即，藉由其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGG或GGG的核苷酸序列或核酸)。

【0048】 在此態樣中，本發明亦係關於一種編碼(融合)蛋白質或融合多肽之核苷酸序列及/或核酸，其中由該核苷酸序列及/或核酸編碼之融合蛋白質或多肽包含經由一或多個GS連接子適合地連接之兩個或更多個肽部分，其中一或多個GS連接子由一或多個本發明之GS連接子-編碼序列編碼(亦即，藉由其中少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGC的核苷酸序列或核酸)。

【0049】 在另一態樣中，本發明係關於一種編碼(融合)蛋白質或融合多肽之核苷酸序列及/或核酸，其中由該核苷酸序列及/或核酸編碼之融合蛋白質或多肽包含經由一或多個GS連接子適合地連接之兩個或更多個肽部分，其中編碼GS連接子之核苷酸序列或核酸的部分為一或多個本發明之GS連接子-編碼序列(亦即，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGG、GGG或GGT/GGU的核苷酸序列或核酸)。

【0050】 在此態樣中，本發明亦係關於一種編碼(融合)蛋白質或融合多肽之核苷酸序列及/或核酸，其中由該核苷酸序列及/或核酸編碼之融合蛋白質或多肽包含經由一或多個GS連接子適合地連接之兩個或更多個肽部分，其中編碼GS連接子之核苷酸序列或核酸的部分為一或多個本發明之GS連接子-編碼序列(亦即，其中超過70%、較佳超過85%、更佳超過

90%，諸如超過95%及高達99%及更多(包括100%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGG或GGG的核苷酸序列或核酸)。

【0051】 在此態樣中，本發明係關於一種編碼(融合)蛋白質或融合多肽之核苷酸序列及/或核酸，其中由該核苷酸序列及/或核酸編碼之融合蛋白質或多肽包含經由一或多個GS連接子適合地連接之兩個或更多個肽部分，其中編碼GS連接子之核苷酸序列或核酸的部分為一或多個本發明之GS連接子-編碼序列(亦即，其中少於30%、較佳少於15%、更佳少於10%，諸如少於5%及高達少於1%或更低(包括0%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGC的核苷酸序列或核酸)。

【0052】 更一般而言，在另一態樣中，本發明係關於一種包含或含有一或多個本發明之GS連接子-編碼序列的核苷酸序列或核酸。此類核苷酸序列或核酸較佳使得在適合的宿主細胞或宿主生物體中表現後，其表現包含至少一個GS連接子(亦即由本發明之GS連接子-編碼序列編碼之GS連接子)之(融合)蛋白質或多肽。

【0053】 在另一態樣中，本發明係關於一種用於表現或產生(融合)蛋白質或多肽之方法，其中該(融合)蛋白質或多肽包含兩個或更多個經由一或多個GS連接子適合地連接之肽部分，該方法包含在適合的宿主細胞或宿主生物體中適合地表現編碼該(融合)蛋白質或多肽之核苷酸序列及/或核酸，其中該核苷酸序列及/或核酸包含或含有一或多個本發明之GS連接子-編碼序列(且進一步描述於本文中)。該方法可進一步包含分離/純化由此表現之(融合)蛋白質或多肽之視情況選用之步驟。

【0054】 在另一態樣中，本發明係關於一種宿主細胞或宿主生物體，其包含編碼(融合)蛋白質或多肽之核苷酸序列及/或核酸，該(融合)蛋

白質或多肽包含一或多個GS連接子，其中該核苷酸序列及/或核酸包含或含有一或多個本發明之GS連接子-編碼序列(且進一步如本文中所描述)。

【0055】 在另一態樣中，本發明係關於一種用於表現或產生(融合)蛋白質或多肽之方法，其中該(融合)蛋白質或多肽包含兩個或更多個經由一或多個GS連接子適合地連接之肽部分，該方法包含在以下條件下培養包含核苷酸序列及/或核酸之適合的宿主細胞或宿主生物體，該核苷酸序列及/或核酸包含或含有一或多個本發明之GS連接子-編碼序列(且進一步如本文中所描述)：使得該宿主細胞或宿主生物體表現/產生該(融合)蛋白質或多肽(其中該融合蛋白質或多肽包含一或多個GS連接子，亦即如本發明之GS連接子-編碼序列編碼)。該方法可進一步包含分離/純化由此表現之(融合)蛋白質或多肽之視情況選用之步驟。

【0056】 在另一態樣中，本發明係關於一種(融合)蛋白質或多肽(且特定言之，關於包含一或多個GS連接子之(融合)蛋白質或多肽)，其已藉由在適合的宿主細胞或宿主生物體中表現編碼該(融合)蛋白質或多肽之核苷酸序列或核酸而獲得，其中該核苷酸序列或核酸含有或包含一或多個本發明之GS連接子-編碼序列(且如本文中進一步描述)。

【0057】 在另一態樣中，本發明提供一種用於降低肽連接子(諸如GS連接子)中之Gly與Asp錯誤配對程度的方法，該方法包含在編碼該肽連接子之核酸序列及/或核酸中用GGG、GGA或GGT/GGU密碼子置換至少一個GGC密碼子之步驟。

【0058】 在此態樣中，本發明亦提供一種用於降低肽連接子(諸如GS連接子)中Gly與Asp錯誤配對程度的方法，該方法包含在編碼該肽連接子之核酸序列及/或核酸中用GGG或GGA置換至少一個GGC密碼子之步

驟。

【0059】 在另一態樣中，本發明提供一種用於降低存在於多價(諸如二價、三價、四價)免疫球蛋白單一可變域或奈米抗體中之肽連接子(諸如GS連接子)中之Gly與Asp錯誤配對程度的方法，該方法包含在編碼該肽連接子之核酸序列及/或核酸中用GGG、GGA或GGT/GGU密碼子置換至少一個GGC密碼子之步驟。

【0060】 在此態樣中，本發明亦提供一種用於降低存在於多價(諸如二價、三價、四價)免疫球蛋白單一可變域或奈米抗體中之肽連接子(諸如GS連接子)中之Gly與Asp錯誤配對程度的方法，該方法包含在編碼該肽連接子之核酸序列及/或核酸中用GGG或GGA置換至少一個GGC密碼子之步驟。

【0061】 本文中所描述之核苷酸序列及核酸可為DNA或RNA (且較佳為雙股DNA)且可呈基因構築體之形式(例如，呈適合載體，諸如表現載體之形式)。此類基因構築體可例如除編碼(融合)蛋白質或多肽之核苷酸序列以外，包含一或多種適用於表現該核苷酸序列之元件，諸如適合啟動子、適合的轉譯起始序列(諸如核糖體結合位點及起始密碼子)、適合的終止密碼子及適合的轉錄終止序列、3'-或5'-UTR序列、前導序列、選擇標記、表現標記/報導基因及/或可有助於或增加轉型或整合(之效率)的元件，所有均適合地(且若適當、可操作地)連接至編碼(融合)蛋白質或多肽之核苷酸序列。此類元件之適合實例將為熟習此項技術者所清楚且可例如視其中表現該(表現)載體之宿主或宿主細胞而定。

【0062】 本文中所描述之基因構築體亦可呈適用於轉型預期宿主細胞或宿主生物體之形式；呈適用於整合至預期宿主細胞之基因組DNA中

之形式；或呈適用於預期宿主生物體中之獨立複製、維持及/或遺傳之形式。舉例而言，本文中所描述之基因構築體可呈載體，諸如質體、黏質體、YAC、病毒載體或轉座子之形式。特定言之，載體可為表現載體，亦即可提供活體外及/或活體內表現之載體(例如在適合的宿主細胞、宿主生物體及/或表現系統中)。此類基因構築體及(表現)載體形成本發明之其他態樣。

【0063】 較佳地，本文中所描述之基因構築體之調節及其他元件使得其能夠在預期宿主細胞或宿主生物體中提供其預期生物功能。

【0064】 舉例而言，啟動子、強化子或終止子在預期宿主細胞或宿主生物體中應為「可操作之」，其意謂(例如)該啟動子應能夠起始或以其他方式控制/調節其可操作地連接(如本文中所定義)之核苷酸序列(例如編碼序列)之轉錄及/或表現。

【0065】 一些尤其較佳啟動子包括但不限於本身已知用於在本文所提及之宿主細胞中表現之啟動子；且尤其用於在細菌細胞，諸如本文所提及之彼等細菌細胞中表現之啟動子。

【0066】 選擇標記應使得其允許(亦即，在適當選擇條件下)已(成功地)用核苷酸序列(如本文中所描述之)轉型之宿主細胞及/或宿主生物體區別於尚未(成功地)轉型之宿主細胞/生物體。此類標記之一些較佳但非限制性實例為提供對抗生素(諸如康微素(kanamycin)或安比西林(ampicillin))之抗性之基因、提供耐熱性之基因或允許將宿主細胞或宿主生物體在不存在某些因素、化合物及/或(食物)組分之下維持於培養基中的基因，該等因素、化合物及/或(食物)組分為非轉型細胞或生物體之存活所必需的。

【0067】 前導序列應使得(在預期宿主細胞或宿主生物體中)其允許

所需轉譯後修飾及/或使得其將經轉錄之mRNA導引至細胞之所需部分或細胞器。前導序列亦可允許自該細胞分泌表現產物。如此，前導序列可為在宿主細胞或宿主生物體中可操作之任何原序列、前序列或前原序列。細菌細胞中之表現可能不需要前導序列。舉例而言，本身已知用於表現及產生抗體及抗體片段(包括但不限於單域抗體及ScFv片段)之前導序列可以基本上類似之方式使用。

【0068】 表現標記或報導基因應使得(在宿主細胞或宿主生物體中)其允許偵測基因構築體(存在於其上之基因或核苷酸序列)之表現。表現標記可視情況亦允許所表現之產物，例如在細胞之特定部分或細胞器中及/或多細胞生物體之(a)特定細胞、組織、器官或部分中定位。此類報導基因亦可表現為具有經編碼之胺基酸序列之蛋白質融合物。一些較佳但非限制性實例包括螢光蛋白質，諸如GFP。

【0069】 適合的啟動子、終止子及其他元件之一些較佳但非限制性實例包括可用於在本文中所提及之宿主細胞中表現之彼等啟動子、終止子及其他元件；且尤其適用於在細菌細胞中表現之彼等啟動子、終止子及其他元件，諸如本文中所提及之彼等啟動子、終止子及其他元件。對於啟動子、選擇標記、前導序列、表現標記及可存在/用於本文中所描述之基因構築體中之其他元件(諸如終止子、轉錄及/或轉譯強化子及/或整合因子)的一些(其他)非限制性實例，參考通用手冊，諸如Sambrook等人，「Molecular Cloning: A Laboratory Manual」(第2版)，第1-3卷，Cold Spring Harbor Laboratory Press (1989)；F. Ausubel等人，編，「Current protocols in molecular biology」，Green Publishing and Wiley Interscience, New York (1987)以及參考WO 95/07463、WO 96/23810、

WO 95/07463、WO 95/21191、WO 97/11094、WO 97/42320、WO 98/06737、WO 98/21355、US-A-7,207,410、US-A-5,693,492及EP 1 085 089中所給出之實例。亦參考上文所引用之一般背景技術及本文中所引用之其他參考文獻。

【0070】 用於產生本文所述之核苷酸序列、核酸及基因構築體之技術將為熟習此項技術者所清楚且可例如包括但不限於自動化DNA合成。本文中所描述之基因構築體一般亦可藉由將本文中所描述之核苷酸序列適合地連接至上文所描述之一或多種其他元件來提供。通常，本文中所描述之基因構築體將藉由將如本文中所描述核苷酸序列或核酸插入本身已知之適合(表現)載體中來獲得。此等及其他技術將為熟習此項技術者所清楚，且再次參考標準手冊，諸如上文所提及之Sambrook等人及Ausubel等人。

【0071】 本文中所描述之核酸及/或本文中所描述之基因構築體可用以轉型宿主細胞或宿主生物體，亦即用於表現及/或產生經編碼之(融合)蛋白質或多肽。適合的宿主或宿主細胞將為熟習此項技術者所清楚，且可例如為任何適合的真菌、原核或真核細胞或細胞株或任何適合的真菌、原核或真核生物體，例如：

-細菌菌株，包括但不限於革蘭氏陰性菌株(gram-negative strain)，諸如大腸桿菌(*Escherichia coli*)之菌株；變形桿菌屬(*Proteus*)之菌株，例如奇異變形桿菌(*Proteus mirabilis*)之菌株；假單胞菌之菌株，例如螢光假單胞菌(*Pseudomonas fluorescens*)之菌株；及革蘭氏陽性菌株(gram-positive strain)，諸如芽孢桿菌(*Bacillus*)之菌株，例如枯草桿菌之菌株或芽孢桿菌(*Bacillus brevis*)之菌株；鏈黴菌之菌株，例如青紫鏈黴菌之菌株；葡萄球菌屬之菌株，例如肉葡萄球菌(*Staphylococcus carnosus*)之菌

株；及乳球菌屬之菌株，例如雷特氏乳球菌(*Lactococcus lactis*)之菌株；

-真菌細胞，包括但不限於來自木黴菌屬(*Trichoderma*)之物種，例如來自里氏木菌(*Trichoderma reesei*)的細胞；例如來自紅黴菌(*Neurospora*)之物種，例如來自粗厚神經孢子菌(*Neurospora crassa*)的細胞；來自糞殼菌屬(*Sordaria*)之物種，例如來自大孢糞殼(*Sordaria macrospora*)的細胞；來自麴菌屬(*Aspergillus*)之物種，例如來自黑麴菌(*Aspergillus niger*)或來自醬油麴菌(*Aspergillus sojae*)的細胞；或來自其他絲狀真菌之物種的細胞；

-酵母細胞，包括但不限於來自酵母菌屬(*Saccharomyces*)，例如釀酒酵母之物種的細胞；裂殖酵母屬(*Schizosaccharomyces*)，例如裂殖酵母屬(*Schizosaccharomyces pombe*)之物種的細胞；畢赤酵母(*Pichia*)，例如嗜甲醇酵母(*Pichia pastoris*)或甲醇畢赤酵母(*Pichia methanolica*)之物種的細胞；漢森酵母(*Hansenula*)，例如多形漢森酵母(*Hansenula polymorpha*)之物種的細胞；克魯維酵母，例如乳酸克魯維酵母(*Kluyveromyces lactis*)之物種的細胞；*Arxula*，例如*Arxula adeninivorans*之物種的細胞；耶氏酵母(*Yarrowia*)，例如解脂耶氏酵母(*Yarrowia lipolytica*)之物種的細胞；

-兩棲動物細胞或細胞株，諸如非洲爪蟾卵母細胞(*Xenopus oocytes*)；

-源自昆蟲之細胞或細胞株，諸如源自鱗翅目之細胞/細胞株，包括但不限於夜蛾SF9及Sf21細胞或源自果蠅之細胞/細胞株，諸如施奈德(Schneider)及Kc細胞；

-植物或植物細胞，例如煙草植物中之植物或植物細胞；及/或

-哺乳動物細胞或細胞株，例如源自人類之細胞或細胞株，包括但不

限於CHO-細胞、BHK-細胞(例如BHK-21細胞)及人類細胞或細胞株，諸如HeLa、COS (例如COS-7)及PER.C6細胞；

-以及本身已知用於表現及產生抗體及抗體片段(包括但不限於(單)域抗體及ScFv片段)之所有其他宿主或宿主細胞，其係熟習此項技術者咸了解者。亦參考上文所引用之一般背景技術，以及例如WO 94/29457；WO 96/34103；WO 99/42077；Frenken 等人(1998, Res. Immunol. 149(6): 589-99)；Riechmann及Muyldermans (1999, J. Immunol. Methods, 231(1-2): 25-38)；van der Linden (2000, J. Biotechnol. 80(3): 261-70)；Joosten等人(2003, Microb. Cell Fact. 2(1): 1)；Joosten等人(2005, Appl. Microbiol. Biotechnol. 66(4): 384-92)；及本文中所引用之其他參考文獻。

【0072】 一些較佳表現宿主為嗜甲醇酵母及用於表現/產生治療蛋白質之人類細胞株。

【0073】 如本文中所使用之術語「GS連接子」一般係指由甘胺酸及絲胺酸殘基構成及/或基本上由甘胺酸及絲胺酸殘基組成之肽連接子。

【0074】 一般而言，此類GS連接子(以及本文中所提及之其他肽連接子)將含有至少5個胺基酸殘基，諸如約10個胺基酸殘基、約15個胺基酸殘基、約20個胺基酸殘基、約25個胺基酸殘基、約35個胺基酸殘基及高達50個或更多個胺基酸殘基(儘管如此，操作時仍經常使用包含約10至40個胺基酸殘基，諸如約15至約35個胺基酸殘基之連接子)。

【0075】 通常，與絲胺酸殘基之數目相比，此類連接子將含有過量甘胺酸殘基，例如每個絲胺酸殘基有3至6個之間的甘胺酸殘基。亦通常，此類連接子將含有序列基元之一或多個(諸如兩個或更多個)重複序列。此

外，儘管在本發明中在其最廣泛意義上不排除存在一或多個其他胺基酸(諸如改用麩胺酸殘基，或蘇胺酸殘基，而非絲胺酸殘基)，但本文中所使用之連接子較佳僅含有(或意欲僅含有)甘胺酸及絲胺酸殘基。

【0076】 如熟習此項技術者將清楚，在蛋白質工程改造之技術中最常用(且在本發明操作中亦較佳)的GS連接子為包含GGGGGS (SEQ ID NO:1)基元之一或多個重複序列的連接子，亦即通式(Gly-Gly-Gly-Gly-Ser)_n之連接子，其中n可為1、2、3、4、5、6、7或更大。一些實例為15GS連接子(n = 3)及35GS連接子(n = 7)。可參考例如：Chen等人, Adv Drug Deliv. Rev. 2013年10月15日; 65(10): 1357-1369；及Klein等人, Protein Eng. Des. Sel. (2014) 27 (10): 325-330。

【0077】 由本發明之GS連接子-編碼序列編碼之GS連接子可用於以合適方式將任何所需蛋白質、肽、肽部分、結合域或結合單元連接在一起，以便形成其中此類蛋白質、肽、肽部分、結合域或結合單元中之兩者或更多者係藉由一或多個GS連接子連接在一起的(融合)蛋白質或多肽。一般而言，且如熟習此項技術者將清楚，由本發明之GS連接子-編碼序列編碼之GS連接子可用於任何目的，此係由於該等GS連接子可用於及/或已用於先前技術中。本發明之GS連接子-編碼序列(及由其編碼之GS連接子)之此類用途及應用將為熟習此項技術者所清楚。

【0078】 在一個特定態樣中，由本發明之GS連接子-編碼序列編碼之GS連接子可適合地用於將兩個或更多個免疫球蛋白單一可變域(諸如兩個或更多個奈米抗體，例如VHH之奈米抗體、人類化VHH之奈米抗體、序列最佳化VHH之奈米抗體或駱駝化VH之奈米抗體，諸如駱駝化人類VH之奈米抗體)連接在一起，以形成二價、三價、雙特異性、三特異性、

雙互補位、四價或其他合適的ISVD構築體。舉例而言，參考Ablynx N.V.之各種申請案，諸如且不限於)WO 2004/062551、WO 2006/122825、WO 2008/020079及WO 2009/068627。GS連接子亦可例如用於將一或多個針對治療標靶之免疫球蛋白單一可變域或奈米抗體連接至提供增加之半衰期(例如增加之 $t_{1/2-\beta}$)之免疫球蛋白單一可變域或奈米抗體，諸如針對血清白蛋白之免疫球蛋白單一可變域或奈米抗體。同樣，在此等用途或應用中，本發明之GS連接子-編碼序列(及由其編碼之GS連接子)可以與編碼GS連接子之已知核苷酸序列基本上相同的方式使用。此類免疫球蛋白單一可變域或奈米抗體構築體之一些特定但非限制性實例示意性地顯示於表III中，且編碼此等構築體之核酸亦示意性地顯示於圖I中(表III之圖例適用)。熟習此項技術者基於本文中之揭示內容將清楚其他實例。

表II：

奈米抗體構築體之類型	結構(示意性地表示)
二價、單特異性	Nb ₁ -L ₁ -Nb ₁
二價、雙特異性	Nb ₁ -L ₁ -Nb ₂
二價、雙互補位	Nb _A -L ₁ -Nb _B
三價、單特異性	Nb ₁ -L ₁ -Nb ₁ -L ₂ -Nb ₁
三價、雙特異性	Nb ₁ -L ₁ -Nb ₁ -L ₂ -Nb ₂
三價、雙特異性	Nb ₁ -L ₁ -Nb ₂ -L ₂ -Nb ₁
三價、雙特異性	Nb ₂ -L ₁ -Nb ₁ -L ₂ -Nb ₁
三價、三特異性	Nb ₁ -L ₁ -Nb ₂ -L ₂ -Nb ₃
三價、雙互補位	Nb _A -L ₁ -Nb _B -L ₂ -Nb ₂
- 「Nb₁」係指第一奈米抗體。 - 「Nb₂」係指結合至不同於Nb₁所結合之目標之目標的第二奈米抗體(Nb₂可例如亦結合至血清白蛋白以便賦予經改良半衰期)。 - 「Nb_A」及「Nb_B」分別係指結合至同一靶標上之不同抗原決定基的第一奈米抗體及第二奈米抗體。 - 「L₁」係指第一GS連接子。 - 「L₂」係指第二GS連接子，其可與或可不與L₁相同。	

【0079】 現將藉助於以下非限制性較佳態樣、實例及圖式進一步描述本發明。

【0080】貫穿本申請案所引用之所有參考文獻(包括文獻參考、頒予之專利、公開之專利申請案及同在申請中之專利申請案)的全部內容以引用之方式明確地併入本文中，尤其對於上文所提及之教示內容。

【0081】

實驗部分

實例1：

構築四價奈米抗體構築體之表現載體

在此實例中，作為非限制性實例，本發明將使用由重鏈大羊駝之四個序列最佳化可變域組成的四價奈米抗體構築體說明，該等域與35GS連接子頭尾融合(參見圖2)。所用總構築體(在本文中亦被稱為「奈米抗體構築體A」)可由下式示意性地表示：

$$[A]-[35GS\text{連接子}]-[B]-[35GS\text{連接子}]-[C]-[35GS\text{連接子}]-[C]$$

其中[A]、[B]及[C]表示三種不同奈米抗體且[35GS連接子]表示35GS連接子(亦參見圖2)。

【0082】將含有奈米抗體構築體A之編碼資訊之DNA片段選殖至含有zeocin™抗性基因之畢赤酵母表現載體(由Näätäsaari等人, PLoS One. 2012;7(6):e39720所描述之原始pPpT4_Alpha_S表現載體之衍生物)的多個選殖位點中，使得Nanobody®序列在α交配因子(aMF)信號肽序列之下游且與其一致。

【0083】

編碼序列之奈米抗體構築體A之轉型、構築體在嗜甲醇酵母中之表現及分泌

轉型及表現研究執行於畢赤酵母菌株NRRL Y-11430 (ARS Patent

Culture Collection 1815 North University St., Peoria)中。此WT菌株用以製造過度表現內因性畢赤酵母輔助蛋白質KAR2 (基因ID:8198455)以及奈米抗體構築體A之衍生菌株。奈米抗體構築體A與Kar2兩者皆處於AOX1甲醇誘導性啟動子之控制下。轉型係藉由標準技術且根據標準手冊執行(參見例如*Methods In Molecular Biology* 2007, Humana Press Inc.)。如已在畢赤酵母協定中描述(再次參見標準手冊)，使轉型體生長於含有吉歐黴素(Zeocin)之選擇性培養基上且選擇多個個別菌落且評估奈米抗體構築體A於BMCM培養基中之5 mL搖晃燒瓶培養物中之表現程度且藉由添加甲醇誘導。在標準分批進料醱酵中使用最佳表現純系。執行甘油進料批次且藉由添加甲醇起始誘導。在2 L規模下在pH 6、30°C下於甲醇進料速率為4 ml/L*h之複合培養基中執行生產。

【0084】

在分批進料醱酵之後純化奈米抗體構築體A

如下純化奈米抗體構築體A：在醱酵之後，經由中空纖維750 kDa使細胞培養液之部分澄清，隨後使用CIEX Poros XS樹脂進行捕捉步驟，使用CIEX Nuvia HR-S樹脂進行拋光步驟且在AIEX Sartobind STIC PA上進行流通步驟。最後，使用Hydrosart 10 kD薄膜，經由UF/DF執行濃縮及緩衝液交換步驟。

【0085】

基於離子交換層析法對經純化奈米抗體構築體A之分析及酸性變體之分子量的測定

藉由強陽離子交換層析，使用pH梯度(pH-IEX)分析經純化奈米抗體構築體A。圖5中所示之層析圖顯示溶離為相對於主峰之預峰之群組的

Nanobody® A之酸性變體。在酸性峰及主峰溶離份收集之後，藉由用電噴霧Q-TOF質譜分析測定其分子量來研究酸性變體之性質。解卷積質譜顯示於圖6中。在酸性溶離份中所觀測到之主要質量為59689.4 Da，其比如在pH-IEX主峰部分中所量測之奈米抗體構築體A之質量高58道爾頓。針對主峰溶離份(59630.9 Da)中之奈米抗體構築體A所量測之質量比奈米抗體構築體A之理論分子量高12 ppm，亦即在儀器之量測誤差內。

【0086】 58道爾頓質量差可藉由甘胺酸與酸性胺基酸天冬胺酸之交換來解釋。

【0087】

藉由與質譜分析(RP-UHPLC-MS)聯合之肽圖譜反相UHPLC分析及鑑別酸性變體

奈米抗體構築體A之酸性變體溶離份之肽圖分析(在胰蛋白酶消化物之後)產生對具有58道爾頓之質量增量之兩種肽的鑑別。如圖2中示意性地顯示，此等兩種肽中之一者(在本文中被稱為「T10肽」)對應於涵蓋構築體中之第一奈米抗體的少數C端胺基酸殘基、第一35Gs連接子及構築體中之第二奈米抗體的少數N端胺基酸殘基之序列的一部分。T10肽之胺基酸序列(SEQ ID NO:10)及核苷酸序列(SEQ ID NO:11)顯示於圖3中。

【0088】 由於質譜儀中之碰撞誘導之片段化僅導致T10肽之部分序列覆蓋，因此胰蛋白酶消化物之T10肽係藉由逆相層析分級分離，且隨後用酶Asp-N消化。酶Asp-N為水解天冬胺酸殘基之N端側上之肽鍵的內切蛋白酶。因為天冬胺酸殘基不在此肽之序列中，因此僅在Gly-> Asp錯誤配對事件之情況下預期裂解。在藉由RP-UHPLC-MS之對T10肽之Asp-N消化物的分析中，用對應於質量增量為58道爾頓之T10肽之片段的質量鑑

別不同片段。如圖7中所示，總共鑑別9個Asp-N片段化位點。非常出乎意料地，觀測到，儘管僅在GGC密碼子處(亦參見圖3)且不在GGT密碼子處發生Asp錯誤配對，但兩個甘胺酸密碼子原則上皆可由天冬胺酸tRNA (具有反密碼子CUG及CUA)錯誤讀取。在兩種情況下皆存在G-(mRNA)/U-(tRNA)錯配，亦即轉譯期間最常見之錯配，連同可引起胺基酸錯配之搖擺位置錯配(C/U及/或U/U)。因此，更一般而言，根據本發明，當編碼除GGA或GGG (亦即非GGA或GGG)以外的甘胺酸之密碼子存在於本發明之核苷酸序列中時，可較佳的係密碼子為GGT或GGU，而非GGC。

【0089】 如所提及，對奈米抗體構築體A之肽圖譜分析亦產生對質量增量為58道爾頓之第二肽的鑑別。發現此肽對應於存在於奈米抗體構築體A中之奈米抗體中之一者的CDR中之一者。進一步分析(資料未示出)證實，亦針對此肽，所觀測到之58道爾頓質量增量最可能歸因於Asp錯誤配對。

【0090】

實例2：35GS連接子之核酸序列中之密碼子最佳化

存在於奈米抗體構築體A之35GS連接子序列中之GGC密碼子序列經GGG、GGA或GGT密碼子序列置換。

【0091】 所獲得之奈米抗體構築體表現於畢赤酵母菌株NRRL Y-11430中且如上文所描述地純化。藉由與上文所描述之相同方法量測所獲得之多肽中Asp錯誤配對程度。設置質譜儀以定量9個錯誤配對位點中之3個。

【0092】 在用參考奈米抗體構築體A獲得之多肽的35GS連接子(無密碼子最佳化)中Asp錯誤配對及用密碼子最佳化奈米抗體構築體A獲得之

多肽的35GS連接子中Asp錯誤配對之相對程度顯示於圖8中。

【0093】

實例3：對其他連接子中Asp錯誤配對之觀測

在此實例中，研究奈米抗體價數及連接子長度對Gly與Asp錯誤配對之影響。為此，產生各自具有9GS、20GS或35GS連接子序列及奈米抗體建構嵌段序列(不同於存在於奈米抗體構築體A中之奈米抗體建構嵌段序列)之二價、三價及四價構築體。亦產生無任何GGC密碼子之額外四價、35GS連接子奈米抗體構築體。十個新構築體顯示於圖9中。9GS連接子含有2個GGC密碼子，20GS連接子含有5個GGC密碼子且35GS連接子含有9個GGC密碼子。

【0094】 Gly與Asp錯誤配對之後的各可能的新肽用如上文所描述之質譜分析方法追蹤。進一步最佳化該方法以允許同時定量全部9個Asp-N片段化位點。錯誤配對之結果顯示於圖10 (9GS連接子)、圖11 (20 GS連接子)及圖12 (35 GS連接子)中。

【0095】 由此等結果可推斷，價數或連接子長度對Gly與Asp錯誤配對程度不具有影響。移除或減少GGC密碼子之數目明確降低Gly與Asp錯誤配對程度。

【0096】 最後，儘管在本文中主要在GS連接子方面描述本發明，但熟習此項技術者應清楚，本發明一般可應用於含有甘胺酸殘基之其他肽連接子。

【0097】 因此，在另一態樣中，本發明係關於一種編碼肽連接子之核苷酸序列及/或核酸，其中由核苷酸序列及/或核酸編碼之肽連接子含有四個或更多個甘胺酸殘基，其中超過70%、較佳超過85%、更佳超過

90%，諸如超過95%及高達99%及更多(包括100%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGA、GGG或GGT/GGU。

【0098】 在此態樣中，本發明亦關於一種編碼肽連接子之核苷酸序列及/或核酸，其中由核苷酸序列及/或核酸編碼之肽連接子含有四個或更多個甘胺酸殘基，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGA或GGG。

【0099】 在此態樣中，本發明亦關於一種編碼肽連接子之核苷酸序列及/或核酸，其中由核苷酸序列及/或核酸編碼之肽連接子含有四個或更多個甘胺酸殘基，其中少於30%、較佳少於1%、更佳少於10%，諸如少於5%及高達小於1%及更低(包括0%)之編碼GS連接子中之甘胺酸殘基的密碼子為GGC。

【序列表】

<110> 比利時商艾伯林克斯公司

<120> 編碼肽連接子之經改良核苷酸序列

<130> P18-001-PCT-1

<150> US 62/634,985

<151> 2018-02-26

<160> 33

<170> PatentIn version 3.5

<210> 1

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> 連接子序列

<400> 1

Gly Gly Gly Gly Ser

1 5

<210> 2

<211> 39

<212> DNA

<213> 人工序列

<220>

<223> 連接子序列

<400> 2

ggaggaggag gacggaggag gaggacggag gaggaggac

39

<210> 3

<211> 42

<212> DNA

<213> 人工序列

<220>

<223> 連接子序列

<400> 3

gggggggggg ggccgggggg ggggggccgg gggggggggg cc

42

<210> 4

<211> 42

<212> DNA

<213> 人工序列

<220>

<223> 連接子序列

<400> 4

ggaggggggg ggcaggaggg ggaggggcagg agggggaggg ca

42

<210> 5

<211> 41

<212> DNA	
<213> 人工序列	
<220>	
<223> 連接子序列	
<400> 5	
ggaggaggag gacggggggg gggggcggga ggaggaggac a	41
<210> 6	
<211> 98	
<212> DNA	
<213> 人工序列	
<220>	
<223> 連接子序列	
<400> 6	
ggaggaggag gaagggaggga ggaaggaggg aggaggaggga agggaggagg aggaaggagg	60
gaggaggagg ggaggaggag gaggaggagg ggaaggagg	98
<210> 7	
<211> 98	
<212> DNA	
<213> 人工序列	
<220>	
<223> 連接子序列	
<400> 7	
gggggggggg ggaggggggg ggggggaggg gggggggggg aggggggggg ggggaggggg	60
ggggggggag gggggggggg ggaggggggg gggggggg	98
<210> 8	
<211> 105	
<212> DNA	
<213> 人工序列	
<220>	
<223> 連接子序列	
<400> 8	
ggagggggag ggaagcggagg gggaggaggc ggaaggaggg ggaagcggagg ggaaggaggc	60
ggaaggaggg ggaagcggagg gggaggaggc ggaaggaggg ggaagc	105
<210> 9	
<211> 97	
<212> DNA	
<213> 人工序列	
<220>	
<223> 連接子序列	
<400> 9	
ggaggaggag gacggggggg gggggccgga ggaggaggac aggggggggg ggaggggggg	60
aggggggagg gggggggggg gagcggaggga ggaggagc	97
<210> 10	

<211> 9
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 10

Val Val Gln Pro Gly Gly Ser Leu Arg
1 5

<210> 11
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 肽片段

<400> 11
gtggtagcagc ctgggggctc tctgaga

27

<210> 12
<211> 35
<212> PRT
<213> 人工序列

<220>
<223> 連接子序列

<400> 12

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
20 25 30

Gly Gly Ser
35

<210> 13
<211> 105
<212> DNA
<213> 人工序列

<220>
<223> 連接子序列

<400> 13
ggaggcggtg ggtaggttg cggaggcagc ggtaggaggag gtagtggcgg tggcggtagt

60

gggggtggag gcagcggagg cggaggcagt gggggcggtg gatca

105

<210> 14
<211> 105
<212> DNA
<213> 人工序列

<220>

<223> 連接子序列

<400> 14
ggaggcggtag ggtcaggtagg cggaggcagc ggtggaggag gtagtggcgg tggcggtagt 60
gggggtggag gcagcggagg cggaggcagt gggggcggtg gatcc 105

<210> 15
<211> 105
<212> DNA
<213> 人工序列

<220>
<223> 連接子序列

<400> 15
ggaggcggtag ggtcaggtagg cggaggcagc ggtggaggag gtagtggcgg tggcggtagt 60
gggggtggag gcagcggagg cggaggcagt gggggcggtg gctca 105

<210> 16
<211> 20
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 16
Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Gly Gly Gly
1 5 10 15
Ser Gly Gly Gly
20

<210> 17
<211> 46
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 17
Asp Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
1 5 10 15
Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Val Gln Leu Val
20 25 30
Glu Ser Gly Gly Gly Val Val Gln Pro Gly Gly Ser Leu Arg
35 40 45

<210> 18
<211> 18
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 18

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Gly Gly Gly
1 5 10 15

Ser Gly

<210> 19
<211> 48
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 19

Asp Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
1 5 10 15

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Val Gln
20 25 30

Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Gly Ser Leu Arg
35 40 45

<210> 20
<211> 29
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 20

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Gly Gly Gly
1 5 10 15

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
20 25

<210> 21
<211> 37
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 21

Asp Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
1 5 10 15

Gly Ser Glu Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro
20 25 30

Gly Gly Ser Leu Arg
35

<210> 22
<211> 35
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 22

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Gly Gly Gly
1 5 10 15

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25 30

Gly Gly Gly
35

<210> 23
<211> 31
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 23

Asp Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Val Gln Leu
1 5 10 15

Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Gly Ser Leu Arg
20 25 30

<210> 24
<211> 38
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 24

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Gly Gly Gly
1 5 10 15

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25 30

Gly Gly Gly Gly Ser Gly

35

<210> 25
<211> 28
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 25

Asp Gly Gly Ser Gly Gly Gly Gly Ser Glu Val Gln Leu Val Glu Ser
1 5 10 15

Gly Gly Gly Val Val Gln Pro Gly Gly Ser Leu Arg
20 25

<210> 26
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 26

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly
1 5 10

<210> 27
<211> 53
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 27

Asp Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
1 5 10 15

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
20 25 30

Gly Ser Glu Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro
35 40 45

Gly Gly Ser Leu Arg
50

<210> 28
<211> 27
<212> PRT
<213> 人工序列

<220>

<223> 肽片段

<400> 28

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Gly Gly Gly
1 5 10 15
Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25

<210> 29
<211> 39
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 29

Asp Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15
Gly Gly Gly Ser Glu Val Gln Leu Val Glu Ser Gly Gly Gly Val Val
20 25 30
Gln Pro Gly Gly Ser Leu Arg
35

<210> 30
<211> 40
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 30

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Gly Gly Gly
1 5 10 15
Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25 30
Gly Gly Gly Gly Ser Gly Gly Gly
35 40

<210> 31
<211> 26
<212> PRT
<213> 人工序列

<220>
<223> 肽片段

<400> 31

Asp Ser Gly Gly Gly Gly Ser Glu Val Gln Leu Val Glu Ser Gly Gly

1

5

10

15

Gly Val Val Gln Pro Gly Gly Ser Leu Arg

2025

<210> 32

<211> 43

<212> PRT

<213> 人工序列

<220>

<223> 肽片段

<400> 32

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly Gly Gly Gly

151015

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

202530

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

3540

<210> 33

<211> 23

<212> PRT

<213> 人工序列

<220>

<223> 肽片段

<400> 33

Asp Gly Gly Ser Glu Val Gln Leu Val Glu Ser Gly Gly Gly Val Val

151015

Gln Pro Gly Gly Ser Leu Arg

20



202000238

【發明摘要】

【中文發明名稱】

編碼肽連接子之經改良核苷酸序列

【英文發明名稱】

IMPROVED NUCLEOTIDE SEQUENCES ENCODING PEPTIDE
LINKERS

【中文】

本發明提供編碼甘胺酸絲胺酸連接子且使用過量GGA、GGG及GGT/GGU密碼子以編碼該等甘胺酸殘基之經改良核苷酸序列及核酸。本發明進一步關於編碼包含甘胺酸絲胺酸連接子之(融合)蛋白質及多肽之核苷酸序列及核酸，該等核苷酸序列及核酸包含本發明此類經改良之核苷酸序列及核酸。

【英文】

The invention provides improved nucleotide sequences and nucleic acids that encode glycine serine linkers and that use an excess of GGA, GGG, and GGT/GGU codons to encode the glycine residues. The invention further relates to nucleotide sequences and nucleic acids that encode (fusion) proteins and polypeptides comprising glycine serine linkers, which nucleotide sequences and nucleic acids comprise such improved nucleotide sequences and nucleic acids of the invention.

【指定代表圖】

圖1

【代表圖之符號簡單說明】

無

【發明申請專利範圍】

【第1項】

一種編碼肽連接子之核苷酸序列及/或核酸，其中由該核苷酸序列或核酸編碼之該肽連接子包含甘胺酸及絲胺酸殘基或(基本上)由甘胺酸及絲胺酸殘基組成，其中：

超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGA、GGG或GGT/GGU；

超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%及更多(包括100%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGA或GGG；及/或

少於30%、較佳少於15%、更佳少於10%，諸如少於5%及至多少於1%及更低(包括0%)之編碼該肽連接子中之甘胺酸殘基的密碼子為GGC。

【第2項】

如請求項1之核苷酸序列及/或核酸，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及至多99%或更多(包括100%)之編碼該肽連接子中之甘胺酸殘基的該等密碼子為GGA、GGG或GGT/GGU。

【第3項】

如請求項1或2中任一項之核苷酸序列及/或核酸，其中超過70%、較佳超過85%、更佳超過90%，諸如超過95%及高達99%或更高(包括100%)之編碼該肽連接子中之甘胺酸殘基的該等密碼子為GGA或GGG。

【第4項】

如請求項1至3中任一項之核苷酸序列及/或核酸，其中少於30%、較

佳少於15%、更佳少於10%，諸如少於5%及至多少於1%或更低(包括0%)之編碼該肽連接子中之甘胺酸殘基的該等密碼子為GGC。

【第5項】

如請求項1至4中任一項之核苷酸序列及/或核酸，其中該肽連接子包含序列基元GGGGS (SEQ ID NO:1)之一或多個(諸如兩個或更多個)重複序列或(基本上)由其組成。

【第6項】

如請求項1至5中任一項之核苷酸序列及/或核酸，其中該肽連接子為9 GS連接子、15 GS連接子、20 GS連接子或35 GS連接子。

【第7項】

如請求項6之核苷酸序列及/或核酸，其中該肽連接子為35 GS連接子。

【第8項】

一種編碼(融合)蛋白質或融合多肽之核苷酸序列及/或核酸，其中由該核苷酸序列及/或核酸編碼之該融合蛋白質或多肽包含經一或多個肽連接子適當連接之兩個或更多個肽部分，其中該一或多個肽連接子係由如請求項1至7中任一項之核苷酸序列或核酸編碼。

【第9項】

如請求項8之核苷酸序列及/或核酸，其中該等兩個或更多個肽部分均為免疫球蛋白單一可變域。

【第10項】

如請求項9之核苷酸序列及/或核酸，其中該等兩個或更多個肽部分均為VHH之肽部分、人類化VHH之肽部分、序列最佳化VHH之肽部分或駱

駝化VH之肽部分，諸如駱駝化人類VH之肽部分。

【第11項】

如請求項8至10中任一項之核苷酸序列及/或核酸，其編碼二價、三價、雙特異性、三特異性、雙互補位或四價構築體。

【第12項】

一種基因構築體，其包含如請求項1至11中任一項之核苷酸序列及/或核酸。

【第13項】

一種用於表現或產生(融合)蛋白質或多肽之方法，其中該方法至少包含在適合的宿主細胞或宿主生物體中表現如請求項8至11中任一項之核苷酸序列或核酸之步驟，且視情況亦包含分離/純化因此表現之該(融合)蛋白質或多肽之步驟。

【第14項】

如請求項13之表現或產生(融合)蛋白質或多肽之方法，其中該宿主為畢赤酵母(*Pichia*)，諸如嗜甲醇酵母(*Pichia pastoris*)。

【第15項】

如請求項13之表現或產生(融合)蛋白質或多肽之方法，其中該宿主為哺乳動物細胞，諸如中國倉鼠卵巢(CHO)細胞。

【第16項】

一種宿主細胞或宿主生物體，其包含如請求項8至11中任一項之編碼(融合)蛋白質或融合多肽的核苷酸序列及/或核酸。

【第17項】

一種降低肽連接子中Gly與Asp錯誤配對(misincorporation)程度的方

法，該方法包含在編碼該肽連接子之核酸序列及/或核酸中用GGG、GGA或GGT/GGU密碼子置換至少一個GGC密碼子之步驟。

【第18項】

如請求項17之降低肽連接子中Gly與Asp錯誤配對程度的方法，其中該至少一個GGC密碼子經GGG或GGA密碼子置換。

【第19項】

如請求項17或18中任一項之降低肽連接子中Gly與Asp錯誤配對程度的方法，其中該肽連接子包含序列基元GGGGS (SEQ ID NO:1)之一或多個(諸如兩個或更多個)重複序列或(基本上)由其組成。

【第20項】

如請求項17至19中任一項之降低肽連接子中Gly與Asp錯誤配對程度的方法，其中該肽連接子為9 GS連接子、15 GS連接子、20 GS連接子或35 GS連接子。

【第21項】

如請求項17至20中任一項之降低肽連接子中Gly與Asp錯誤配對程度的方法，其中該肽連接子為35 GS連接子。

【第22項】

如請求項17至21中任一項之降低肽連接子中Gly與Asp錯誤配對程度的方法，其中該肽連接子連接兩個或更多個肽部分。

【第23項】

如請求項22之降低肽連接子中Gly與Asp錯誤配對程度的方法，其中該等肽部分為免疫球蛋白單一可變域。

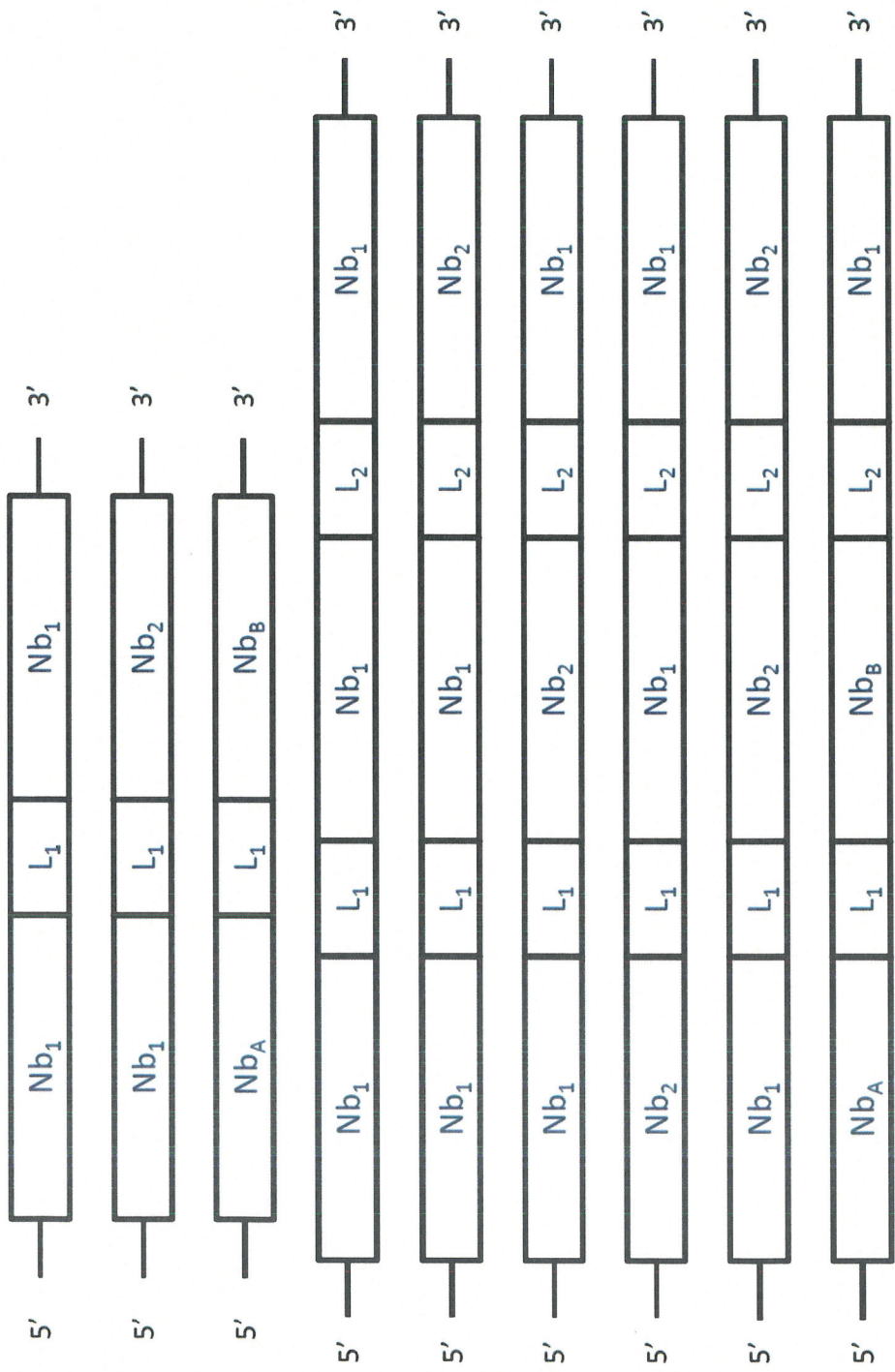
【第24項】

如請求項23之降低肽連接子中Gly與Asp錯誤配對程度的方法，其中該等肽部分為VHH之肽部分、人類化VHH之肽部分、序列最佳化VHH之肽部分或駱駝化VH之肽部分，諸如駱駝化人類VH之肽部分。

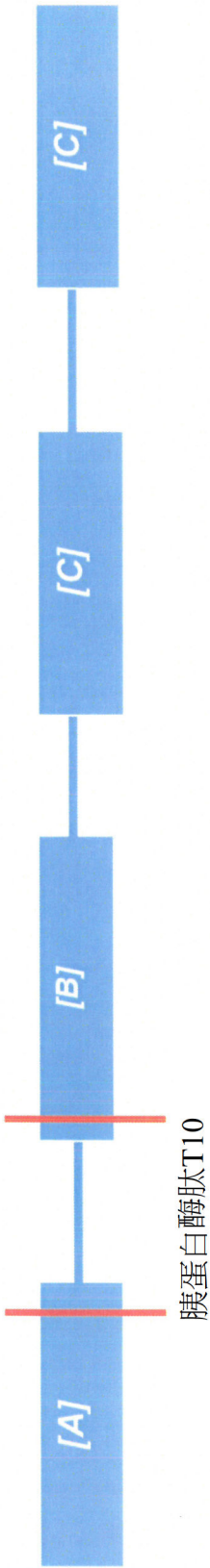
【第25項】

如請求項22至24中任一項之降低肽連接子中Gly與Asp錯誤配對程度的方法，其中該肽連接子包含於二價、三價、雙特異性、三特異性、雙互補位或四價構築體中。

【發明圖式】



【圖1】



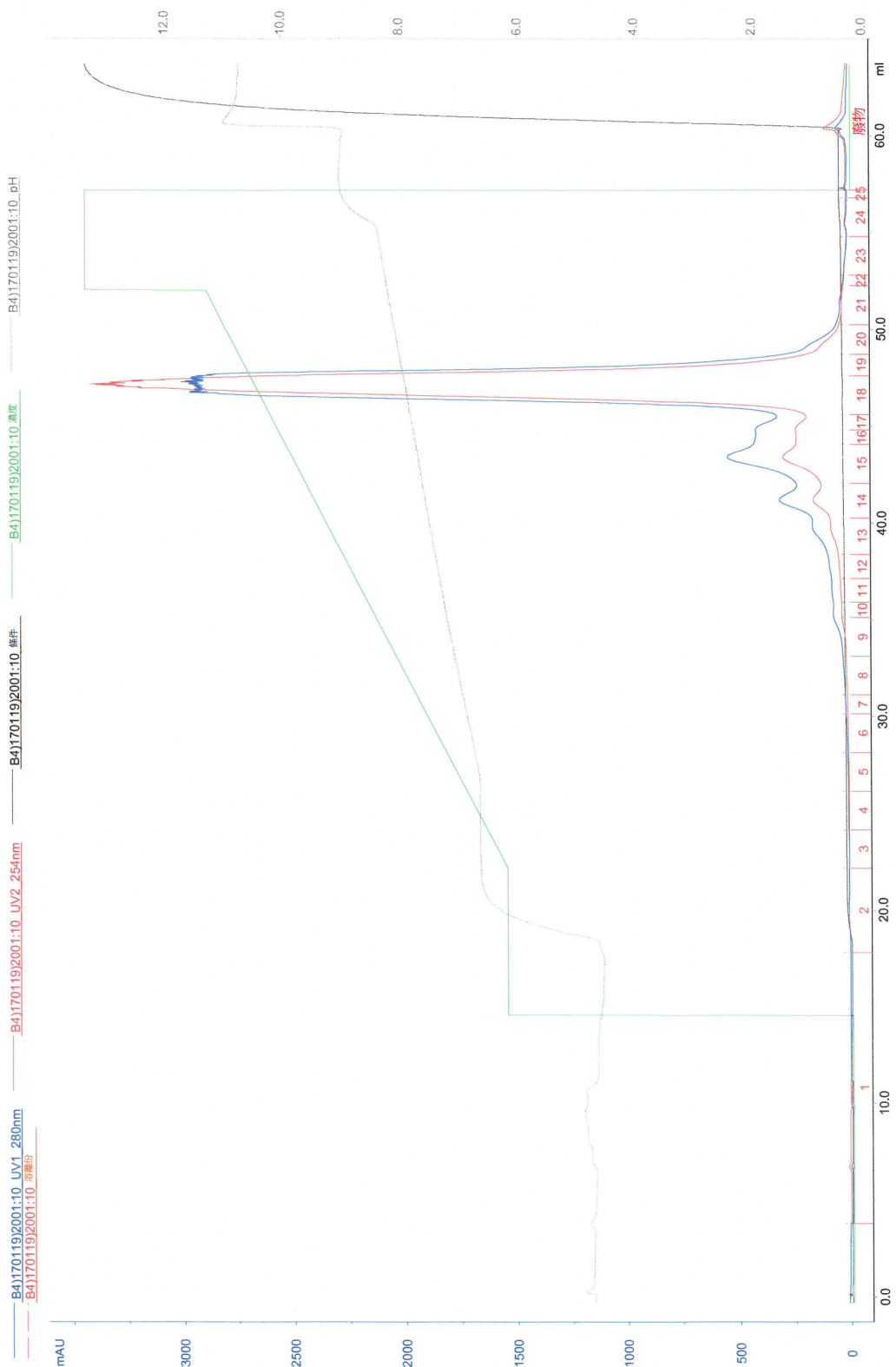
【圖2】

Y W <u>G</u> Q G T L V T V S S G <u>G</u> G S G <u>G</u>
tac-tgg-ggc-cag-ggg-acc-ctg-gtc-acg-gtc-tcc-tcc-gga-ggc-ggt-ggg-tca-ggt-ggc
G <u>G</u> S G G G S <u>G</u> G S <u>G</u> G S G G <u>G</u> S G
gga-ggc-agc-ggt-gga-gga-ggt-agt-ggc-ggt-ggt-agt-ggg-ggt-gga-ggc-agc-gga
<u>G</u> G <u>G</u> S G <u>G</u> G S E V Q L V E S G G G
ggc-gga-ggc-agt-ggg-ggc-ggt-gga-tca-gag-gtg-cag-ttg-gtg-gag-tct-ggg-gga-gga
V V Q P G <u>G</u> S L R [SEQ ID NO:10]
gtg-gtg-cag-cct-ggg-ggc-tct-ctg-aga [SEQ ID NO:11]

【圖3】

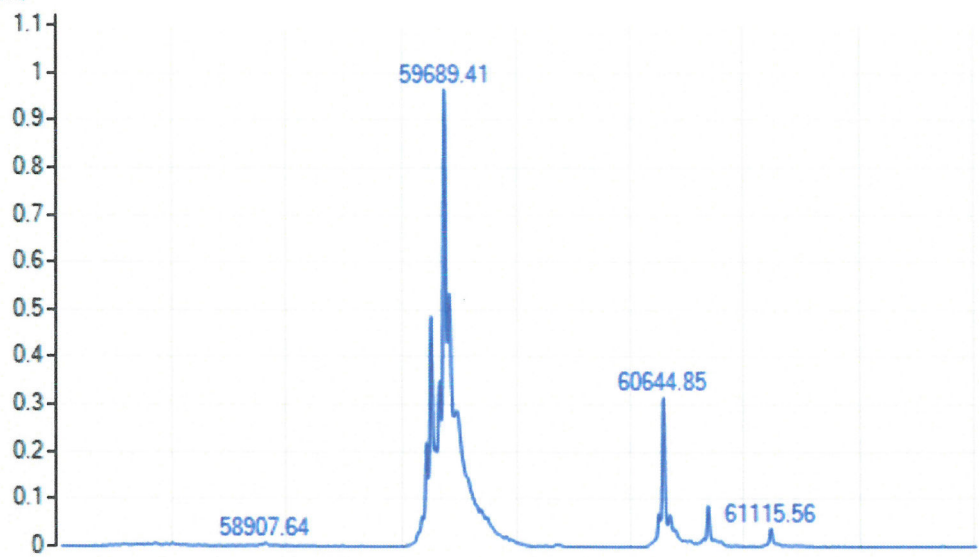
35GS連接子： GGGSGGGSGGGSGGGSGGGSGGGSGGGGS [SEQ ID NO:12]
第一35GS連接子編碼序列： GGA GGC <u>GGT</u> GGG tca <u>GGT</u> <u>GGC</u> GGA <u>GGC</u> agc <u>GGT</u> GGA GGA <u>GGT</u> agt <u>GGC</u> <u>GGT</u> <u>GGC</u> <u>GGT</u> agt GGG <u>GGT</u> GGA <u>GGC</u> agc GGA <u>GGC</u> GGA <u>GGC</u> agt GGG <u>GGC</u> <u>GGT</u> GGA tca [SEQ ID NO:13]
第二35GS連接子編碼序列： GGA <u>GGC</u> <u>GGT</u> GGG tca <u>GGT</u> <u>GGC</u> GGA <u>GGC</u> agc <u>GGT</u> GGA GGA <u>GGT</u> agt <u>GGC</u> <u>GGT</u> <u>GGC</u> <u>GGT</u> agt GGG <u>GGT</u> GGA <u>GGC</u> agc GGA <u>GGC</u> GGA <u>GGC</u> agt GGG <u>GGC</u> <u>GGT</u> GGA tcc [SEQ ID NO:14]
第三35GS連接子編碼序列 GGA <u>GGC</u> <u>GGT</u> GGG tca <u>GGT</u> <u>GGC</u> GGA <u>GGC</u> agc <u>GGT</u> GGA GGA <u>GGT</u> agt <u>GGC</u> <u>GGT</u> <u>GGC</u> <u>GGT</u> agt GGG <u>GGT</u> GGA <u>GGC</u> agc GGA <u>GGC</u> GGA <u>GGC</u> agt GGG <u>GGC</u> <u>GGT</u> <u>GGC</u> tca [SEQ ID NO:15]

【圖4】

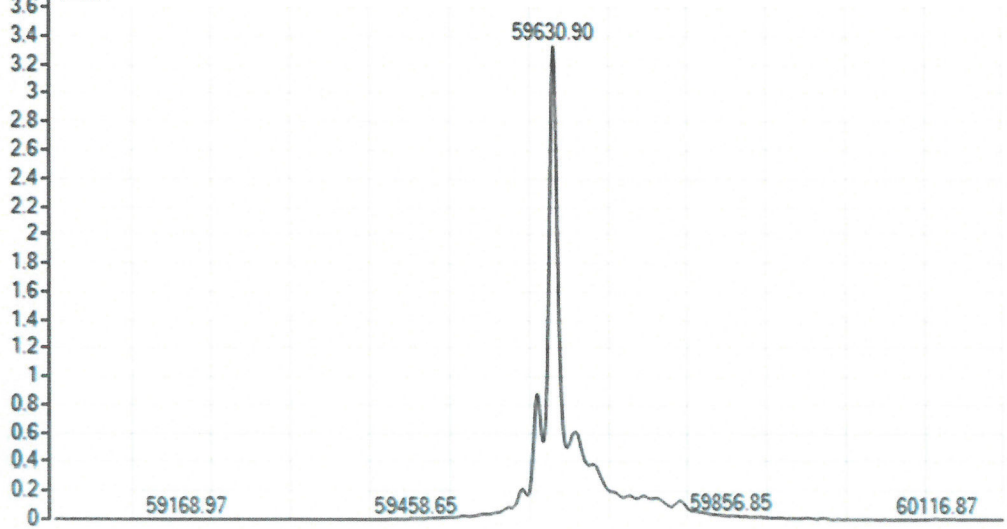


【圖5】

x10⁶ +ESI掃描(rt:2.195-2.595 min , 25次掃描) 片段化=200.0 V LMSC0101T0758.d 解卷積



x10⁶ +ESI掃描(rt:2.2.9-2.593 min , 24次掃描) 片段化=200.0 V LMSC0101T0755.d 解卷積(同位素)



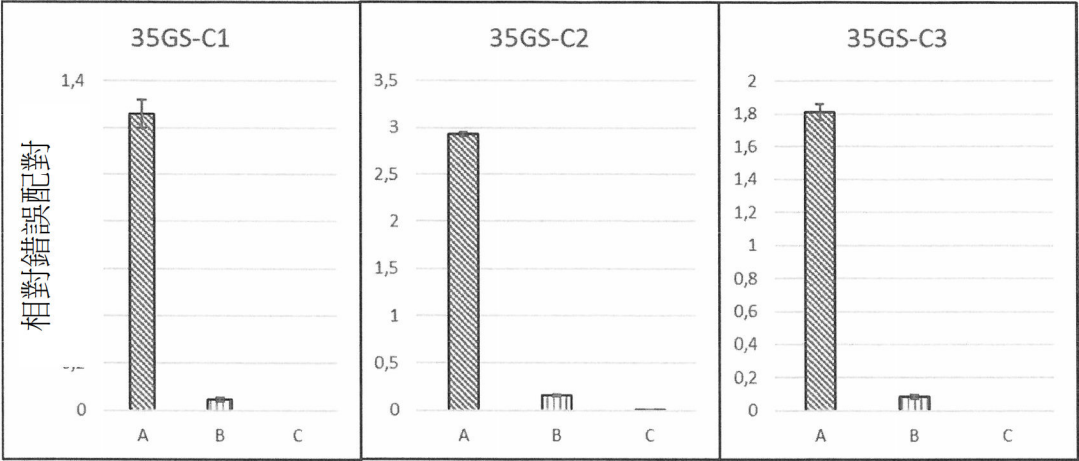
【圖6】

所量測之 質量	肽(AA開始 - AA結束)	序列
1782.8	T10 (110- 129)	YWGQGTLVTVSSGGGGSGGG [SEQ ID NO:16]
3644.7	T10 (130- 175)	DSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLVESGGGVVQP GGSLR [SEQ ID NO:17]
1668.8	T10 (110- 127)	YWGQGTLVTVSSGGGGSG [SEQ ID NO:18]
3758.7	T10 (128- 175)	DGGSGGGSGGGGSGGGGSGGGGSGGGGSEVQLVESGGGV VQPGGSLR [SEQ ID NO:19]
2356.0	T10 (110- 138)	YWGQGTLVTVSSGGGGSGGGGSGGGGSGG [SEQ ID NO:20]
3071.4	T10 (139- 175)	DGSGGGSGGGGSGGGGSEVQLVESGGGVVQPGGSLR [SEQ ID NO:21]
2728.2	T10 (110- 144)	YWGQGTLVTVSSGGGGSGGGGSGGGGSGGGGSGGG [SEQ ID NO:22]
2699.3	T10 (145- 175)	DSGGGGSGGGGSEVQLVESGGGVVQPGGSLR [SEQ ID NO:23]
2929.2	T10 (110- 147)	YWGQGTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSG [SEQ ID NO:24]
2498.2	T10 (148- 175)	DGGSGGGGSEVQLVESGGGVVQPGGSLR [SEQ ID NO:25]
1353.7	T10 (110- 122)	YWGQGTLVTVSSG [SEQ ID NO:26]
4073.8	T10 (123- 175)	DGGSGGGSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLVES GGGVVQPGGSLR [SEQ ID NO:27]
2242	T10 (110- 136)	YWGQGTLVTVSSGGGGSGGGGSGGGGS [SEQ ID NO:28]
3185.5	T10 (137- 175)	DGGSGGGSGGGGSGGGGSEVQLVESGGGVVQPGGSLR [SEQ ID NO:29]

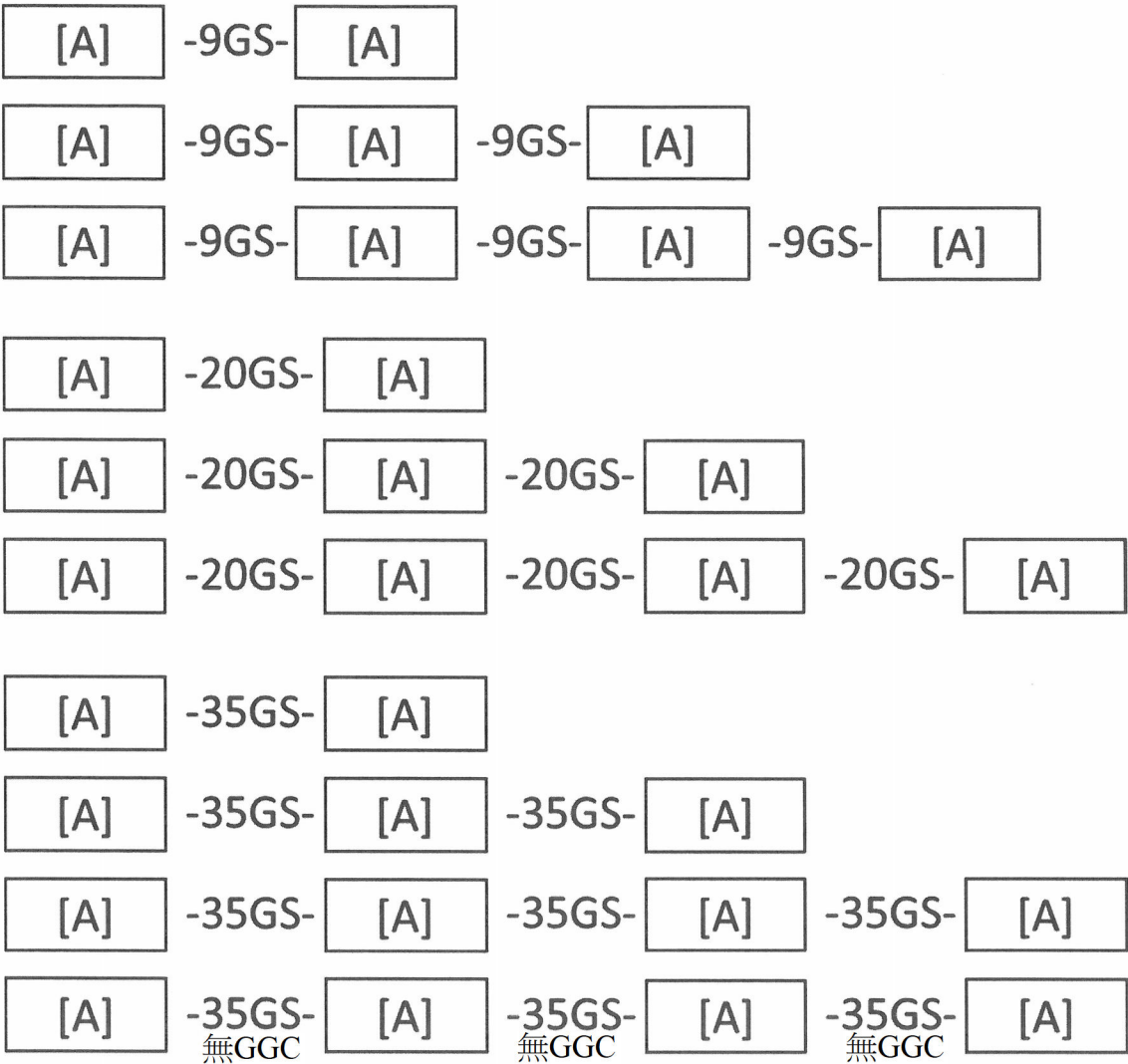
【圖7】

所量測之 質量	肽(AA開 始 - AA結 束)	序列
3043.3	T10 (110- 149)	YWGQGTLVTVSSGGGSGGGGSGGGGSGGGGSGGGGSGGG [SEQ ID NO:30]
2384.2	T10 (150- 175)	DSGGGGSEVQLVESGGGVVQPGGSLR [SEQ ID NO:31]
3244.4	T10 (110- 152)	YWGQGTLVTVSSGGGSGGGGSGGGGSGGGGSGGGGSGGG GSG [SEQ ID NO:32]
2183.1	T10 (153- 175)	DGGSEVQLVESGGGVVQPGGSLR [SEQ ID NO:33]

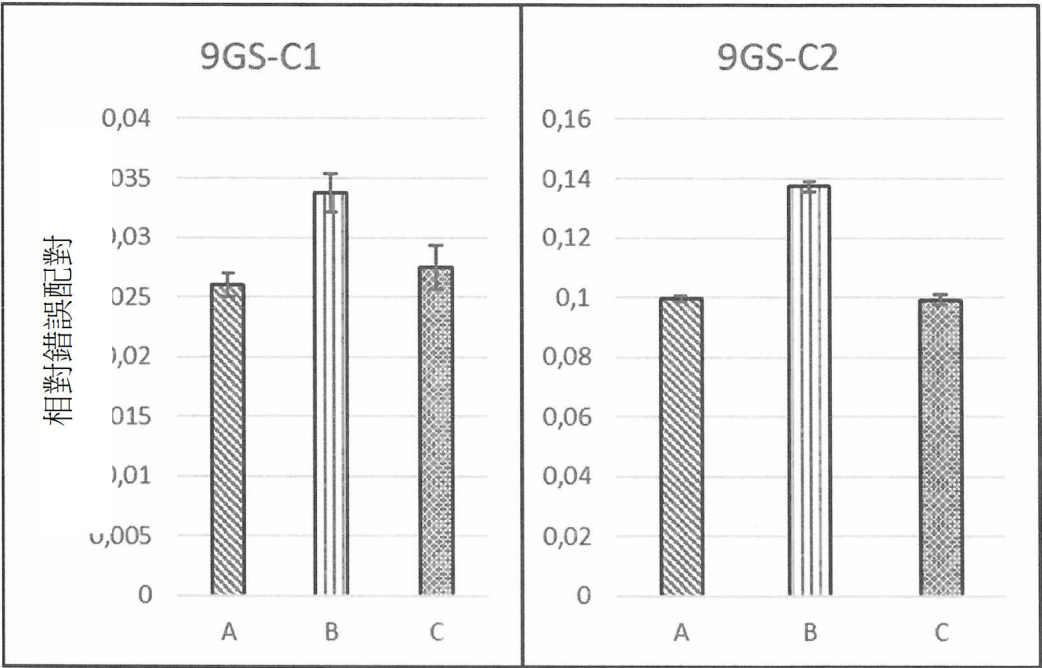
【圖7】(續)



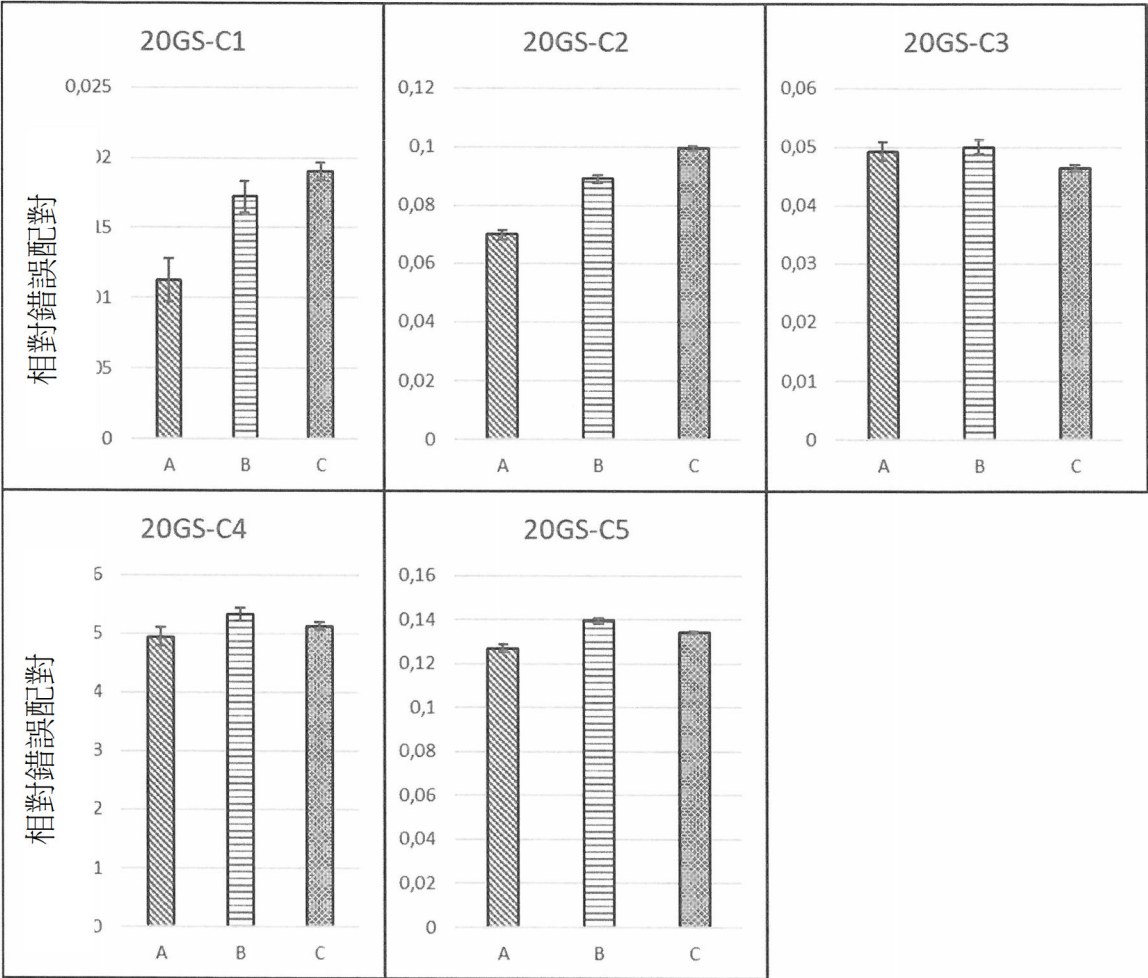
【圖8】



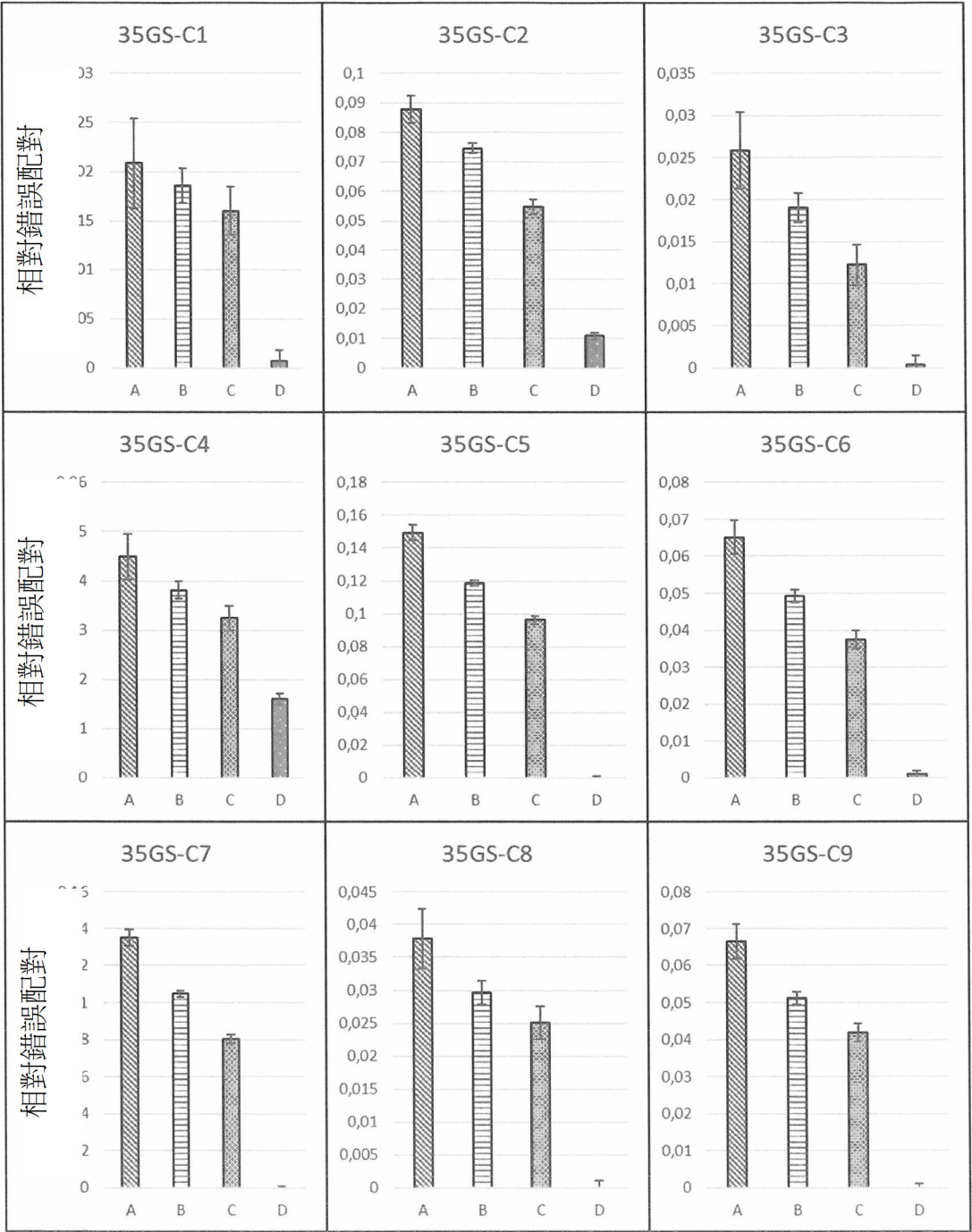
【圖9】



【圖10】



【圖11】



【圖12】