

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6814132号  
(P6814132)

(45) 発行日 令和3年1月13日(2021.1.13)

(24) 登録日 令和2年12月22日(2020.12.22)

(51) Int.Cl.

F I

C O 7 K 19/00 (2006.01)

C O 7 K 19/00 Z N A

C 1 2 N 15/12 (2006.01)

C 1 2 N 15/12

C 1 2 Q 1/02 (2006.01)

C 1 2 Q 1/02

C 1 2 Q 1/66 (2006.01)

C 1 2 Q 1/66

G O 1 N 33/532 (2006.01)

G O 1 N 33/532

B

請求項の数 9 (全 94 頁) 最終頁に続く

(21) 出願番号 特願2017-514354 (P2017-514354)  
 (86) (22) 出願日 平成27年9月11日 (2015.9.11)  
 (65) 公表番号 特表2017-533181 (P2017-533181A)  
 (43) 公表日 平成29年11月9日 (2017.11.9)  
 (86) 国際出願番号 PCT/US2015/049741  
 (87) 国際公開番号 W02016/040835  
 (87) 国際公開日 平成28年3月17日 (2016.3.17)  
 審査請求日 平成30年9月4日 (2018.9.4)  
 (31) 優先権主張番号 62/049,875  
 (32) 優先日 平成26年9月12日 (2014.9.12)  
 (33) 優先権主張国・地域又は機関  
 米国 (US)

前置審査

(73) 特許権者 593089149  
 プロメガ コーポレーション  
 Promega Corporation  
 アメリカ合衆国 ウィスコンシン 537  
 11-5399 マディソン ウッズ ホ  
 ロー ロード 2800番地  
 (74) 代理人 100094569  
 弁理士 田中 伸一郎  
 (74) 代理人 100109070  
 弁理士 須田 洋之  
 (74) 代理人 100119013  
 弁理士 山崎 一夫  
 (74) 代理人 100123777  
 弁理士 市川 さつき

最終頁に続く

(54) 【発明の名称】 内部タンパク質タグ

(57) 【特許請求の範囲】

【請求項 1】

対象とするタンパク質の N 末端断片と、対象とするタンパク質の C 末端断片と、前記 N 末端断片及び前記 C 末端断片の間の内部タグとを含むポリペプチドであって、前記内部タグが、対象とするタンパク質の中に挿入された配列番号 390 及び配列番号 2271 から選択されるアミノ酸配列を含み、前記内部タグが配列番号 440 のポリペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成され、前記 N 末端断片、C 末端断片及び内部タグを含むポリペプチドが、天然のポリペプチドではない、前記ポリペプチド。

【請求項 2】

前記 N 末端断片及び / 又は前記 C 末端断片が、少なくとも 20 アミノ酸の長さである、請求項 1 に記載のポリペプチド。

【請求項 3】

前記 N 末端断片と前記 C 末端断片とが、前記内部タグの非存在下で直接連結されたとしたならば、対象とする第 1 のタンパク質の配列を含む、請求項 1 に記載のポリペプチド。

【請求項 4】

請求項 1 に記載のポリペプチドをコードする配列を含む、核酸。

【請求項 5】

(a) 請求項 1 に記載のポリペプチドと、

(b) 配列番号 1292 の相補体ポリペプチドを含む第 2 のポリペプチドと、

を含む、生物発光複合体。

【請求項 6】

前記第 2 のポリペプチドが、対象とする分子に連結されている、請求項 5 に記載の生物発光複合体。

【請求項 7】

セレンテラジン基質を更に含む、請求項 5 に記載の生物発光複合体。

【請求項 8】

前記第 2 のポリペプチドが、融合ポリペプチドではないか、又は対象とする分子に連結されていない、請求項 5 に記載の生物発光複合体。

【請求項 9】

対象とするタンパク質の N 末端断片と、対象とするタンパク質の C 末端断片と、2 つ以上の内部タグとを含むポリペプチドであって、前記内部タグが、対象とするタンパク質の中に挿入された配列番号 3 9 0 及び配列番号 2 2 7 1 から選択されるアミノ酸配列を含み、前記内部タグの 1 つ以上が配列番号 4 4 0 のポリペプチドに接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、前記ポリペプチド。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、その全体が参照によって本明細書に組み入れられる 2014 年 9 月 12 日に  
出願された米国仮特許出願シリアル番号 62 / 049 , 875 号に対する優先権を主張する。

【0002】

分野

本明細書で提供されるのは、対象とするタンパク質の中で、または N 末端ペプチド / ポリペプチドと C 末端ペプチド / ポリペプチドとの間で内部に挿入される実質的に非発光性のペプチド / ポリペプチドタグである。内部に挿入されたタグの相補体ポリペプチド / ペプチドとの相互作用により生物発光レポーター複合体が生じる。

【背景技術】

【0003】

レポーターまたは親和性タグによってタンパク質にタグを付けることはタンパク質の機能及び挙動を解析するのに一般に使用される。一般に、対象とするタンパク質の C 末端または N 末端を用いて遺伝子融合を生成する。

【発明の概要】

【0004】

本明細書で提供されるのは、対象とするタンパク質の中で、または N 末端ペプチド / ポリペプチドと C 末端ペプチド / ポリペプチドとの間で内部に挿入される実質的に非発光性のペプチド / ポリペプチドタグである。内部に挿入されたタグの相補体ポリペプチド / ペプチドとの相互作用により生物発光レポーター複合体が生じる。

【0005】

一部の実施形態では、本明細書で提供されるのは、その中に挿入された内部タグを伴ったタンパク質またはポリペプチドを含む組成物、系、方法等である。一部の実施形態では、本明細書で提供されるのは、C 末端ペプチド / ポリペプチドと N 末端ペプチド / ポリペプチドとの間に挿入された内部タグを含む組成物、系、方法等である。特定の実施形態では、構造相補体配列（または内部タグ）も提供される（たとえば、遊離のまたは融合として（たとえば、内部のもしくは末端の））。一部の実施形態では、内部タグ及び構造相補体は双方とも実質的に不活性である（たとえば、酵素活性を欠いている（たとえば、実質的に非発光性等））。一部の実施形態では、内部タグ及び構造相補体は互いに高い親和性を有し、溶液にて一緒であれば、複合体（たとえば、安定な複合体）を形成する。他の実施形態では、内部タグ及び構造相補体は互いに低い親和性を有し、外部の因子 / 力（たと

10

20

30

40

50

えば、内部タグ及び構造相補体に融合される相互作用要素）によって一緒にされない限り、複合体（たとえば、安定な複合体）を形成しない。一部の実施形態では、内部タグと構造相補体との複合体は検出可能な活性（たとえば、基質の存在下での発光）を生じる。

【 0 0 0 6 】

一部の実施形態では、内部融合体が提供される。一部の実施形態では、内部タグの配列（たとえば、構造相補体と複合体形成すると検出可能な活性を生じるもの）は、対象とするタンパク質の配列またはポリペプチド配列の中で内部に存在する。一部の実施形態では、内部タグは、（ i ）構造相補体と共に活性のある複合体を形成する内部タグの能力及び（ i i ）対象とするタンパク質またはポリペプチドの構造上または活性上の特徴を維持する位置にて対象とするタンパク質またはポリペプチドの配列の中に存在する。一部の実施形態では、対象とするタンパク質またはポリペプチドの対象とする 1 以上のドメインの構造または活性は、対象とするポリペプチドまたはタンパク質の配列の中での内部タグの存在によって中断されない。一部の実施形態では、内部タグは、それが対象とするタンパク質の表面で表面にアクセス可能に露出されるように対象とするタンパク質の中の位置にて存在する。一部の実施形態では、内部タグは、対象とするタンパク質の構造及び / または活性に対する崩壊を低減する / 最少にするように対象とするタンパク質のループにて存在する。

10

【 0 0 0 7 】

一部の実施形態では、内部タグの配列（たとえば、構造相補体と共に複合体形成すると検出可能な活性を生じるもの）は、C末端ペプチド / ポリペプチドとN末端ペプチド / ポリペプチドとの間で内部に存在する。一部の実施形態では、内部タグは、（ i ）構造相補体と共に活性のある複合体を形成する内部タグの能力及び（ i i ）C末端及びN末端のペプチド / ポリペプチドの構造上または活性上の特徴を維持する位置にてC末端及びN末端のペプチド / ポリペプチドに連結される。

20

【 0 0 0 8 】

一部の実施形態では、内部タグ及び相補体ペプチド / ポリペプチドは同一の多重ドメインタンパク質の別々のドメインに融合される及び / またはドメイン内に挿入される。タンパク質の折り畳みまたは配座変化の際、内部タグと相補体との複合体に由来する活性が検出可能である。

【 0 0 0 9 】

一部の実施形態では、内部タグの使用方法が提供される。以下の実施形態の一部は内部タグ及び対象とするタンパク質と共に使用するために記載されている。しかしながら、適宜、それらはN末端ペプチド / ポリペプチドとC末端ペプチド / ポリペプチドとの間での内部タグと共に使用されてもよい。

30

【 0 0 1 0 】

一部の実施形態では、会合の際、活性のある複合体を形成する内部タグ及び構造相補体を用いて対象とするタンパク質 / ポリペプチドの存在及び / または位置を検出する方法が提供される。一部の実施形態では、内部タグを伴ったポリペプチドの試料（たとえば、細胞、細胞内の位置、試験管内試料等）における存在及び / または位置は、内部タグに対する高い親和性を有する遊離の構造相補体を試料に加えることによって判定される。複合体形成によって生じる活性の検出は対象とするタンパク質 / ポリペプチドの存在及び / または位置を示す。一部の実施形態では、内部タグと構造相補体との活性のある複合体は、試験管内試料、細胞溶解物、生きている細胞の中、生きている生物の中等を含むが、これらに限定されない環境にて検出される。

40

【 0 0 1 1 】

一部の実施形態では、会合の際、活性のある複合体を形成する内部タグ及び構造相補体を用いて対象とするタンパク質と選択される相互作用の相手（たとえば、核酸、ペプチド、タンパク質、ポリペプチド、小分子、脂質等）との間での相互作用を検出する方法が提供される。一部の実施形態では、（ i ）内部タグを伴った対象とするタンパク質と（ i i ）内部タグに対する低い親和性を有する構造相補体に連結された相互作用の相手（たとえ

50

ば、核酸、ペプチド、タンパク質、ポリペプチド、小分子、脂質等)との相互作用は、内部タグと構造相補体との間で活性のある複合体を形成する際に検出される。一部の実施形態では、相互作用は試験管内試料、細胞溶解物、生きている細胞の中、生きている生物の中等を含むが、これらに限定されない環境にて検出される。

【0012】

一部の実施形態では、内部タグ及び相補体ペプチド/ポリペプチド(相補体による内部標識または末端標識)によって対象とするタンパク質を標識することによって対象とするタンパク質の中での分子内相互作用を検出する方法が提供される。内部タグと相補体との複合体の形成によって配座変化、折り畳みまたは他の分子内相互作用が検出される。

【0013】

一部の実施形態では、対象とするタンパク質またはポリペプチドの中に存在する内部タグとその構造相補体の相補性の効率(及び活性のある複合体の形成及び検出可能な発光)は、(i)対象とするタンパク質またはポリペプチドにおける配座変化(たとえば、内部タグの構造相補体へのアクセス性に影響を与える配座変化)、(ii)対象とするタンパク質またはポリペプチドの分子内相互作用(たとえば、薬物との)、及び/または環境変化(たとえば、条件に対する変化)の1以上によって影響を受ける。

【0014】

一部の実施形態では、本明細書で提供されるのは、(i)前から存在するタンパク質の断片ではない、(ii)実質的に非発光性である、(iii)対象とするタンパク質の中で内部に挿入される、及び(iv)構造的に相補性のポリペプチド及び/またはペプチドと共に発光複合体を形成するペプチドタグ及び/またはポリペプチドタグを含む組成物である。本明細書で提供されるのはまた、その相補体ポリペプチド及び/またはペプチドとの内部ペプチド及び/またはポリペプチドタグの発光複合体、及びそのような複合体の形成の際、任意で検出可能な生物発光シグナルを生成する方法である。一部の実施形態では、本明細書で提供されるのは、その1以上が、一緒にすると生物発光複合体を作る内部タンパク質タグとして提供される、2以上の実質的に非発光性のペプチド及び/またはポリペプチドである。一部の実施形態では、実質的に非発光性のペプチド及び/またはポリペプチド内部タグとその相補体ポリペプチド/ペプチドとが生物発光複合体を組み立てる。一部の実施形態では、相補体ペプチド/ポリペプチドも内部タグである。他の実施形態では、相補体は末端(たとえば、C末端のN末端)タグである。他の実施形態では、相補体は別のペプチド、ポリペプチドまたはタンパク質と会合しない(たとえば、遊離である)。一部の実施形態では、その1以上が内部タンパク質タグである3以上の実質的に非発光性のペプチド及び/またはポリペプチド単位が生物発光複合体を組み立てる(たとえば、三元複合体、三次複合体等)。一部の実施形態では、本明細書で提供されるのは、さもなければ実質的に非発光性の内部タグとその実質的に非発光性の構造相補体との生物発光複合体の形成を介して、内部でタグ付けされたタンパク質またはポリペプチドを検出する技術である。一部の実施形態では、対象とするタンパク質と別の部分(たとえば、タンパク質、ペプチド、核酸、脂質、小分子等)との間の相互作用は、対象とするタンパク質の実質的に非発光性の内部タグと内部タグの実質的に非発光性の構造相補体との間での生物発光複合体の形成を検出することによって特定される。一部の実施形態では、そのような組成物は、試験管内試料、細胞溶解物、生きている細胞の中、生きている生物の中等を含むが、これらに限定されない環境にて提供される。

【0015】

一部の実施形態では、タンパク質の異なる領域間、または多重ドメインタンパク質のドメイン間での相互作用は異なる領域/ドメインを内部タグと相補体で標識することによって検出される。内部タグと相補体の複合体に由来する活性はタンパク質内の相互作用(たとえば、配座変化、折り畳み等)を示す。

【0016】

一部の実施形態では、実質的に非発光性の内部タグと実質的に非発光性のその構造相補体との複合体は適当な基質の高エネルギー状態への化学反応を触媒し、光が放射される。

10

20

30

40

50

一部の実施形態では、内部タンパク質タグとその構造相補体との生物発光複合体は基質（たとえば、セレンテラジン、フリマジン等）の存在下で発光を示す。

【0017】

本明細書に記載されている実施形態は主として、（たとえば、少なくとも1つの実質的に非発光性のタグと実質的に非発光性のその相補体を含む）相補性の発光複合体の形成を記載し、それを参照するけれども、本技術が他の検出可能な特質（たとえば、他の酵素活性、蛍光色素分子の生成、発色団の生成等）に同等に適用できることが言及される。発光に関して本明細書に記載されている実施形態は、実質的に非酵素的に活性があるアミノ酸鎖（たとえば、前から存在するタンパク質の断片ではないペプチド及び／またはポリペプチド）である内部タグと、特定された検出可能な活性（たとえば、酵素活性）も欠く、その実質的に相補性のポリペプチド／ペプチドと、酵素的に活性があるそれらの複合体とに適用されると見なされるべきである。本明細書で提供されるのは、実質的に非酵素的に活性がある内部タグと実質的に非酵素的に活性があるその相補体ペプチド／ポリペプチドとの会合の際、検出可能な活性（たとえば、酵素活性）を生成する方法である。

【0018】

本発明はさらに、対象とするタンパク質に内部タグを挿入し、内部タグの構造相補体によって他の部分にタグ付けする（たとえば、内部で標識される、末端標識される等）ことによってタンパク質（またはポリペプチド）と別の部分（たとえば、ペプチド、ポリペプチド、タンパク質、核酸、小分子等）との間での分子相互作用（たとえば、一時的な会合、安定な会合、複合体の形成等）の検出のためのアッセイに関し、その際、対象とするタンパク質と他の部分との間での分子相互作用の非存在下ではシグナルは生じない（たとえば、実質的にシグナルはない）が、対象とするタンパク質と他の部分との間での相互作用の際に内部タグとその相補体との検出可能な（たとえば、生物発光性の）複合体が生成される。そのような実施形態では、生物発光複合体の構築は対象とするタンパク質と他の部分の分子相互作用によって行われる。対象とするタンパク質と他の部分が十分に安定な相互作用で関わり合うのであれば、内部タグとその相補体との生物発光複合体が生じ、生物発光シグナルが生成される。対象とするタンパク質と他の部分が十分に安定な相互作用で関わり合うことができなければ、生物発光複合体は生じない、またはごく弱く生じ、生物発光シグナルは生成されない、または実質的に低減される（たとえば、実質的に検出できない、本質的に検出できない、安定な対照シグナルに比べて差別的に検出できる等）。一部の実施形態では、検出可能な生物発光シグナルの大きさは、対象とするタンパク質と他の部分との間での分子相互作用の量、強度、好ましさ及び／または安定性に比例する（たとえば、直接的に比例する）。

【0019】

一部の実施形態では、本明細書で提供されるのは、配列番号2との100%未満（たとえば、20%...30%...40%...50%...60%...70%...80%、90%以上）の配列同一性を有するアミノ酸配列を含む内部タグであり、その際、該ペプチドが配列番号440の相補体ポリペプチドに接触すると検出可能な生物発光シグナルが生成される。一部の実施形態では、本明細書で提供されるのは、配列番号2との100%未満でかつ30%を超える（たとえば、>30%、>40%、>45%、>50%、>55%、>60%、>65%、>70%、>75%、>80%、>85%、>90%、>95%、>98%、>99%）配列同一性を有するアミノ酸配列を含む内部タグであり、その際、該ペプチドが配列番号440の構造的に相補性のポリペプチドに接触すると検出可能な生物発光シグナルが生成される。一部の実施形態では、内部タグが、配列番号440との100%未満でかつ30%を超える（たとえば、>30%、>40%、>45%、>50%、>55%、>60%、>65%、>70%、>75%、>80%、>85%、>90%、>95%、>98%、>99%）配列同一性を有するポリペプチドと接触すると、検出可能な生物発光シグナルが生成される。特定の実施形態では、内部タグが配列番号440もしくはその一部を含むまたは配列番号440もしくはその一部からなるポリペプチドと会合すると検出可能な生物発光シグナルが生成される、またはそれが実質的

に上昇する。これらの配列に限定されないが、ペプチドのアミノ酸配列は配列番号 3 ~ 4 3 8 及び 2 1 6 2 ~ 2 3 6 5 のアミノ酸配列から選択されてもよい。一部の実施形態では、対象とするタンパク質またはポリペプチドに挿入された（たとえば、タンパク質またはポリペプチドの C 末端または N 末端ではない）上述の内部タグを含むポリペプチドが提供される。一部の実施形態では、内部タグの構造相補体は単独で、または別の部分（たとえば、タンパク質、ペプチド、ポリペプチド、核酸、脂質、小分子等）の（たとえば、内部のまたは末端の）タグとして提供される。特定の実施形態では、（ a ）内部タグ（たとえば、N 末端または C 末端に位置しない）を有する第 1 のポリペプチドと、（ b ）内部タグの構造相補体を含むペプチドまたはポリペプチドとを含む生物発光複合体が提供され、その際、会合すれば、内部タグ及びその構造相補体は適当な基質の存在下で検出可能な生物発光シグナルを放射する。一部の実施形態では、内部タグは、配列番号 2 との 1 0 0 % 未満でかつ 3 0 % を超える配列同一性を有するアミノ酸配列を含み、内部タグが基質の存在下で配列番号 4 4 0 の構造的に相補性のポリペプチドと接触すると、検出可能な生物発光シグナルが生成される。

10

#### 【 0 0 2 0 】

一部の実施形態では、本明細書で提供されるのは、配列番号 4 4 0 との 1 0 0 % 未満の配列同一性を有するアミノ酸配列を含む内部タグであり、その際、内部タグが基質の存在下で配列番号 2 のペプチドと接触すると、検出可能な生物発光シグナルが生成される。一部の実施形態では、本発明は、配列番号 4 4 0 との 1 0 0 % 未満でかつ 3 0 % を超える（たとえば、> 3 0 %、> 4 0 %、> 4 5 %、> 5 0 %、> 5 5 %、> 6 0 %、> 6 5 %、> 7 0 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 5 %、> 9 8 %、> 9 9 %）配列同一性を有するアミノ酸配列を含む内部タグを提供し、その際、内部タグが配列番号 2 のペプチドと接触すると、検出可能な生物発光シグナルが生成される。一部の実施形態では、内部タグが、配列番号 2 との 1 0 0 % 未満でかつ 3 0 % を超える（たとえば、> 3 0 %、> 4 0 %、> 4 5 %、> 5 0 %、> 5 5 %、> 6 0 %、> 6 5 %、> 7 0 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 5 %、> 9 8 %、> 9 9 %）配列同一性を有するペプチドと接触すると、検出可能な生物発光シグナルが生成される。そのような配列に限定されないが、内部タグの配列は、配列番号 4 4 1 ~ 2 1 5 6 のアミノ酸配列の 1 つから選択されてもよい。一部の実施形態では、適当な基質の存在下で内部タグが配列番号 2 のペプチドと会合すると、検出可能な生物発光シグナルが生成される。特定の実施形態では、（ a ）内部タグを有する（N 末端または C 末端に位置しない）第 1 のポリペプチドと、（ b ）内部タグの構造相補体を含むペプチドまたはポリペプチドとを含む生物発光複合体が提供され、その際、会合すれば、内部タグ及びその構造相補体は適当な基質の存在下で検出可能な生物発光シグナルを放射する。一部の実施形態では、内部タグは配列番号 4 4 0 との 1 0 0 % 未満でかつ 3 0 % を超える配列同一性を有するアミノ酸配列を含み、内部が基質の存在下で配列番号 2 の構造的に相補性のペプチドに接触すると、検出可能な生物発光シグナルが生成される。

20

30

#### 【 0 0 2 1 】

一部の実施形態では、本明細書で提供されるのは、本明細書に記載されているペプチド、ポリペプチド（たとえば、内部タグを含む、末端タグを含む、等）、タンパク質（たとえば、内部タグを含む、末端タグを含む、等）、融合タンパク質等のいずれかをコードする核酸（たとえば、DNA、RNA 等）、オリゴヌクレオチド、ベクター等である。一部の実施形態では、配列番号 3 ~ 4 3 8 及び 2 1 6 2 ~ 2 3 6 5（たとえば、ペプチド内部タグをコードする、ペプチド構造相補体をコードする）及び / または配列番号 4 4 1 ~ 2 1 5 6（たとえば、ポリペプチド内部タグをコードする、ポリペプチド構造相補体をコードする）の核酸配列の 1 つを含むまたは 1 つからなる核酸が提供される。一部の実施形態では、配列番号 3 ~ 4 3 8 及び 2 1 6 2 ~ 2 3 6 5 及び / または配列番号 4 4 1 ~ 2 1 5 6 のアミノ酸配列をコードする他の核酸配列が提供される。

40

#### 【 0 0 2 2 】

特定の実施形態では、本明細書で提供されるのは、（ a ）配列番号 2 との 1 0 0 % 未満

50

の配列同一性（＜ 9 9 %、＜ 9 5 %、＜ 9 0 %、＜ 8 0 %、＜ 7 0 %、＜ 6 0 %、＜ 5 0 %、等）を有するアミノ酸配列を含む内部タグ（N末端またはC末端ではない）を有する第1のポリペプチドと、（b）配列番号440との100%未満でかつ30%を超える（たとえば、＞ 3 0 %、＞ 4 0 %、＞ 4 5 %、＞ 5 0 %、＞ 5 5 %、＞ 6 0 %、＞ 6 5 %、＞ 7 0 %、＞ 7 5 %、＞ 8 0 %、＞ 8 5 %、＞ 9 0 %、＞ 9 5 %、＞ 9 8 %、＞ 9 9 %）配列同一性を有するアミノ酸配列を含む（たとえば、内部でまたは末端で）第2のポリペプチドとを含む生物発光複合体であり、その際、生物発光複合体は基質の存在下で検出可能な生物発光を示す。特定の実施形態では、本明細書で提供されるのは、（a）配列番号2との100%未満でかつ30%を超える（たとえば、＞ 3 0 %、＞ 4 0 %、＞ 4 5 %、＞ 5 0 %、＞ 5 5 %、＞ 6 0 %、＞ 6 5 %、＞ 7 0 %、＞ 7 5 %、＞ 8 0 %、＞ 8 5 %、＞ 9 0 %、＞ 9 5 %、＞ 9 8 %、＞ 9 9 %）配列同一性を有するアミノ酸配列を含む内部タグを含むポリペプチドと、（b）配列番号440との100%未満でかつ30%を超える（たとえば、＞ 3 0 %、＞ 4 0 %、＞ 4 5 %、＞ 5 0 %、＞ 5 5 %、＞ 6 0 %、＞ 6 5 %、＞ 7 0 %、＞ 7 5 %、＞ 8 0 %、＞ 8 5 %、＞ 9 0 %、＞ 9 5 %、＞ 9 8 %、＞ 9 9 %）配列同一性を有するアミノ酸配列を含むポリペプチドとを含む生物発光複合体であり、その際、生物発光複合体は基質の存在下で検出可能な生物発光を示す。

10

#### 【0023】

種々の実施形態では、（a）前から存在するタンパク質の断片ではない内部タグ配列を含む第1のアミノ酸配列と、（b）前から存在するタンパク質の断片ではないアミノ酸配列を含む第2のアミノ酸配列とを含む生物発光複合体が提供され、その際、生物発光複合体は、第1のアミノ酸配列と第2のアミノ酸配列が会合すると検出可能な生物発光を示す。

20

#### 【0024】

一部の実施形態では、ポリペプチドを相互作用させることが提供され、その際、第1のポリペプチドは内部タグ（その配列内の一部で挿入されたアミノ酸配列（たとえば、N末端またはC末端ではない）を含み、かつ第2のポリペプチドは構造相補体タグ（たとえば、内部タグに構造的に相補性である内部にまたは末端に位置するアミノ酸配列）を含む。第1と第2のポリペプチドの相互作用の際に、内部タグと構造相補体タグが会合して生物発光複合体を形成する。一部の実施形態では、相互作用するポリペプチドの相互作用は生物発光複合体の生物発光に基づいて評価される。特定の実施形態では、内部タグと構造相補体タグとの間での相互作用（たとえば、非共有結合性の相互作用（たとえば、水素結合、イオン結合、ファンデルワールス力、疎水性の相互作用等）、共有結合性の相互作用（たとえば、ジスルフィド結合）等）は、相互作用するポリペプチドの非存在下では有意な生物発光複合体の形成を生じない。一部の実施形態では、そのような系は細胞内に存在する（たとえば、発現する）。

30

#### 【0025】

一部の実施形態では、本明細書で提供されるのは、（a）ポリペプチド配列内に挿入された第1の実質的に非発光性の要素と、（b）分子実体（たとえば、小分子等）に連結された第2の実質的に非発光性の要素（たとえば、遊離の、ポリペプチドに連結された（内部にまたは末端に）とを含む生物発光複合体であり、その際、各非発光性の要素は前から存在するタンパク質の断片ではない。

40

#### 【0026】

本明細書に記載されている種々の実施形態は、たとえば、（a）第1のアミノ酸配列の中に内部タグを挿入し、第2のアミノ酸配列に相補体配列を連結する（たとえば、内部にまたは末端に）工程であって、その際、内部タグ及び相補体配列は前から存在するタンパク質の断片ではなく、内部タグ及び相補体配列の複合体は検出可能な生物発光シグナル（たとえば、内部タグ及び相補体配列と別々に対して実質的に上昇した生物発光）を放射し、内部タグ及び相補体配列の相互作用（たとえば、非共有結合性）は追加の安定化する及び/または凝集する条件の非存在下では複合体を形成するには不十分であり、またはごく弱く形成し、第1のアミノ酸配列と第2のアミノ酸配列の相互作用が追加の安定化する及

50

び／または凝集する力を提供して内部タグと相補体配列との複合体を生成することと、(b)工程(a)のタグ付けされた第1のアミノ酸配列及び第2のアミノ酸配列を第1のアミノ酸配列と第2のアミノ酸配列の間で相互作用を生じさせる条件に置く工程と、(c)適当な基質の存在下で内部タグと相補体配列との複合体によって放射される生物発光シグナルを検出する工程とを含む、第1のアミノ酸配列と第2のアミノ酸配列の間での相互作用を検出する方法を提供し、その際、生物発光シグナルの検出は第1のアミノ酸配列と第2のアミノ酸配列の間での相互作用を指し示す。一部の実施形態では、第1のアミノ酸配列及び内部タグは内部融合を含む。一部の実施形態では、第2のアミノ酸配列及び相補体配列は内部融合または従来の融合を含む。一部の実施形態では、第1の融合タンパク質(たとえば、内部タグを含む)及び第2の融合タンパク質(たとえば、相補体配列を含む)はさらに、融合された要素間でリンカーを含む。特定の実施形態では、融合タンパク質は前記融合タンパク質をコードする核酸から発現する。一部の実施形態では、単一のベクターが双方の融合タンパク質を含む。他の実施形態では、第1と第2の融合タンパク質は別々のベクターから発現する。

#### 【0027】

一部の実施形態では、本明細書で提供されるのは、N末端断片と、C末端断片と、内部タグとを含むポリペプチドであり、その際、内部タグは対象とするタンパク質の中に挿入された配列番号2との100%未満でかつ30%を超える配列同一性を有するアミノ酸配列を含み、内部タグが配列番号440のポリペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成される。一部の実施形態では、N末端断片及びC末端断片は双方とも長さ少なくとも20アミノ酸である。一部の実施形態では、N末端断片及び／またはC末端断片は長さ少なくとも50アミノ酸である。一部の実施形態では、内部タグは配列番号2のペプチドに比べて1以上の形質の向上を示し、その際、形質は、配列番号440のポリペプチドに対する親和性、配列番号440のポリペプチドと組み合わせた場合の発現、細胞内溶解性、細胞内安定性及び生物発光活性から選択される。一部の実施形態では、内部タグは表1のペプチドから選択される。一部の実施形態では、N末端断片及びC末端断片は、内部タグの非存在下で直接連結されるならば対象とする第1のタンパク質の配列を含む。一部の実施形態では、内部でタグ付けされたポリペプチドをコードする配列を含む核酸が提供される。一部の実施形態では、(a)内部でタグ付けされたポリペプチドと、(b)配列番号440との100%未満でかつ30%を超える配列同一性を有する相補体ポリペプチドを含む第2のポリペプチドとを含む生物発光複合体が提供される。一部の実施形態では、内部タグ及び相補体ポリペプチドは互いに対して低い親和性を有する。一部の実施形態では、第2のポリペプチドは対象とする第2のタンパク質との融合体である。一部の実施形態では、融合体は内部融合体または従来の融合体である。一部の実施形態では、対象とする第2のタンパク質はN末端断片及び／またはC末端断片のすべてまたは一部に対する親和性を有する。一部の実施形態では、親和性は、第1もしくは第2のタンパク質もしくは双方に対する構造的修飾(たとえば、翻訳後修飾)によって、または第3の分子(たとえば、薬物、核酸、タンパク質等)との相互作用によって変えられてもよい。一部の実施形態では、第2のポリペプチドは対象とする分子に連結される。一部の実施形態では、N末端断片及び／またはC末端断片のすべてまたは一部は、対象とする分子に対する親和性を有する。一部の実施形態では、生物発光複合体はさらにセレンテラジン基質(たとえば、フリマジン)を含む。一部の実施形態では、内部タグ及び相補体ポリペプチドは互いに対して高い親和性を有する。一部の実施形態では、第2のポリペプチドは融合ポリペプチドではなく、または対象とする分子に連結されない。一部の実施形態では、相補体ポリペプチドは表2のペプチドから選択される。

#### 【0028】

一部の実施形態では、本明細書で提供されるのは、N末端断片と、C末端断片と、内部タグとを含むポリペプチドであり、その際、内部タグは対象とするタンパク質の中に挿入された配列番号440との100%未満でかつ30%を超える配列同一性を有するアミノ酸配列を含み、検出ペプチドが配列番号2のポリペプチドと接触すると、基質の存在下で

検出可能な生物発光シグナルが生成される。一部の実施形態では、N末端断片及びC末端断片は双方とも長さ少なくとも20アミノ酸である。一部の実施形態では、N末端断片及び/またはC末端断片は長さ少なくとも50アミノ酸である。一部の実施形態では、内部タグは配列番号440のペプチドに比べて1以上の形質の向上を示し、その際、形質は、配列番号2のポリペプチドに対する親和性、配列番号2のポリペプチドと組み合わせた場合の発現、細胞内溶解性、細胞内安定性及び生物発光活性から選択される。一部の実施形態では、内部タグは表2のペプチドから選択される。一部の実施形態では、N末端断片及びC末端断片は、内部タグの非存在下で直接連結されるならば対象とする第1のタンパク質の配列を含む。一部の実施形態では、内部でタグ付けされたポリペプチドをコードする配列を含む核酸が提供される。一部の実施形態では、(a)内部でタグ付けされたポリペプチドと、(b)配列番号2との100%未満かつ30%を超える配列同一性を有する相補体ポリペプチドとを含む生物発光複合体が提供される。一部の実施形態では、内部タグ及び相補体ポリペプチドは互いに対して低い親和性を有する。一部の実施形態では、相補体ペプチドは対象とする第2のタンパク質との融合体である。一部の実施形態では、融合体は内部融合体または従来の融合体である。一部の実施形態では、対象とする第2のタンパク質はN末端断片及び/またはC末端断片のすべてまたは一部に対する親和性を有する。一部の実施形態では、親和性は、第1もしくは第2のタンパク質もしくは双方に対する構造的修飾(たとえば、翻訳後修飾)によって、または第3の分子(たとえば、薬物、核酸、タンパク質等)との相互作用によって変えられてもよい。一部の実施形態では、相補体ペプチドは対象とする分子に連結される。一部の実施形態では、N末端断片及び/またはC末端断片のすべてまたは一部は、対象とする分子に対する高い親和性を有する。一部の実施形態では、生物発光複合体はさらにセレンテラジン基質を含む。一部の実施形態では、内部タグ及び相補体ペプチドは互いに対して高い親和性を有する。一部の実施形態では、相補体ペプチドは表1のペプチドから選択される。一部の実施形態では、相補体ペプチドは融合ポリペプチドではなく、または対象とする分子に連結されない。

#### 【0029】

一部の実施形態では、本明細書で提供されるのは、(a)第1のアミノ酸配列の中に内部タグを挿入することによって内部融合体を作り出すので、前記内部タグが第1のアミノ酸配列のN末端にもなく、C末端にもないこと、その際、内部タグが配列番号2との100%未満かつ30%を超える配列同一性を有し、内部タグが配列番号440のポリペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、(b)第2のアミノ酸配列と相補体ポリペプチドとの第2の融合体を作り出すこと、その際、相補体ポリペプチドが配列番号440との100%未満かつ30%を超える配列同一性を有し、相補体ポリペプチドが配列番号2のペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、(c)第1のアミノ酸配列と第2のアミノ酸配列との間で見込まれる相互作用を生じさせる条件に内部融合体と第2の融合体とセレンテラジン基質とを置くことと、(d)存在するならば、放射された生物発光シグナルを検出すること、その際、生物発光シグナルの検出は第1のアミノ酸配列と第2のアミノ酸配列との間での相互作用を指し示すこととを含む、第1のアミノ酸配列と第2のアミノ酸配列との間での相互作用を検出する方法である。一部の実施形態では、相互作用は、生きている細胞または生物にて、細胞または生物によって放射された生物発光シグナルを検出することによって検出される。一部の実施形態では、細胞の環境の変化の結果生じる相互作用の変化は、変化した環境が存在しない対照細胞に対する放射された生物発光シグナルの差異を検出することによって検出される。一部の実施形態では、変化した環境は、培養培地に分子(たとえば、薬物)を加えたまたはそれから分子を取り除いた結果である。一部の実施形態では、第2の融合体は内部融合体または従来の融合体である。一部の実施形態では、内部融合体は第1のアミノ酸配列と内部タグをコードする第1の核酸配列から発現し、第2の融合体は第2のアミノ酸配列と相補体ポリペプチドをコードする第2の核酸配列から発現する。一部の実施形態では、単一のベクターが第1の核酸配列と第2の核酸配列とを含む。一部の実施形態では、第1の核酸配列及

10

20

30

40

50

び第2の核酸配列は別々のベクターにある。一部の実施形態では、工程(a)及び(b)は細胞内にて内部融合体と第2の融合体を発現させることを含む。

【0030】

一部の実施形態では、本明細書で提供されるのは、(a)第1のアミノ酸配列の中に内部タグを挿入することによって内部融合体を作り出すので、前記内部タグが第1のアミノ酸配列のN末端にもなく、C末端にもないこと、その際、内部タグが配列番号440との100%未満でかつ30%を超える配列同一性を有し、内部タグが配列番号2のペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、(b)第2のアミノ酸配列と相補体ペプチドとの第2の融合体を作り出すこと、その際、相補体ペプチドが配列番号2との100%未満でかつ30%を超える配列同一性を有し、相補体ペプチドが配列番号2のポリペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、(c)第1のアミノ酸配列と第2のアミノ酸配列との間で見込まれる相互作用を生じさせる条件に内部融合体と第2の融合体とセレンテラジン基質とを置くことと、(d)存在するならば、放射された生物発光シグナルを検出すること、その際、生物発光シグナルの検出は第1のアミノ酸配列と第2のアミノ酸配列との間での相互作用を指し示すこととを含む、第1のアミノ酸配列と第2のアミノ酸配列との間での相互作用を検出する方法である。一部の実施形態では、第2の融合体は内部融合体または従来の融合体である。一部の実施形態では、内部融合体は第1のアミノ酸配列と内部タグをコードする第1の核酸配列から発現し、第2の融合体は第2のアミノ酸配列と相補体ペプチドをコードする第2の核酸配列から発現する。一部の  
10  
20  
実施形態では、単一のベクターが第1の核酸配列と第2の核酸配列とを含む。一部の実施形態では、第1の核酸配列及び第2の核酸配列は別々のベクターにある。一部の実施形態では、工程(a)及び(b)は細胞内にて内部融合体と第2の融合体を発現させることを含む。

【0031】

一部の実施形態では、本明細書で提供されるのは、(a)標的ポリペプチドの中に内部タグを挿入することによって内部融合体を作り出すので、前記内部タグが標的ポリペプチドのN末端にもなく、C末端にもないこと、その際、内部タグが配列番号440との100%未満でかつ30%を超える配列同一性を有し、内部タグが配列番号2のペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、(b)前記試料に(i)配列番号2との100%未満でかつ30%を超える配列同一性を有する相補体ペプチドと(ii)セレンテラジン基質とを加えることと、(c)存在するならば、放射された生物発光シグナルを検出すること、その際、生物発光シグナルの検出は試料における標的ポリペプチドの存在を指し示すこととを含む、試料にて標的ポリペプチドを検出する方法である。一部の実施形態では、試料は細胞を含む。一部の実施形態では、工程(a)は前記細胞にて前記内部融合体を発現させることを含む。一部の実施形態では、工程(b)(i)は前記細胞にて前記相補体ペプチドを含む。  
30

【0032】

一部の実施形態では、本明細書で提供されるのは、(a)標的ポリペプチドの中に内部タグを挿入することによって内部融合体を作り出すので、前記内部タグが標的ポリペプチドのN末端にもなく、C末端にもないこと、その際、内部タグが配列番号2との100%未満でかつ30%を超える配列同一性を有し、内部タグが配列番号440のペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、(b)前記試料に(i)配列番号440との100%未満でかつ30%を超える配列同一性を有する相補体ポリペプチドと(ii)セレンテラジン基質とを加えることと、(c)存在するならば、放射された生物発光シグナルを検出すること、その際、生物発光シグナルの検出は試料における標的ポリペプチドの存在を指し示すこととを含む、試料にて標的ポリペプチドを検出する方法である。一部の実施形態では、試料は細胞を含む。一部の  
40  
50  
実施形態では、工程(a)は前記細胞にて前記内部融合体を発現させることを含む。一部の実施形態では、工程(b)(i)は前記細胞にて前記相補体ポリペプチドを含む。

## 【 0 0 3 3 】

一部の実施形態では、本明細書で提供されるのは、( a ) ポリペプチドが配列番号 2 のペプチドと接触すると、基質の存在下で検出可能な生物発光シグナルが生成される、配列番号 4 4 0 との 1 0 0 % 未満でかつ 3 0 % を超える配列同一性を有するアミノ酸配列を含む相補体ポリペプチドと、( b ) 前記ポリペプチドと配列番号 2 のペプチドとによって生成される生物発光複合体のための基質とを含む検出試薬である。

## 【 0 0 3 4 】

一部の実施形態では、本明細書で提供されるのは、( a ) ペプチドが配列番号 4 4 0 のポリペプチドと接触すると、基質の存在下で検出可能な生物発光シグナルが生成される、配列番号 2 との 1 0 0 % 未満でかつ 3 0 % を超える配列同一性を有するアミノ酸配列を含む相補体ペプチドと、( b ) 前記ペプチドと配列番号 4 4 0 のポリペプチドとによって生成される生物発光複合体のための基質とを含む検出試薬である。

## 【 0 0 3 5 】

特定の実施形態では、内部タグ及び/または構造相補体は配列番号 2 または配列番号 4 4 0 との 1 0 0 % の配列同一性を有するアミノ酸を含む、またはそれからなる。一部の実施形態では、そのような内部タグ及び構造相補体は本明細書に記載されている任意の実施形態で使用され、本明細書に記載されている他のペプチドまたはポリペプチドと共に使用される。

## 【 0 0 3 6 】

一部の実施形態では、本明細書で提供されるのは、( a ) 第 1 のアミノ酸配列の中に内部タグを挿入することによって内部融合体を作り出すので、前記内部タグが第 1 のアミノ酸配列の N 末端にもなく、C 末端にもないこと、その際、内部タグが配列番号 2 との 1 0 0 % 未満でかつ 3 0 % を超える配列同一性を有し、内部タグが配列番号 4 4 0 のポリペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、( b ) 第 2 のアミノ酸配列と相補体ポリペプチドとの第 2 の融合体を作り出すこと、その際、相補体ポリペプチドが配列番号 4 4 0 との 1 0 0 % 未満でかつ 3 0 % を超える配列同一性を有し、相補体ポリペプチドが配列番号 2 のペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、( c ) 第 1 のアミノ酸配列と第 2 のアミノ酸配列との間で見込まれる相互作用を生じさせる条件に内部融合体と第 2 の融合体とセレンテラジン基質を置くことと、( d ) 存在するならば、放射された生物発光シグナルを検出することと、( e ) 内部融合体と第 2 の融合体とセレンテラジン基質とに作用剤を加えることと、( f ) 存在するならば、放射された生物発光シグナルを検出することと、( g ) 工程 ( d ) と工程 ( f ) の生物発光シグナルを比較すること、その際、工程 ( d ) から工程 ( f ) までの生物発光シグナルの変化は、第 1 のアミノ酸配列と第 2 のアミノ酸配列との間での相互作用の作用剤による変化を指し示すこととを含む、第 1 のアミノ酸配列と第 2 のアミノ酸配列との間での相互作用の変化を検出する方法である。一部の実施形態では、工程 ( a ) 及び ( b ) は細胞内で内部融合体及び第 2 の融合体を発現させることを含む。一部の実施形態では、作用剤はペプチドまたは小分子である。一部の実施形態では、作用剤は相互作用の阻害剤であり、その際、低下した相互作用は生物発光シグナルの低下によって検出される。一部の実施形態では、作用剤は相互作用の活性化剤であり、増大した相互作用は生物発光シグナルの上昇によって検出される。

## 【 0 0 3 7 】

一部の実施形態では、本明細書で提供されるのは、( a ) 第 1 のアミノ酸配列の中に内部タグを挿入することによって内部融合体を作り出すので、前記内部タグが第 1 のアミノ酸配列の N 末端にもなく、C 末端にもないこと、その際、内部タグが配列番号 4 4 0 との 1 0 0 % 未満でかつ 3 0 % を超える配列同一性を有し、内部タグが配列番号 2 のペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、( b ) 第 2 のアミノ酸配列と相補体ポリペプチドとの第 2 の融合体を作り出すこと、その際、相補体ポリペプチドが配列番号 2 との 1 0 0 % 未満でかつ 3 0 % を超える配

列同一性を有し、相補体ポリペプチドが配列番号440のポリペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成されることと、(c)第1のアミノ酸配列と第2のアミノ酸配列との間で見込まれる相互作用を生じさせる条件に内部融合体と第2の融合体とセレンテラジン基質とを置くことと、(d)存在するならば、放射された生物発光シグナルを検出することと、(e)内部融合体と第2の融合体とセレンテラジン基質とに作用剤を加えることと、(f)存在するならば、放射された生物発光シグナルを検出することと、(g)工程(d)と工程(f)の生物発光シグナルを比較すること、その際、工程(d)から工程(f)までの生物発光シグナルの変化は、第1のアミノ酸配列と第2のアミノ酸配列との間での相互作用の作用剤による変化を指し示すこととを含む、第1のアミノ酸配列と第2のアミノ酸配列との間での相互作用の変化を検出する方法である。一部の実施形態では、工程(a)及び(b)は細胞内で内部融合体及び第2の融合体を発現させることを含む。一部の実施形態では、作用剤はペプチドまたは小分子である。一部の実施形態では、作用剤は相互作用の阻害剤であり、その際、低下した相互作用は生物発光シグナルの低下によって検出される。一部の実施形態では、作用剤は相互作用の活性化剤であり、増大した相互作用は生物発光シグナルの上昇によって検出される。

10

**【0038】**

一部の実施形態では、本明細書で提供されるのは、(a)第1のアミノ酸配列の中に内部タグを挿入することによって内部融合体を作り出すので、前記内部タグが第1のアミノ酸配列のN末端にもなく、C末端にもないこと、その際、内部タグが配列番号2との100%未満でかつ30%を超える配列同一性を有し、内部タグが配列番号440のポリペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成され、第1のアミノ酸配列の第1の構造配座が第1のアミノ酸配列の第2の構造配座に対して内部タグへのアクセスを変化させることと、(b)内部タグと(i)配列番号440との100%未満でかつ30%を超える配列同一性を有する相補体ポリペプチドまたは(ii)第2のアミノ酸配列と相補体ポリペプチドとの第2の融合体とをセレンテラジン基質の存在下に置くことと、(c)存在するならば、放射された生物発光シグナルを検出することと、(d)内部融合体と第2の融合体とセレンテラジン基質とに作用剤を加えることと、(e)存在するならば、放射された生物発光シグナルを検出することと、(f)工程(c)と工程(e)の生物発光シグナルを比較すること、その際、工程(c)から工程(e)までの生物発光シグナルの変化は作用剤による第1のアミノ酸配列の配座の変化を指し示すこととを含む、第1のアミノ酸配列の構造配座における作用剤による変化を検出する方法である。一部の実施形態では、配座変化を誘導することは、第1のアミノ酸配列の一部を切断するプロテアーゼを加えること、第1のアミノ酸配列に結合する作用剤の添加、及びアッセイ条件を変えることから選択される。

20

30

**【0039】**

一部の実施形態では、本明細書で提供されるのは、(a)第1のアミノ酸配列の中に内部タグを挿入することによって内部融合体を作り出すので、前記内部タグが第1のアミノ酸配列のN末端にもなく、C末端にもないこと、その際、内部タグが配列番号440との100%未満でかつ30%を超える配列同一性を有し、内部タグが配列番号2のペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成され、第1のアミノ酸配列の第1の構造配座が第1のアミノ酸配列の第2の構造配座に比べて内部タグへのアクセスを変化させることと、(b)内部タグと(i)配列番号2との100%未満でかつ30%を超える配列同一性を有する相補体ペプチドまたは(ii)第2のアミノ酸配列と相補体ペプチドとの第2の融合体とをセレンテラジン基質の存在下に置くことと、(c)存在するならば、放射された生物発光シグナルを検出することと、(d)内部融合体と第2の融合体とセレンテラジン基質とに作用剤を加えることと、(e)存在するならば、放射された生物発光シグナルを検出することと、(f)工程(c)と工程(e)の生物発光シグナルを比較すること、その際、工程(c)から工程(e)までの生物発光シグナルの変化は作用剤による第1のアミノ酸配列の配座の変化を指し示すことと含む

40

50

、第1のアミノ酸配列の構造配座における作用剤による変化を検出する方法である。一部の実施形態では、配座変化を誘導することは、第1のアミノ酸配列の一部を切断するプロテアーゼを加えること、第1のアミノ酸配列に結合する作用剤を加えること、及びアッセイ条件を変えることから選択される。

#### 【0040】

一部の実施形態では、本明細書で提供されるのは、N末端断片とC末端断片と2以上の内部タグを含むポリペプチドであり、その際、内部タグは対象とするタンパク質の中に挿入された配列番号2との100%未満でかつ30%を超える配列同一性を有するアミノ酸配列を含み、内部タグの1以上が配列番号440のポリペプチドと接触すると、セレンテラジン基質の存在下で検出可能な生物発光シグナルが生成される。一部の実施形態では、2以上の内部タグは2つの内部タグ（たとえば、直列タグ）である。一部の実施形態では、2以上の内部タグは互いに直接接続される。一部の実施形態では、2以上の内部タグは1以上のリンカー（たとえば、ペプチドリinker）によって分離される。一部の実施形態では、2以上の内部タグは対象とするタンパク質またはポリペプチドの中での単一の位置にて挿入される。一部の実施形態では、2以上の内部タグは対象とするタンパク質またはポリペプチドの中での2以上の位置にて挿入される。一部の実施形態では、2以上の内部タグは同一アミノ酸配列を含む。一部の実施形態では、2以上のまたは2以上の内部タグは同一ではないアミノ酸配列を含む。一部の実施形態では、2以上の内部タグは、相補体配列との親和性に影響を有してもよいし、または有さなくてもよいが、内部タグまたは直列タグの電荷全体をさらに荷電するようにまたは中性に近いように変化させるアミノ酸置換を有する。

#### 【図面の簡単な説明】

#### 【0041】

【図1】内部タグ（たとえば、NL pepまたはNL poly）が対象とする第1のタンパク質（POI1）に内部ループとして挿入され、相補体配列（たとえば、NL pepまたはNL poly）が対象とする第2のタンパク質（POI2）に末端で融合される例となる実施形態を示す模式図である。内部タグ及び相補体配列は互いに低い親和性を有するので、それらは外部の力の非存在下では複合体を形成することにおいて効果はない（たとえば、検出できないまたは無視できる量の複合体を生じる）。内部タグ及び相補体配列は別々に実質的に非発光性であるけれども、POI1及びPOI2の相互作用の際、内部タグと相補体配列との間で生物発光複合体が形成される。

【図2】内部タグ（たとえば、NL pepまたはNL poly）が対象とする第1のタンパク質（POI1）に挿入され、遊離の相補体配列（たとえば、NL pepまたはNL poly）が提供される例となる実施形態を示す模式図である。この例では、内部タグは、対象とするタンパク質が相補体配列によって内部タグをアクセス可能にする配座変化を経験した後でのみアクセス可能である位置に置かれる。内部タグ及び相補体配列は、内部タグ及び相補体配列が会合すると複合体が生じるように互いに高い親和性を有する。内部タグ及び相補体配列が同一試料に存在する場合、それらの間で生物発光複合体が形成される。

【図3】内部タグ（たとえば、NL pepまたはNL poly）及びその相補体配列が対象とするポリペプチド（POI1）に挿入される／融合される例となる実施形態を示す模式図である。この例では、内部タグは、対象とするタンパク質が相補体配列によって内部タグをアクセス可能にする配座変化を経験した後でのみアクセス可能である位置に置かれる。内部タグ及び相補体配列は、内部タグ及び相補体配列が会合すると複合体が生じるように互いに高い親和性を有する。内部タグ及び相補体配列が同一試料に存在する場合、それらの間で生物発光複合体が形成される。

【図4】HALOTAGタンパク質（Promega Corp）に挿入された高親和性NLペプチド86とNL poly 1 ISの構造的な相補性を示す代表的な実験の結果を示す図である。この実験では、指示された構築物についての発現構築物でHeLa細胞に形質移入した。細胞を24時間インキュベートした。NANOLUC基質フリマジンの添加

に続いて各試料の発光を測定した。

【図5】それに挿入された高親和性NLペプチド86を有するHALOTAGタンパク質(Promega Corp)の機能を示す画像である。

【図6】プロテインGに融合されたNLpoly11Sと内部でタグ付けされたNLpep114標的タンパク質との間での抗体主導のNANOLUC生物発光相補性を手段として、アッセイ成分をどのように使用して標的タンパク質への抗体の結合をスクリーニングするのかを示す模式図である。内部のpep114タグを含有する標的はIL6シグナルペプチド(SP)と共に哺乳類細胞の培養物で発現する。SPは標的を分泌経路に向ける。培地(+/-)細胞にて標的をアッセイすることができる。例では、ラージビット(11S-プロテインG融合タンパク質)の精製調製物及び試験抗体を哺乳類細胞の培養物に直接加えた。ラージビットのプロテインGドメインは重鎖定常ドメイン2と3の間の試験抗体のFc領域に結合する。試験抗体が標的に結合するのであれば、ラージビットとスモールビットが一体となりフリマジンによって検出される活性のある発光複合体を形成することができる。

10

【図7】対照として役立つC末端上での邪魔のない、または内部タグとしてポリペプチド間に置かれたNLpep114タグを伴った対象とする標的タンパク質の構成を示す図である。VEGFAは、この標的タンパク質が概念データの証明で役立つのでここで示される。対象とする可溶性標的タンパク質を使用することができる。HT(Halotag)、114(スモールビット)、VEGF(血管内皮増殖因子)、FLAG(FLAGオクタペプチド)、ATG1915:末端スモールビットを伴う対照標的; ATG1917: 2つの大きなドメイン間でスモールビットを伴う実験標的、及びATG1946:大型ドメイン1つと小型ドメイン1つの間でスモールビットを伴う実験標的。

20

【図8】抗体濃度依存性での生物発光の上昇によって測定されるようなNLpep114でタグ付けされたVEGFA標的タンパク質への抗VEGFA特異抗体の結合の検出による標的抗体特異的なNANOLUC生物発光相補性を示す図である。生物発光におけるこの上昇は、アイソタイプ対照が光を生じなかった(データは示さず)ので、抗体:標的特異的であった。

【図9】シグナル倍率/バックグラウンドとしての標的抗体特異的なNANOLUC生物発光相補性を示す図である。抗体が存在しないアッセイ成分のバックグラウンドを超える抗体の存在下での生物発光の上昇によって測定されるようなNLpep114でタグ付けされたVEGFA標的タンパク質に結合する抗VEGFA特異抗体の検出のシグナルウィンドウ。RLUにおけるバックグラウンドを超えるシグナルは、図8で得られたデータから算出し、使用した濃度範囲にわたって抗VEGFA抗体を増やすことに応答して75~450倍増大することが見いだされた。

30

【図10】生物発光の相補性の用量反応を介した抗VEGFA抗体の相対的な親和性によって示されるように、標的に対する抗体の親和性はNLpep114の位置によって変化しないことを示す図である。3つのVEGFA標的構築物についてのEC<sub>50</sub>値は図8で生成された用量反応曲線から算出されたように示される。

【図11】NLpep114でタグ付けされた標的タンパク質との内在性標的タンパク質の競合及びNLpoly11Sを融合したプロテインGと内部でタグ付けされたNLpep114標的タンパク質との間での抗体が主導するNANOLUC生物発光相補性を手段として、アッセイ成分をどのように使用して予測データ解析と共に内在性の標的タンパク質を定量するのかを示す模式図である。対象とする内在性の標的タンパク質の定量を可能にするように構成されたアッセイ成分。同じ標的タンパク質pep114融合体と11S-プロテインG融合体を用いて、内在性の標的タンパク質が増えるにつれてシグナルの減少を生じる結合競合を介して内在性の標的タンパク質の量を定量することができる。

40

【図12】内部タグ付けで使用される114直列ペプチドの模式図及び配列を示す。

【図13】FKBP/Frb融合タンパク質の模式図を示す。示されるのは、FKBPまたはFrbのC末端またはN末端に融合されたNLpoly11Sの融合体、及びFKBPまたはFrbの中にて異なる位置での内部タグ(NLpep114直列ペプチド=2x

50

N L p e p 1 1 4 ) の統合である。内部タグの統合部位の位置は、宿主タンパク質の隣接するアミノ酸位置によって示される(たとえば、A A 1 2 / 1 3 は F K B P のアミノ酸 1 2 と 1 3 の間での 2 x N L p e p 1 1 4 の統合を示す)。

【図 1 4】本明細書に記載されている内部タグ付けを用いた、ラパマイシンが誘導する F K B P / F r b のタンパク質 / タンパク質の相互作用を明らかにする図である。N L p e p 1 1 4 直列ペプチドの位置は、C 末端として ( C - 1 1 4 )、または宿主タンパク質の N 末端側で隣接するアミノ酸によって(たとえば、1 2 は F K B P 内の A A 1 2 と 1 3 の間での統合を示す)示される。

【図 1 5】図 1 5 で示した結果の無処理試料とラパマイシンで処理した試料との間の相対的な変化(応答比)への変換を示す図である。応答比は、方程式: 応答比 = R L U ラパマイシン / R L U 無処理を用いて算出される。

10

【図 1 6】本明細書に記載されている内部タグ付けを用いたラパマイシンによる F K B P / F r b 相互作用の用量依存性の誘導を示す図である。

【図 1 7】本明細書に記載されている内部タグ付けを用いたラパマイシン誘導の F K B P / F r b 相互作用の動態測定を示す図である。結果はすべて、対数尺度(左)または線形尺度(右)にてプロットした測定の単位としての相対光単位(R L U)を用いて示す。

【図 1 8】図 1 7 で示した結果(左のグラフ)の標準化した図を示す。標準化のために、各個々の分量について得られた最低値と最高値を参照点(それぞれ 0 % と 1 0 0 %)として使用した。

【0 0 4 2】

20

定義

本明細書で使用されるとき、用語「内部タグ」は別のポリペプチドまたはタンパク質の中に(たとえば、N 末端または C 末端ではない)挿入されるペプチドまたはポリペプチドの配列を指す。内部タグは、それがその中に挿入されるペプチドまたはポリペプチドの配列に検出、単離、局在化、会合等の 1 以上の特徴を提供してもよい。内部タグはポリペプチドまたはタンパク質の N - 及び C - 末端部分に直接接続されてもよいし、または 1 以上のリンカーによって接続されてもよい。一部の実施形態では、リンカー自体が官能性を提供してもよい。

【0 0 4 3】

本明細書で使用されるとき、用語「実質的に」は、引用される特色、パラメータ及び / または値が正確に達成される必要はないが、たとえば、忍容性、測定誤差、測定精度の限度及び当業者に既知の他の因子を含む偏差または変動が、特色が提供するように意図された効果を不可能にしない量で存在してもよいことを意味する。実質的に非存在である特色または特徴(たとえば、実質的に非発光性)は、使用されるアッセイのノイズの範囲内にある、バックグラウンドに達しない、または検出能を下回るものであってもよい。

30

【0 0 4 4】

本明細書で使用されるとき、用語「生物発光」は、酵素、タンパク質、タンパク質複合体または他の生体分子(たとえば、生物発光複合体)によって触媒される、または可能にされる化学反応による光の生成及び放射を指す。典型的な実施形態では、生物発光実体(たとえば、生物発光タンパク質または生物発光複合体)のための基質は生物発光実体によって高エネルギー反応生成物に変換され、反応生成物はそれに続いて、それがさらに安定な形態に変換するにつれて光を放射する。

40

【0 0 4 5】

本明細書で使用されるとき、用語「相補性の」は、互いにハイブリッド形成する、二量体化するまたはさもなくば複合体を形成することができる 2 以上の構造要素(たとえば、ペプチド、ポリペプチド、核酸、小分子等)の特色を指す。たとえば、「相補性のペプチドまたはポリペプチド」は一緒になって複合体を形成することができる。相補性の要素は、(たとえば、相互作用要素から)安定な複合体を形成するために、たとえば、相補性のために適正な配座に要素を置くために、相補性の要素を同時局在させるために、相補性のために相互作用エネルギーを低下させるために等で支援を必要としてもよい。一部の実

50

施形態では、「相補体配列」、「相補体」または「構造相補体」は、別の配列（たとえば、内部タグの）の構造相補体であるアミノ酸配列である。

【0046】

本明細書で使用されるとき、用語「複合体」は、互いに直接及び／または間接的に接触する分子（たとえば、ペプチドまたはポリペプチド）の集合体または凝集体を指す。態様の1つでは、「接触」、さらに詳しくは「直接接触」は、たとえば、ファンデルワールス力、水素結合、イオン性相互作用及び疎水性相互作用等のような分子間の非共有結合性の引力相互作用が分子の相互作用に影響を及ぼすように2以上の分子が十分に密接であることを意味する。本明細書で使用されるとき、用語「複合体」は、特に記載されない限り、2つ以上の分子（たとえば、ペプチド、ポリペプチドまたはそれらの組み合わせ）の集合体

10

【0047】

本明細書で使用されるとき、用語「非発光性の」は、可視スペクトルでの光としてのエネルギーを放射しない（たとえば、基質の存在下または非存在下で）特色を示す実体（たとえば、ペプチド、ポリペプチド、複合体、タンパク質等）を指す。所与のアッセイで実体が検出可能な発光を示さない場合、それは非発光性と呼ばれてもよい。本明細書で使用されるとき、用語「非発光性の」は用語「実質的に非発光性の」と同義である。特定のアッセイの意図する目的を妨害しないように発光が十分に最少限である場合、実体は「非発光性」である。

【0048】

20

本明細書で使用されるとき、用語「非発光性ペプチド」(NL pep)及び「非発光性ポリペプチド」(NL poly)は、標準的な条件下（たとえば、生理的条件、アッセイ条件等）でかつ典型的な計装（たとえば、ルミノメータ等）にて、（たとえば、基質の存在下または非存在下で）実質的に発光を示さない、または実際検出できない（たとえば、ノイズに満たない）量を示すペプチドまたはポリペプチド（たとえば、内部タグ、相補体配列等）を指す。一部の実施形態では、そのような非発光性のペプチド及びポリペプチドが、本明細書に記載されている基準に従って集まって生物発光複合体を形成する。本明細書で使用されるとき、「非発光性要素」は非発光性のペプチドまたは非発光性のポリペプチドである。用語「生物発光複合体」は2以上の非発光性のペプチド及び／または非発光性のポリペプチドの構築された複合体を指す。生物発光複合体は生物発光複合体の基質を高エネルギーの反応生成物に変換することを触媒し、または可能にし、反応生成物はその後、それがさらに安定な形態に変換するにつれて光を放射する。複合体化しない場合、生物発光複合体を形成する2つの非発光性要素は「非発光性対」と呼ばれてもよい。生物発光複合体が3以上の非発光性のペプチド及び／または非発光性のポリペプチドによって形成されるのであれば、生物発光複合体の複合体化しない構成成分は非発光性基と呼ばれてもよい。本明細書で使用されるとき、用語「非発光性複合体」は、生物発光複合体の基質を高エネルギーの反応生成物に変換することを実質的に触媒しない2以上の要素（たとえば、ペプチド、ポリペプチド等）の複合体を指す。一部の実施形態では、「非発光性複合体」は、発光複合体を形成するのに追加の非発光性要素（たとえば、第3の要素）を必要とする。

30

40

【0049】

本明細書で使用されるとき、用語「相互作用要素」は、非発光性要素（たとえば、内部タグ及び相補体配列）または、非発光性基（たとえば、内部タグ及び相補体配列）の対を一緒にして生物発光複合体を形成するのを支援する部分を指す。典型的な実施形態では、本発明はそのようなメカニズムに限定されず、メカニズムの理解は本発明を実践するのに必要とされないが、相互作用要素の対（別称「相互作用の対」）は非発光性要素の対（たとえば、非発光性のペプチド／ポリペプチドの対）に連結され、2つの相互作用要素の間での引力相互作用は生物発光複合体の形成を促す。相互作用要素は好適なメカニズム（たとえば、非発光性の対／基を極めて近接させること、非発光性の対／基を相互作用のための適正な配座に置くこと、複合体形成の活性化エネルギーを減らすこと、それらの組み合

50

わせ等)による生物発光複合体の形成を促してもよい。相互作用要素は、タンパク質、ポリペプチド、ペプチド、小分子、補因子、核酸、脂質、炭水化物、抗体、ポリマー、粒子等であってもよい。相互作用の対は2つの同一の相互作用要素(すなわち、ホモペア)または2つの異なる相互作用要素(すなわち、ヘテロペア)で作られてもよい。ヘテロペアの場合、相互作用要素は同じ種類の部分(たとえば、ポリペプチド)または異なる種類の部分(たとえば、ポリペプチドと小分子)であってもよい。相互作用の対による複合体形成が検討される一部の実施形態では、相互作用の対は「標的対」または「対象とする対」と呼ばれてもよく、個々の相互作用要素は「標的要素」(たとえば、「標的ペプチド」、「標的ポリペプチド」等)または対象とする要素(「対象とするペプチド」、「対象とするポリペプチド」等)と呼ばれる。

10

**【0050】**

本明細書で使用されるとき、用語「前から存在するタンパク質」は、特定の事象または日付に先立って物理的な存在にあったアミノ酸配列を指す。「前から存在するタンパク質の断片ではないペプチド」は、ペプチドの設計及び/または合成に先立って物理的な存在にあったタンパク質の(たとえば、合成のまたは天然に存在する)断片またはサブ配列ではない短いアミノ酸鎖である。

**【0051】**

本明細書で使用されるとき、用語「断片」は、さらに大きな全実体(たとえば、タンパク質、ポリペプチド、酵素等)の解離または「断片化」の結果生じるペプチドもしくはポリペプチド、またはそのようなものとして同じ配列を有するように調製されたペプチドもしくはポリペプチドを指す。したがって、断片は、それが作られ、及び/または設計される全実体(たとえば、タンパク質、ポリペプチド、酵素等)のサブ配列である。前から存在する全タンパク質のサブ配列ではないペプチドまたはポリペプチドは断片ではない(たとえば、前から存在するタンパク質の断片ではない)。「前から存在する生物発光タンパク質の断片ではない」ペプチドまたはポリペプチドは、(1)ペプチドまたはポリペプチドの設計及び/または合成に先立って物理的な存在にあった、かつ(2)実質的な生物発光活性を示すタンパク質(天然または合成)のサブ配列ではないアミノ酸鎖である。

20

**【0052】**

本明細書で使用されるとき、用語「サブ配列」は、別のさらに大きなペプチドまたはポリペプチドとの100%配列同一性を有するペプチドまたはポリペプチドを指す。サブ配列はさらに大きなアミノ酸鎖の一部についての完全な配列一致である。

30

**【0053】**

本明細書で使用されるとき、用語「配列同一性」は、2つのポリマー配列(たとえば、ペプチド、ポリペプチド、核酸等)がモノマーサブユニットの同一の連続的な組成を有する程度を指す。用語「配列類似性」は2つのポリマー配列(たとえば、ペプチド、ポリペプチド、核酸等)が類似のポリマー配列を有する程度を指す。たとえば、類似のアミノ酸は、同じ生物物理学的な特徴を共有し、かつファミリー、たとえば、酸性(たとえば、アスパラギン酸、グルタミン酸)、塩基性(たとえば、リジン、アルギニン、ヒスチジン)、非極性(たとえば、アラニン、バリン、ロイシン、イソロイシン、プロリン、フェニルアラニン、メチオニン、トリプトファン)、及び非荷電で極性(たとえば、グリシン、アスパラギン、グルタミン、システイン、セリン、スレオニン、チロシン)にグループ分けすることができるものである。「パーセント配列同一性」(または「パーセント配列類似性」)は、(1)比較のウィンドウ(たとえば、さらに長い配列の長さ、さらに短い配列の長さ、特定されたウィンドウ)にわたって2つの最適に並べた配列を比較し、(2)同一の(または類似の)モノマーを含有する位置の数を決定して(たとえば、双方の配列に同じアミノ酸が存在する、双方の配列に類似のアミノ酸が存在する)、一致した位置の数を得、(3)一致した位置の数を比較ウィンドウ(たとえば、さらに長い配列の長さ、さらに短い配列の長さ、特定されたウィンドウ)の位置の総数で割り、(4)結果に100を乗じてパーセント配列同一性またはパーセント配列類似性を得ることによって算出される。たとえば、ペプチドAとペプチドBが双方とも長さ20アミノ酸であり、1つの位置

40

50

を除いてすべての位置で同一アミノ酸を有するのであれば、ペプチドAとペプチドBは95%の配列同一性を有する。同一でない位置でのアミノ酸が同一の生物物理学的特徴（たとえば、双方とも酸性だった）を共有していたとすると、ペプチドAとペプチドBは100%の配列類似性を有することになる。別の例として、ペプチドCが長さ20アミノ酸であり、ペプチドDが長さ15アミノ酸であり、ペプチドDにおける15アミノ酸のうち14がペプチドCの一部のそれと同一であるとすると、ペプチドCとペプチドDは70%の配列同一性を有するが、ペプチドDはペプチドCの最適な比較ウィンドウに対して93.3%の配列同一性を有する。本明細書で「パーセント配列同一性」（または「パーセント配列類似性」）を算出する目的では、並べられた配列におけるギャップはその位置でのミスマッチとして処理される。

10

**【0054】**

本明細書で使用されるとき、用語「生理的条件」は、たとえば、生きている細胞と適合性のある温度、pH、塩分濃度、化学組成等の大部分の水性の条件等の、生きている細胞と適合性のあるいかなる条件も包含する。

**【0055】**

本明細書で使用されるとき、用語「試料」は最も広義で使用される。ある意味で、それは、任意の供給源と同様に生物試料及び環境試料から得られる検体または培養物を含むことにする。生物試料は動物（ヒトを含む）から得られてもよく、流体、固体、組織及び気体を包含する。生物試料には、血漿、血清等のような血液製剤が含まれる。試料はまた、細胞溶解物、または本明細書に記載されているペプチド及び／またはポリペプチドの精製された形態を指してもよい。細胞溶解物には、溶解剤で溶解されている細胞、または、たとえば、ウサギ網状赤血球もしくはコムギ胚芽の溶解物のような溶解物が含まれてもよい。試料には無細胞発現系も含まれてもよい。環境試料には、表面物質、土壌、水及び工業試料のような環境物質が挙げられる。しかしながら、そのような試料は本発明に適用できる試料の種類を限定すると解釈されるべきではない。

20

**【0056】**

本明細書で使用されるとき、特定されない限り、用語「ペプチド」及び「ポリペプチドの」はペプチドアミド結合（ $-C(O)NH-$ ）によって主鎖を介して連結される2以上のアミノ酸のポリマー化合物を指す。用語「ペプチド」は通常短いアミノ酸ポリマー（たとえば、25未満のアミノ酸を有する鎖）を指すのに対して用語「ポリペプチド」は通常長いアミノ酸ポリマー（たとえば、25を超えるアミノ酸を有する鎖）を指す。

30

**【0057】**

本明細書で使用されるとき、用語「融合体」、「融合ポリペプチド」及び「融合タンパク質」は、第2の異なるペプチド、ポリペプチドまたはタンパク質（たとえば、検出可能な配列、単離可能な配列、タグ等）に連結された対象とする第1のタンパク質またはポリペプチド（たとえば、標的配列等）を含有するキメラタンパク質を指す。用語「内部融合体」は本明細書で使用されるとき、第2のペプチド、ポリペプチドまたはタンパク質が第1の配列内のある位置に（たとえば、N末端またはC末端ではない）挿入される融合体を指す。用語「従来の融合体」は第1のポリペプチドまたはタンパク質と第2のペプチド、ポリペプチドまたはタンパク質が末端から末端までで融合される（たとえば、C末端からN末端まで、またはN末端からC末端まで）融合体を指す。

40

**【0058】**

本明細書で使用されるとき、用語「セレンテラジン」または「セレンテラジン基質」は天然に存在する（「ネイティブな」）セレンテラジンを指す。本明細書で使用されるとき、用語「セレンテラジン」または「セレンテラジン基質」は、その開示が全体として参照によって本明細書に組み入れられるWO2003/040100、米国特許出願番号12/056,073（段落[0086]）、及び米国特許第8,669,103号にて開示されたものに加えて、フリマジン、セレンテラジン-n、セレンテラジン-f、セレンテラジン-h、セレンテラジン-hcp、セレンテラジン-cp、セレンテラジン-c、セレンテラジン-e、セレンテラジン-fcp、ビス-デオキシセレンテラジン（「セレン

50

テラジン - h h」)、セレンテラジン - i、セレンテラジン - i c p、セレンテラジン - v、及び 2 - メチルセレンテラジンを含む、ネイティブなセレンテラジンと同様に合成の、たとえば、その誘導体または変異体及び天然の類似体を指す。

【 0 0 5 9 】

本明細書で使用されるとき、用語「低親和性」は、生理的条件またはアッセイ条件よりも実質的に高い（たとえば、2 倍、5 倍、1 0 倍、1 0 0 倍、1 0 0 0 倍以上）濃度を除いて、弱すぎて実体間で十分な複合体形成を生じない 2 つの実体（たとえば、タンパク質 / タンパク質）間の分子間相互作用を記載する。

【 0 0 6 0 】

本明細書で使用されるとき、用語「高親和性」は、生理的条件下またはアッセイ条件下で検出可能な複合体形成を生成するために十分な強度である 2 つの実体間の分子間相互作用を記載する。

【 発明を実施するための形態 】

【 0 0 6 1 】

レポーターによるタンパク質のタグ付けはタンパク質の機能及び挙動を分析するために一般的に使用される。一般に、遺伝子融合は、対象とするタンパク質の C 末端または N 末端を用いて生成される。しかしながら、特定の場合、双方の末端が対象とするタンパク質の機能に関わるので、タンパク質の生理的機能を変えことなく修飾することができない。本明細書に記載されている実施形態は、たとえば、N 末端または C 末端の修飾を必要とすることなくタンパク質 / タンパク質の相互作用の解析を可能にする。実施形態はさらに N 末端または C 末端の修飾を必要とすることなく、タンパク質の検出及び / または局在確認（たとえば、細胞性のまたは細胞内の局在確認）を可能にする。種々のタンパク質は構成の変化をもたらす修飾を受け、宿主タンパク質の構成に基づいた構造的相補性にアクセスできる内部ペプチド / ポリペプチドタグを使用することは、完全長のタンパク質を用いたバイオセンサーの生成を可能にする。したがって、本明細書で提供されるのは、対象とするタンパク質の中で内部に挿入される実質的に非発光性のペプチド / ポリペプチドタグである。内部に挿入されたタグと実質的に非発光性である相補体ポリペプチド / ペプチドとの相互作用により生物発光レポーター複合体の形成が生じる。

【 0 0 6 2 】

本明細書で提供されるのは、タンパク質またはポリペプチドの内部タグ（たとえば、ペプチドまたはポリペプチド）とその構造相補体（たとえば、遊離のまたは融合体（たとえば、内部または末端の））からの生物発光複合体の構築のための組成物及び方法である。一部の実施形態では、内部タグ及び / または構造相補体は前から存在するタンパク質の断片ではない（たとえば、既知のポリペプチド配列の相補性のサブ配列ではない）。特に、生物発光活性は、内部タグの実質的に非発光性のペプチド構造相補体の配列による構造的な相補性を介してタンパク質 / ポリペプチドの実質的に非発光性の内部タグに付与される。

【 0 0 6 3 】

一部の実施形態では、本明細書で提供されるのは、対象とするタンパク質 / ポリペプチドの存在を検出するために使用するための、及び分子の相互作用（たとえば、タンパク質 / タンパク質、タンパク質 / DNA、タンパク質 / RNA、タンパク質 / 小分子等）をモニタリングするための、実質的に非発光性の内部タグ及び構造相補体である。本明細書で提供されるのはまた、種々の生物発光複合体の形成の際に可変の親和性及び発光を有する相互交換可能な内部タグ及び構造相補体の配列（たとえば、ペプチド及びポリペプチド）の相補性パネルである（たとえば、高親和性 / 高発光の対、中程度の親和性 / 高発光の対、低親和性 / 中程度の発光の対等）。内部タグ及び構造相補体の様々な組み合わせを利用することは、低いものから高いものに及ぶ親和性、発光及び他の可変の特徴の種々の対を含む適合性の系を提供する。この適合性によって、対象とするタンパク質及びその分子相互作用の検出 / モニタリングが対象とする特定の分子（複数可）に微調整されるのが可能になり、モニターすることができる分子相互作用の範囲が非常に高いまたは非常に低い親

10

20

30

40

50

和性を伴う相互作用を含むように拡大される。さらに本明細書で提供されるのは、内部タグ、構造相補体及びそれらのパネルを開発し、調べる方法である。

【0064】

一部の実施形態では、内部タグと構造相補体単独との間での親和性は、活性のある（たとえば、生物発光性の）複合体を形成し、結果として生じるシグナル（たとえば、生物発光シグナル）を生成には不十分である。しかしながら、構造相補体が内部でタグ付けされたポリペプチドと相互作用する相互作用部分（たとえば、ペプチド、タンパク質、核酸、小分子等）に融合され、繋がれ、または連結されれば、そのときは、その相互作用（たとえば、対象とするポリペプチドと相互作用部分との複合体形成）は生物発光複合体の形成を促進する。そのような実施形態では、基質の存在下での生物発光複合体からのシグナルは、対象とするポリペプチドと統合部分（別称、相互作用複合体）との複合体の形成についての指標として役立つ。相互作用複合体が形成されれば、次いで生物発光複合体が形成され、生物発光シグナルが生成され、次いでそれは（たとえば、基質の存在下で）検出する／測定する／モニターすることができる。相互作用複合体が形成できなければ（たとえば、不都合な条件のせいで、相互作用要素間の不安定な相互作用のせいで、不適合な相互作用要素のせいで等）、そのときは安定な生物発光複合体は生じず、生物発光シグナルは生成されない。

10

【0065】

特定の実施形態では、内部でタグ付けされたポリペプチド及びそれと相互作用する（たとえば、複合体を形成する）第2の部分は相互作用対として知られる。一部の実施形態では、相互作用対は対象とする2つの分子（たとえば、対象とするタンパク質）を含む。一部の実施形態では、相互作用対の少なくとも一方のメンバーは内部でタグ付けされる。一部の実施形態では、相互作用対の双方のメンバーが内部でタグ付けされる（たとえば、構造的に相補性の内部タグによって）。一部の実施形態では、相互作用対の一方のメンバーは内部でタグ付けされ、他方は末端でタグ付けされる。たとえば、対象とするタンパク質に内部タグを挿入し、（たとえば、内部タグ、末端タグ等）対象とする分子を内部タグの構造相補体に繋ぐことによって対象とするタンパク質と対象とする第2の分子（たとえば、ペプチド、タンパク質、核酸、小分子等）との相互作用を検出するためにアッセイが実施される。対象とするタンパク質と対象とする分子が相互作用するのであれば（たとえば、一時的に相互作用する、安定して相互作用する、等）、内部タグと構造相補体は好適な配座で極めて近接して活性のある複合体（たとえば、生物発光複合体）を形成し、シグナルが生成され／検出される（たとえば、基質の存在下で）。対象とするタンパク質と対象とする分子との間での相互作用の非存在下では、内部タグと構造相補体は十分に安定な方法で相互作用せず、シグナルは生成されない、またはごく弱く生成されるにすぎない。そのような実施形態は、複合体形成に対する阻害剤の効果、複合体形成に対する変異の効果、複合体形成に対する条件（たとえば、温度、pH等）の効果、小分子（たとえば、治療剤候補）の標的分子との相互作用、等を検討するために使用される。

20

30

【0066】

一部の実施形態では、遊離の構造相補体との活性のある（たとえば、生物発光性の）複合体の形成によって、対象とする内部でタグ付けされたタンパク質がモニターされる（たとえば、検出される、局在確認される、等）。そのような実施形態では、内部タグ及び構造相補体は、内部でタグ付けされたタンパク質と遊離の構造相補体の双方が存在する場合、検出可能な複合体が生じるように互いに十分に高い親和性で選択される。

40

【0067】

異なる内部タグ及び構造相補体の対は、相互作用複合体の異なる強度、持続時間及び／または安定性を必要として活性のある（たとえば、生物発光性の）複合体の形成を生じてよい。一部の実施形態では、検出可能な（たとえば、生物発光性の）シグナルを生成するには安定な相互作用複合体が必要とされる。他の実施形態では、弱いまたは一時的な相互作用複合体でさえ活性のある（たとえば、生物発光性の）複合体の形成を生じる。一部の実施形態では、相互作用複合体の強度は結果として生じる（たとえば、生物発光性の）

50

シグナルの強度に直接比例する。一部の内部タグと構造相補体の対は、高いミリモル解離定数（たとえば、 $K_d > 100 \text{ mM}$ ）を持つ相互作用対（たとえば、対象とする内部でタグ付けされたタンパク質と相互作用相手）と組み合わせると検出可能なシグナルを生成する。他の内部タグと構造相補体の対は、検出可能なシグナルを伴った生物発光複合体を生成するために、低いミリモル（たとえば、 $K_d < 100 \text{ mM}$ ）、マイクロモル（たとえば、 $K_d < 1 \text{ mM}$ ）、ナノモル（たとえば、 $K_d < 1 \text{ }\mu\text{M}$ ）またはさらにピコモル（たとえば、 $K_d < 1 \text{ nM}$ ）の解離定数を持つ相互作用対を必要とする。さらに他の内部タグと構造相補体の対は、相互作用対の非存在下で活性のある複合体を形成する。

#### 【0068】

一部の実施形態では、内部タグと構造相補体の一方または双方は前から存在するタンパク質の断片ではない。一部の実施形態では、内部タグと構造相補体の一方または双方は前から存在する生物発光タンパク質の断片ではない。一部の実施形態では、内部タグも構造相補体も前から存在するタンパク質の断片ではない。一部の実施形態では、内部タグも構造相補体も前から存在する生物発光タンパク質の断片ではない。

10

#### 【0069】

一部の実施形態では、内部タグと構造相補体は双方とも単独では実質的に不活性（たとえば、非発光性）である。特定の実施形態では、好適な条件（たとえば、生理的条件下に置かれると、実質的に非発光性の内部タグ及び実質的に非発光性のその構造相補体は相互作用して生物発光複合体を形成し、基質の存在下で生物発光シグナルを生成する。一部の実施形態では、内部タグとその構造相補体は互いの存在下にて低レベルの活性（たとえば、生物発光）を生じるが、特定の設定の条件下で検出可能な活性（たとえば、生物発光）の有意な上昇を経験する。

20

#### 【0070】

一部の実施形態では、本明細書に記載されている組成物及び方法は1以上の相互作用要素を含む。典型的な実施形態では、相互作用要素は、内部タグの構造相補体に連結され、内部でタグ付けされたタンパク質との複合体を会合してまたは形成して、内部タグとその構造相補体の複合体の構築を促す部分（たとえば、ペプチド、ポリペプチド、タンパク質、小分子、核酸、脂質、炭水化物等）である。

#### 【0071】

一部の実施形態では、相互作用対は、内部でタグ付けされたタンパク質またはポリペプチドと、内部でタグ付けされたタンパク質またはポリペプチドと相互作用する他の好適な化学部分とを含んで、内部タグとその構造相補体との活性のある複合体の構築を促す。相互作用対は、たとえば、内部でタグ付けされたタンパク質と、核酸、ポリペプチド、タンパク質、リガンド、小分子、抗体、脂質等とからなってもよい。内部でタグ付けされたタンパク質またはポリペプチドと相互作用することが可能である分子実体は本明細書の一部の実施形態で使用されてもよい。

30

#### 【0072】

一部の実施形態では、本明細書の組成物及び方法は、標的分子（たとえば、内部でタグ付けされたタンパク質及び（潜在的な）相互作用相手）の対間での相互作用を検討するための有用なアッセイ（たとえば、試験管内、生体内、そのままで、動物全体等）を提供する。

40

#### 【0073】

一部の実施形態では、内部タグと相互作用対に連結された構造相補体との間での複合体（たとえば、生物発光複合体）の形成を促すために、内部でタグ付けされたタンパク質とその相互作用相手との間で相互作用を誘導するためにリガンド、基質、補因子等の存在が必要である。一部の実施形態では、生物発光複合体からのシグナルを検出することは、リガンド、基質、補因子等の存在を示す。

#### 【0074】

一部の実施形態では、内部タグ及びその構造相補体は単一のアミノ酸鎖（たとえば、N - （アミノ酸配列1） - （内部タグ） - （アミノ酸配列2） - （構造相補体） - C等）に

50

存在する。一部の実施形態では、対象とするタンパク質の折り畳みは活性のある複合体（たとえば、生物発光複合体）の形成を生じる。

【0075】

一部の実施形態では、内部でタグ付けされたタンパク質と、相互作用ペプチドまたはポリペプチド及び内部タグの構造相補体の融合体とが同一細胞内で発現する。そのような実施形態では、内部でタグ付けされたタンパク質と、相互作用ペプチドまたはポリペプチド及び内部タグの構造相補体の融合体とが細胞から精製され、及び／または単離され、または細胞内にて相互作用がアッセイされる。一部の実施形態では、内部でタグ付けされたタンパク質と、相互作用ペプチドまたはポリペプチド及び内部タグの構造相補体の融合体とが安定して発現する。一部の実施形態では、内部でタグ付けされたタンパク質と、相互作用ペプチドまたはポリペプチド及び内部タグの構造相補体の融合体とが一時的に発現する。他の実施形態では、内部でタグ付けされたタンパク質と、相互作用ペプチドまたはポリペプチド及び内部タグの構造相補体の融合体とが別々の細胞で発現し、シグナル検出のために（たとえば、精製及び／または単離に続いて）組み合わせられる。一部の実施形態では、内部でタグ付けされたタンパク質と、相互作用ペプチドまたはポリペプチド及び内部タグの構造相補体と融合体とが細胞溶解物（たとえば、ウサギ網状赤血球溶解物）または無細胞系で発現する。

10

【0076】

特定の実施形態では、本明細書に記載されているペプチド、ポリペプチド、融合ポリペプチド、融合タンパク質等をコードする核酸、DNA、RNA、ベクター等が提供される。そのような核酸及びベクターは発現、形質転換、形質移入、注入等に使用されてもよい。

20

【0077】

一部の実施形態では、内部タグは、リンカーによってポリペプチド配列に連結される（たとえば、そのN末端で、そのC末端で、両末端で）。一部の実施形態では、構造相補体はリンカーによって対象とする分子（たとえば、対象とするタンパク質）に連結される（たとえば、そのN末端で、そのC末端で、両末端で）。一部の実施形態では、リンカーは接続を提供し、要素間で所望の量の空間／間隔を可能にする。特定の実施形態では、リンカーは結合される要素間で適当な連結の化学的性質を提供する。一部の実施形態では、リンカーは2つの要素（たとえば、ペプチド、ポリペプチド、小分子等）を結合する、接続するまたは繋ぐことが可能である好適な化学部分である。一部の実施形態では、リンカーは、1以上の反復または非反復のモノマー単位のポリマー（たとえば、核酸、アミノ酸、炭素含有ポリマー、炭素鎖等）である。多種多様なリンカーが使用されてもよい。一部の実施形態では、リンカーは単一の共有結合である。一部の実施形態では、リンカーは、1～20の非水素原子（たとえば、C、N、P、O及びS）を有する線状または分岐状、環状または複素環状の、飽和または不飽和の構造を含み、アルキル、エーテル、チオエーテル、イミン、カルボン酸、アミン、エステル、カルボキサミド、スルホンアミド、ヒドラジドの結合及び芳香族またはヘテロ芳香族の結合の組み合わせで構成される。一部の実施形態では、リンカーは20の非水素原子より長い（たとえば、21の非水素原子、25の非水素原子、30の非水素原子、40の非水素原子、50の非水素原子、100の非水素原子等）。一部の実施形態では、リンカーは、（水素原子に加えて）C、N、P、O及びS（たとえば、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、または50の非水素原子）の群から選択される1～50の非水素原子を含む。

30

40

【0078】

本発明は利用できるリンカーの種類によって限定されない。シグナル要素と相互作用要素は直接（たとえば、単一の共有結合から成るリンカー）結合され、または好適なリンカーを介して結合される。本発明は特定のリンカー基に限定されない。種々のリンカー基が

50

熟考され、好適なリンカーは、アルキル基、メチレン炭素鎖、エーテル、ポリエーテル、アルキルアミドリンカー、ペプチドリンカー、修飾されたペプチドリンカー、ポリ（エチレングリコール）（PEG）リンカー、ストレプトアビジン／ビオチンリンカー、またはアビジン／ビオチンリンカー、ポリアミノ酸（たとえば、ポリリジン）、官能化PEG、多糖類、グリコサミノグリカン、樹状ポリマー（その全体が参照によって本明細書に組み入れられるWO93/06868及びTomaliaらによる、in Angew. Chem. Int. Ed. Engl. 29:138-175（1990）、PEG/キレート剤ポリマー（その全体が参照によって本明細書に組み入れられるW94/08629、W094/09056及びW096/26754）、オリゴヌクレオチドリンカー、リン脂質誘導体、アルケニル鎖、アルキニル鎖、ジスルフィドまたはそれらの組み合わせを含み得るが、これらに限定されない。

10

#### 【0079】

一部の実施形態では、リンカーは切断可能である（たとえば、酵素的に（たとえば、TEVプロテアーゼ部位）、化学的に、光誘導で等）。

#### 【0080】

一部の実施形態では、実質的に非発光性の内部タグはペプチド及び／またはポリペプチドの配列に直接結合される。一部の実施形態では、2以上の内部タグが対象とするポリペプチドの内部の位置に存在する。一部の実施形態では、1以上の内部タグがレポーター機能ではなくリンカー機能の役目を果たす。

#### 【0081】

20

一部の実施形態では、実質的に非発光性の内部タグ及びその構造相補体には、既存のルシフェラーゼ（その全体が参照によって本明細書に組み入れられる米国特許第8,557,970号、米国特許出願2014/0120548、米国特許第8,669,103号、米国特許出願14/160,278、及び米国特許出願14/160,282にて記載されたようなホタルルシフェラーゼ、Renillaルシフェラーゼ、Oplophorusルシフェラーゼ、増強Oplophorusルシフェラーゼ）の一部に対する100%未満の配列同一性及び／または配列類似性が提供される。本発明の特定の実施形態には、配列番号2157（たとえば、完全なNANOLUC配列）のすべてまたは一部（たとえば、>5、>8、>10、>12、>15、>20、<15、<18、<20、<22、<25、<30、<40、及びそれによって定義される範囲）との100%未満、しかし、30%を超える（たとえば、>30%、>40%、>45%、>50%、>55%、>60%、>65%、>70%、>75%、>80%、>85%、>90%、>95%、>98%、>99%）配列同一性を持つ実質的に非発光性の内部タグと実質的に非発光性の構造相補体との生物発光複合体の形成が関与する。一部の実施形態では、実質的に非発光性の内部タグ及び実質的に非発光性の構造相補体には、配列番号2157（たとえば、生物発光複合体と相互作用するペプチド及び／またはポリペプチド）の一部（たとえば、>5、>8、>10、>12、>15、>20、<15、<18、<20、<22、<25、<30、<40、及びそれによって定義される範囲）との100%未満の、しかし、30%を超える（たとえば、>30%、>40%、>45%、>50%、>55%、>60%、>65%、>70%、>75%、>80%、>85%、>90%、>95%、>98%、>99%）配列類似性が提供される。一部の実施形態では、配列番号2157の約25アミノ酸以下の部分との100%未満の、しかし、30%を超える（たとえば、>30%、>40%、>45%、>50%、>55%、>60%、>65%、>70%、>75%、>80%、>85%、>90%、>95%、>98%、>99%）配列同一性及び／または配列類似性を有する実質的に非発光性の内部タグと実質的に非発光性の構造相補体とが提供され、その際、そのようなペプチドは、適当な条件下（たとえば、相互作用対によって安定化される）にて、配列番号2157の別の部分との100%未満の、しかし、30%を超える（たとえば、>30%、>40%、>45%、>50%、>55%、>60%、>65%、>70%、>75%、>80%、>85%、>90%、>95%、>98%、>99%）配列同一性及び／または配列類似性を有するポリペプチドと組み合わ

30

40

50

せると生物発光複合体を形成する。同様に、配列番号 2 1 5 7 の一部との 1 0 0 % 未満の、しかし、4 0 % を超える（たとえば、> 4 0 %、> 4 5 %、> 5 0 %、> 5 5 %、> 6 0 %、> 6 5 %、> 7 0 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 5 %、> 9 8 %、> 9 9 %）配列同一性及び／または配列類似性を有する実質的に非発光性の内部タグと実質的に非発光性の構造相補体とが提供され、その際、そのような実質的に非発光性のタグと実質的に非発光性の構造相補体は、適当な条件下（たとえば、相互作用対によって安定化される）にて配列番号 2 1 5 7 の別の部分との 1 0 0 % 未満の、しかし、3 0 % を超える（たとえば、> 3 0 %、> 4 0 %、> 4 5 %、> 5 0 %、> 5 5 %、> 6 0 %、> 6 5 %、> 7 0 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 5 %、> 9 8 %、> 9 9 %）配列同一性及び／または配列類似性を有するペプチドと組み合わせると生物発光複合体を形成する。一部の実施形態では、配列番号 2 との 1 0 0 % 未満の、しかし、3 0 % を超える（たとえば、> 3 0 %、> 4 0 %、> 4 5 %、> 5 0 %、> 5 5 %、> 6 0 %、> 6 5 %、> 7 0 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 5 %、> 9 8 %、> 9 9 %）配列同一性及び／または配列類似性を持つ実質的に非発光性のタグと実質的に非発光性の構造相補体とが提供される。一部の実施形態では、配列番号 4 4 0 との 1 0 0 % 未満の、しかし、3 0 % を超える（たとえば、> 3 0 %、> 4 0 %、> 4 5 %、> 5 0 %、> 5 5 %、> 6 0 %、> 6 5 %、> 7 0 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 5 %、> 9 8 %、> 9 9 %）配列同一性及び／または配列類似性を持つ実質的に非発光性のタグと実質的に非発光性の構造相補体とが提供される。

10

**【 0 0 8 2 】**

20

一部の実施形態では、本発明の実施形態で使用される内部タグ及び／または構造相補体には、G V T G W R L C K R I L A（配列番号 2）に由来する 1 以上のアミノ酸の置換、欠失または付加を伴うペプチドが含まれる。一部の実施形態では、本明細書で提供されるのは、表 1 のアミノ酸配列、及び／または表 1 の核酸配列（表 1 のペプチド配列をコードする）を含む核酸を含む内部タグ及び／または構造相補体である。

**【 0 0 8 3 】**

表 1．例となる内部タグ及び／または構造相補体のペプチド配列

配列 番号	ペプチド番号	ポリマ ー	配列
3	NLpep2 (w/ Met)	N.A.	ATGGACGTGACCGGCTGGCGGCTGTGCGAACGCAT TCTGGCG
4	NLpep2 (w/ Met)	A.A.	MDVTGWRLCERILA
5	NLpep3 (w/ Met)	N.A.	ATGGGAGTGACCGCCTGGCGGCTGTGCGAACGCAT TCTGGCG
6	NLpep3 (w/ Met)	A.A.	MGVTAWRLCERILA
7	NLpep4 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCAT TCTGGCG
8	NLpep4 (w/ Met)	A.A.	MGVTGWRLCKRILA
9	NLpep5 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCGAACGCAT TAGCGCG
10	NLpep5 (w/ Met)	A.A.	MGVTGWRLCERISA
11	NLpep6 (w/ Met)	N.A.	ATGGACGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
12	NLpep6 (w/ Met)	A.A.	MDVTGWRLCKRISA
13	NLpep7 (w/ Met)	N.A.	ATGGACGTGACCGGCTGGCGGCTGTGCAAGCGCAT TCTGGCG
14	NLpep7 (w/ Met)	A.A.	MDVTGWRLCKRILA
15	NLpep8 (w/ Met)	N.A.	ATGGACGTGACCGGCTGGCGGCTGTGCGAACGCAT TAGCGCG
16	NLpep8 (w/ Met)	A.A.	MDVTGWRLCERISA
17	NLpep9 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
18	NLpep9 (w/ Met)	A.A.	MGVTGWRLCKRISA
19	NLpep10 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGAACGAACGCAT TCTGGCG

10

20

30

20	NLpep10 (w/ Met)	A.A.	MGVTGWRLNERILA
21	NLpep11 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGCAGGAACGCAT TCTGGCG
22	NLpep11 (w/ Met)	A.A.	MGVTGWRLQERILA
23	NLpep12 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGAAGAAGCGCC GGAGCCGG
24	NLpep12 (w/ Met)	A.A.	MGVTGWRLKKRRSR
25	NLpep13 (w/ Met)	N.A.	ATGAACGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
26	NLpep13 (w/ Met)	A.A.	MNVTGWRLCKRISA
27	NLpep14 (w/ Met)	N.A.	ATGAGCGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
28	NLpep14 (w/ Met)	A.A.	MSVTGWRLCKRISA
29	NLpep15 (w/ Met)	N.A.	ATGGAGGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
30	NLpep15 (w/ Met)	A.A.	MEVTGWRLCKRISA
31	NLpep16 (w/ Met)	N.A.	ATGGGCGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
32	NLpep16 (w/ Met)	A.A.	MHVTGWRLCKRISA
33	NLpep17 (w/ Met)	N.A.	ATGGGACACACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
34	NLpep17 (w/ Met)	A.A.	MGITGWRLCKRISA
35	NLpep18 (w/ Met)	N.A.	ATGGGAGCCACCGGCTGGCGGCTGTGCAAGCGCA TTAGCGCG
36	NLpep18 (w/ Met)	A.A.	MGATGWRLCKRISA
37	NLpep19 (w/ Met)	N.A.	ATGGGAAAGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
38	NLpep19 (w/ Met)	A.A.	MGKTGWRLCKRISA
39	NLpep20 (w/ Met)	N.A.	ATGGGACAGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG

10

20

30

40	NLpep20 (w/ Met)	A.A.	MGQTGWRLCKRISA
41	NLpep21 (w/ Met)	N.A.	ATGGGAAGCACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
42	NLpep21 (w/ Met)	A.A.	MGSTGWRLCKRISA
43	NLpep22 (w/ Met)	N.A.	ATGGGAGTGGTGGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
44	NLpep22 (w/ Met)	A.A.	MGVVGWRLCKRISA
45	NLpep23 (w/ Met)	N.A.	ATGGGAGTGAAGGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
46	NLpep23 (w/ Met)	A.A.	MGVKGWRLCKRISA
47	NLpep24 (w/ Met)	N.A.	ATGGGAGTGCAGGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
48	NLpep24 (w/ Met)	A.A.	MGVQGWRLCKRISA
49	NLpep25 (w/ Met)	N.A.	ATGGGAGTGACCGGCACCGGCTGTGCAAGCGCAT TAGCGCG
50	NLpep25 (w/ Met)	A.A.	MGVTGTRLCKRISA
51	NLpep26 (w/ Met)	N.A.	ATGGGAGTGACCGGCAAGCGGCTGTGCAAGCGCAT TAGCGCG
52	NLpep26 (w/ Met)	A.A.	MGVTGKRLCKRISA
53	NLpep27 (w/ Met)	N.A.	ATGGGAGTGACCGGCGTGCGGCTGTGCAAGCGCAT TAGCGCG
54	NLpep27 (w/ Met)	A.A.	MGVTGVRLCKRISA
55	NLpep28 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCACTGCAAGCGCAT TAGCGCG
56	NLpep28 (w/ Met)	A.A.	MGVTGWRICKRISA
57	NLpep29 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGGTGTGCAAGCGCAT TAGCGCG
58	NLpep29 (w/ Met)	A.A.	MGVTGWRVCKRISA
59	NLpep30 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGACCTGCAAGCGCAT TAGCGCG

10

20

30

60	NLpep30 (w/ Met)	A.A.	MGVTGWRTCKRISA
61	NLpep31 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTACTGCAAGCGCAT TAGCGCG
62	NLpep31 (w/ Met)	A.A.	MGVTGWRYCKRISA
63	NLpep32 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGAAGTGCAAGCGCAT TAGCGCG
64	NLpep32 (w/ Met)	A.A.	MGVTGWRKCKRISA
65	NLpep33 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGAACAAGCGCAT TAGCGCG
66	NLpep33 (w/ Met)	A.A.	MGVTGWRLNKRISA
67	NLpep34 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGACCAAGCGCAT TAGCGCG
68	NLpep34 (w/ Met)	A.A.	MGVTGWRLTKRISA
69	NLpep35 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGAAGAT TAGCGCG
70	NLpep35 (w/ Met)	A.A.	MGVTGWRLCKKISA
71	NLpep36 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGAACAT TAGCGCG
72	NLpep36 (w/ Met)	A.A.	MGVTGWRLCKNISA
73	NLpep37 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCGT GAGCGCG
74	NLpep37 (w/ Met)	A.A.	MGVTGWRLCKRVSA
75	NLpep38 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCC AGAGCGCG
76	NLpep38 (w/ Met)	A.A.	MGVTGWRLCKRQSA
77	NLpep39 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCG AGAGCGCG
78	NLpep39 (w/ Met)	A.A.	MGVTGWRLCKRESA
79	NLpep40 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCC GGAGCGCG

10

20

30

80	NLpep40 (w/ Met)	A.A.	MGVTGWRLCKRRSA
81	NLpep41 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCTT CAGCGCG
82	NLpep41 (w/ Met)	A.A.	MGVTGWRLCKRFSA
83	NLpep42 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCAAC
84	NLpep42 (w/ Met)	A.A.	MGVTGWRLCKRISN
85	NLpep43 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCACC
86	NLpep43 (w/ Met)	A.A.	MGVTGWRLCKRIST
87	NLpep44 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCCGG
88	NLpep44 (w/ Met)	A.A.	MGVTGWRLCKRISR
89	NLpep45 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCCTG
90	NLpep45 (w/ Met)	A.A.	MGVTGWRLCKRISL
91	NLpep46 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGAG
92	NLpep46 (w/ Met)	A.A.	MGVTGWRLCKRISE
93	NLpep47 (w/ Met)	N.A.	ATGGGAGTGACCGGCTTCCGGCTGTGCAAGCGCAT TAGCGCG
94	NLpep47 (w/ Met)	A.A.	MGVTGFRLCKRISA
95	NLpep48 (w/ Met)	N.A.	ATGGGAGTGACCGGCTACCGGCTGTGCAAGCGCAT TAGCGCG
96	NLpep48 (w/ Met)	A.A.	MGVTGYRLCKRISA
97	NLpep49(w/ Met)	N.A.	ATGGGAGTGACCGGCGAGCGGCTGTGCAAGCGCAT TAGCGCG
98	NLpep49(w/ Met)	A.A.	MGVTGERLCKRISA
99	NLpep50 (w/ Met)	N.A.	ATGCAGGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG

10

20

30

100	NLpep50 (w/ Met)	A.A.	MQVTGWRLCKRISA
101	NLpep51 (w/ Met)	N.A.	ATGACCGTGACCGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
102	NLpep51 (w/ Met)	A.A.	MTVTGWRLCKRISA
103	NLpep52 (w/ Met)	N.A.	ATGGGAGTGGAGGGCTGGCGGCTGTGCAAGCGCAT TAGCGCG
104	NLpep52 (w/ Met)	A.A.	MGVEGWRLCKRISA
105	NLpep53 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTTCAAGCGCAT TAGCGCG
106	NLpep53 (w/ Met)	A.A.	MGVTGWRLFKRISA
107	NLpep54 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTACAAGCGCAT TAGCGCG
108	NLpep54 (w/ Met)	A.A.	MGVTGWRLYKRISA
109	NLpep55 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGAGCAAGCGCAT TAGCGCG
110	NLpep55 (w/ Met)	A.A.	MGVTGWRLSKRISA
111	NLpep56 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGGGCAAGCGCA TTAGCGCG
112	NLpep56 (w/ Met)	A.A.	MGVTGWRLHKRISA
113	NLpep57 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGATGAAGCGCAT TAGCGCG
114	NLpep57 (w/ Met)	A.A.	MGVTGWRLMKRISA
115	NLpep58 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGGCCAAGCGCA TTAGCGCG
116	NLpep58 (w/ Met)	A.A.	MGVTGWRLAKRISA
117	NLpep59 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGCAGAAGCGCAT TAGCGCG
118	NLpep59 (w/ Met)	A.A.	MGVTGWRLQKRISA
119	NLpep60 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGCTGAAGCGCAT TAGCGCG

10

20

30

120	NLpep60 (w/ Met)	A.A.	MGVTGWRLKLRISA
121	NLpep61 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGAAGAAGCGCAT TAGCGCG
122	NLpep61 (w/ Met)	A.A.	MGVTGWRLKKRISA
123	NLpep62 (w/ Met)	N.A.	ATGAACCACACCGGCTGGCGGCTGAACAAGAAGGT GAGCAAC
124	NLpep62 (w/ Met)	A.A.	MNITGWRLNKKVSN
125	NLpep63 (w/ Met)	N.A.	ATGAACCACACCGGCTACCGGCTGAACAAGAAGGT GAGCAAC
126	NLpep63 (w/ Met)	A.A.	MNITGYRLNKKVSN
127	NLpep64 (w/ Met)	N.A.	ATGTGCGTGACCGGCTGGCGGCTGTTCAAGCGCAT TAGCGCG
128	NLpep64 (w/ Met)	A.A.	MCVTGWRLFKRISA
129	NLpep65 (w/ Met)	N.A.	ATGCCCGTGACCGGCTGGCGGCTGTTCAAGCGCAT TAGCGCG
130	NLpep65 (w/ Met)	A.A.	MPVTGWRLFKRISA
131	NLpep66 (w/ Met)	N.A.	ATGAACCACACCGGCTACCGGCTGTTCAAGAAGGT GAGCAAC
132	NLpep66 (w/ Met)	A.A.	MNITGYRLFKKVSN
133	NLpep67 (w/ Met)	N.A.	ATGAACGTGACCGGCTACCGGCTGTTCAAGAAGGT GAGCAAC
134	NLpep67 (w/ Met)	A.A.	MNVTGYRLFKKVSN
135	NLpep68 (w/ Met)	N.A.	ATGAACGTGACCGGCTGGCGGCTGTTCAAGAAGGT GAGCAAC
136	NLpep68 (w/ Met)	A.A.	MNVTGWRLFKKVSN
137	NLpep69 (w/ Met)	N.A.	ATGAACGTGACCGGCTGGCGGCTGTTCAAGAAGAT TAGCAAC
138	NLpep69 (w/ Met)	A.A.	MNVTGWRLFKKISN
139	NLpep70 (w/ Met)	N.A.	ATGAACGTGACCGGCTGGCGGCTGTTCAAGCGCAT TAGCAAC

10

20

30

140	NLpep70 (w/ Met)	A.A.	MNVTGWRLFKRISN
141	NLpep71 (w/ Met)	N.A.	ATGGGAGTGACCGGCTGGCGGCTGTTCAAGCGCAT TAGCAAC
142	NLpep71 (w/ Met)	A.A.	MGVTGWRLFKRISN
143	NLpep72 (w/ Met)	N.A.	ATGAACGTGACCGGCTGGCGGCTGTTCAAGCGCAT TAGCAAC
144	NLpep72 (w/ Met)	A.A.	MNVTGWRLFERISN
145	NLpep73 (w/ Met)	N.A.	ATGAACGTGACCGGCTGGCGGCTGTTCAAGCGCAT TCTGAAC
146	NLpep73 (w/ Met)	A.A.	MNVTGWRLFKRILN
147	NLpep74 (w/ Met)	N.A.	ATGAACGTGACCGGCTGGCGGCTGTTCAAGCGCAT TAGCGCG
148	NLpep74 (w/ Met)	A.A.	MNVTGWRLFKRISA
149	NLpep75 (w/ Met)	N.A.	ATGAACGTGACCGGCTGGCGGCTGTTCAAGAAAGAT TAGCAAC
150	NLpep75 (w/ Met)	A.A.	MNVTGWRLFEEKISN
151	NLpep76 (w/ Met)	N.A.	ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAAGAT TAGCAAC
152	NLpep76 (w/ Met)	A.A.	MNVSGWRLFEEKISN
153	NLpep77 (w/ Met)	N.A.	ATG-GTGACCGGCTGGCGGCTGTTCAAGAAGATT GCAAC
154	NLpep77 (w/ Met)	A.A.	M-VTGWRLFKKISN
155	NLpep78 (w/ Met)	N.A.	ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGAT TAGCAAC
156	NLpep78 (w/ Met)	A.A.	MNVSGWRLFKKISN
157	NLpep79 (w/ Met)	N.A.	ATGAACGTGACCGGCTACCGGCTGTTCAAGAAGAT TAGCAAC
158	NLpep79 (w/ Met)	A.A.	MNVTGYRLFKKISN
159	NLpep80(w/ Met)	N.A.	ATGGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAG CAAC

10

20

30

160	NLpep80(w/ Met)	A.A.	MVSGWRLFKKISN
161	NLpep81 (w/ Met)	N.A.	ATGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAAC
440	NLpep81 (w/ Met)	A.A.	MSGWRLFKKISN
163	NLpep82 (w/ Met)	N.A.	ATGGGCTGGCGGCTGTTCAAGAAGATTAGCAAC
164	NLpep82 (w/ Met)	A.A.	MGWRLFKKISN
165	NLpep83 (w/ Met)	N.A.	ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGC
166	NLpep83 (w/ Met)	A.A.	MNVSGWRLFKKIS
167	NLpep84 (w/ Met)	N.A.	ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATT
168	NLpep84 (w/ Met)	A.A.	MNVSGWRLFKKI
169	NLpep85 (w/ Met)	N.A.	ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAG
170	NLpep85 (w/ Met)	A.A.	MNVSGWRLFKK
171	NLpep86 (w/ Met)	N.A.	ATGGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGC
172	NLpep86 (w/ Met)	A.A.	MVSGWRLFKKIS
173	NLpep87 (w/ Met)	N.A.	ATGAGCGGCTGGCGGCTGTTCAAGAAGATT
174	NLpep87 (w/ Met)	A.A.	MSGWRLFKKI
175	NLpep88 (w/ Met)	N.A.	ATGAACGTGAGCGGCTGGGGCCTGTTCAAGAAGATTAGCAAC
176	NLpep88 (w/ Met)	A.A.	MNVSGWGLFKKISN
177	NLpep89 (w/ Met)	N.A.	ATGCCCGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAAC
178	NLpep89 (w/ Met)	A.A.	MPVSGWRLFKKISN
179	NLpep90 (w/ Met)	N.A.	ATGAACCCCGTGAGCGGCTGGCGGCTGTTCAAGAAATTAGCAAC

10

20

30

180	NLpep90 (w/ Met)	A.A.	MNPVSGWRLFKKISN
181	NLpep91 (w/ Met)	N.A.	ATGATCAACCCCGTGAGCGGCTGGCGGCTGTTCAA GAAGATTAGCAAC
182	NLpep91 (w/ Met)	A.A.	MINPVSGWRLFKKISN
183	NLpep92 (w/ Met)	N.A.	ATGACCATCAACCCCGTGAGCGGCTGGCGGCTGTT CAAGAAGATTAGCAAC
184	NLpep92 (w/ Met)	A.A.	MTINPVSGWRLFKKISN
185	NLpep93 (w/ Met)	N.A.	ATGGTGACCATCAACCCCGTGAGCGGCTGGCGGCT GTTCAAGAAGATTAGCAAC
186	NLpep93 (w/ Met)	A.A.	MVTINPVSGWRLFKKISN
187	NLpep94 (w/ Met)	N.A.	ATGCGGGTGACCATCAACCCCGTGAGCGGCTGGCG GCTGTTCAAGAAGATTAGCAAC
188	NLpep94 (w/ Met)	A.A.	MRVTINPVSGWRLFKKISN
189	NLpep95 (w/ Met)	N.A.	ATGAGCGGCTGGCGGCTGCTGAAGAAGATT
190	NLpep95 (w/ Met)	A.A.	MSGWRLKKI
191	NLpep96 (w/ Met)	N.A.	ATGACCGGCTACCGGCTGCTGAAGAAGATT
192	NLpep96 (w/ Met)	A.A.	MTGYRLKKI
193	NLpep97 (w/ Met)	N.A.	ATGAGCGGCTGGCGGCTGTTCAAGAAG
194	NLpep97 (w/ Met)	A.A.	MSGWRLFKK
195	NLpep98 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGTTCAAGAAGATTAG C
196	NLpep98 (w/ Met)	A.A.	MVTGYRLFKKIS
197	NLpep99 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGTTGAGAGAAGATTAG C
198	NLpep99 (w/ Met)	A.A.	MVTGYRLFEEKIS
199	NLpep100(w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGTTGAGCAGATTAG C

10

20

30

200	NLpep100(w/ Met)	A.A.	MVTGYRLFQIS
201	NLpep101(w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGTTCGAGAAGGAGAG C
202	NLpep101(w/ Met)	A.A.	MVTGYRLFQES
203	NLpep102(w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGTTCGAGCAGGAGAG C
204	NLpep102(w/ Met)	A.A.	MVTGYRLFQES
205	NLpep103(w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGTTCGAGCAGGAGCT G
206	NLpep103(w/ Met)	A.A.	MVTGYRLFQEL
207	NLpep104(w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAGAAGATTAG C
208	NLpep104(w/ Met)	A.A.	MVEGYRLFQIS
209	NLpep105(w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAGCAGATTAG C
210	NLpep105(w/ Met)	A.A.	MVEGYRLFQIS
211	NLpep106(w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAGAAGGAGAG C
212	NLpep106(w/ Met)	A.A.	MVEGYRLFQES
213	NLpep107(w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAGCAGGAGAG C
214	NLpep107(w/ Met)	A.A.	MVEGYRLFQES
215	NLpep108(w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAGCAGGAGCT G
216	NLpep108(w/ Met)	A.A.	MVEGYRLFQEL
217	NLpep109(w/ Met)	N.A.	ATGATTAGCGGCTGGCGGCTGATGAAGAACATTAG C
218	NLpep109(w/ Met)	A.A.	MISGWRLMKNIS
219	NLpep110(w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAAGAAGATTAG C

10

20

30

220	NLpep110(w/ Met)	A.A.	MVEGYRLFKKIS
221	NLpep2 (w/o Met)	N.A.	GACGTGACCGGCTGGCGGCTGTGCGAACGCATTCT GGCG
222	NLpep2 (w/o Met)	A.A.	DVTGWRLCERILA
223	NLpep3 (w/o Met)	N.A.	GGAGTGACCGCCTGGCGGCTGTGCGAACGCATTCT GGCG
224	NLpep3 (w/o Met)	A.A.	GVTAWRLCERILA
225	NLpep4 (w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCATTCT GGCG
226	NLpep4 (w/o Met)	A.A.	GVTGWRLCKRILA
227	NLpep5 (w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCGAACGCATTAG CGCG
228	NLpep5 (w/o Met)	A.A.	GVTGWRLCERISA
229	NLpep6 (w/o Met)	N.A.	GACGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
230	NLpep6 (w/o Met)	A.A.	DVTGWRLCKRISA
231	NLpep7 (w/o Met)	N.A.	GACGTGACCGGCTGGCGGCTGTGCAAGCGCATTCT GGCG
232	NLpep7 (w/o Met)	A.A.	DVTGWRLCKRILA
233	NLpep8 (w/o Met)	N.A.	GACGTGACCGGCTGGCGGCTGTGCGAACGCATTAG CGCG
234	NLpep8 (w/o Met)	A.A.	DVTGWRLCERISA
235	NLpep9 (w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
236	NLpep9 (w/o Met)	A.A.	GVTGWRLCKRISA
237	NLpep10(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGAACGAACGCATTCT GGCG
238	NLpep10(w/o Met)	A.A.	GVTGWRLNERILA
239	NLpep11(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGCAGGAACGCATTCT GGCG

10

20

30

240	NLpep11(w/o Met)	A.A.	GVTGWRLQERILA
241	NLpep12(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGAAGAAGCGCCGGA GCCGG
242	NLpep12(w/o Met)	A.A.	GVTGWRLKKRRSR
243	NLpep13(w/o Met)	N.A.	AACGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
244	NLpep13(w/o Met)	A.A.	NVTGWRLCKRISA
245	NLpep14(w/o Met)	N.A.	AGCGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
246	NLpep14(w/o Met)	A.A.	SVTGWRLCKRISA
247	NLpep15(w/o Met)	N.A.	GAGGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
248	NLpep15(w/o Met)	A.A.	EVTGWRLCKRISA
249	NLpep16(w/o Met)	N.A.	GGCGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
250	NLpep16(w/o Met)	A.A.	HVTGWRLCKRISA
251	NLpep17(w/o Met)	N.A.	GGACACACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
252	NLpep17(w/o Met)	A.A.	GITGWRLCKRISA
253	NLpep18(w/o Met)	N.A.	GGAGCCACCGGCTGGCGGCTGTGCAAGCGCATTAG GCGCG
254	NLpep18(w/o Met)	A.A.	GATGWRLCKRISA
255	NLpep19(w/o Met)	N.A.	GGAAAGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
256	NLpep19(w/o Met)	A.A.	GKTGWRLCKRISA
257	NLpep20(w/o Met)	N.A.	GGACAGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
258	NLpep20(w/o Met)	A.A.	GQTGWRLCKRISA
259	NLpep21(w/o Met)	N.A.	GGAAGCACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG

10

20

30

260	NLpep21(w/o Met)	A.A.	GSTGWRLCKRISA
261	NLpep22(w/o Met)	N.A.	GGAGTGGTGGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
262	NLpep22(w/o Met)	A.A.	GVVGWRLCKRISA
263	NLpep23(w/o Met)	N.A.	GGAGTGAAGGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
264	NLpep23(w/o Met)	A.A.	GVKGWRLCKRISA
265	NLpep24(w/o Met)	N.A.	GGAGTGCAGGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
266	NLpep24(w/o Met)	A.A.	GVQGWRLCKRISA
267	NLpep25(w/o Met)	N.A.	GGAGTGACCGGCACCCGGCTGTGCAAGCGCATTAG CGCG
268	NLpep25(w/o Met)	A.A.	GVTGTRLCKRISA
269	NLpep26(w/o Met)	N.A.	GGAGTGACCGGCAAGCGGCTGTGCAAGCGCATTAG CGCG
270	NLpep26(w/o Met)	A.A.	GVTGKRLCKRISA
271	NLpep27(w/o Met)	N.A.	GGAGTGACCGGCGTGCGGCTGTGCAAGCGCATTAG CGCG
272	NLpep27(w/o Met)	A.A.	GVTGVRLCKRISA
273	NLpep28(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCACTGCAAGCGCATTAG CGCG
274	NLpep28(w/o Met)	A.A.	GVTGWRICKRISA
275	NLpep29(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGGTGTGCAAGCGCATTAG CGCG
276	NLpep29(w/o Met)	A.A.	GVTGWRVCKRISA
277	NLpep30(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGACCTGCAAGCGCATTAG CGCG
278	NLpep30(w/o Met)	A.A.	GVTGWRTCKRISA
279	NLpep31(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGTACTGCAAGCGCATTAG CGCG

10

20

30

280	NLpep31(w/o Met)	A.A.	GVTGWRYCKRISA
281	NLpep32(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGAAGTGCAAGCGCATTAG CGCG
282	NLpep32(w/o Met)	A.A.	GVTGWRKCKRISA
283	NLpep33(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGAACAAGCGCATTAG CGCG
284	NLpep33(w/o Met)	A.A.	GVTGWRLNKRISA
285	NLpep34(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGACCAAGCGCATTAG CGCG
286	NLpep34(w/o Met)	A.A.	GVTGWRLTKRISA
287	NLpep35(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGAAGATTAG CGCG
288	NLpep35(w/o Met)	A.A.	GVTGWRLCKKISA
289	NLpep36(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGAACATTAG CGCG
290	NLpep36(w/o Met)	A.A.	GVTGWRLCKNISA
291	NLpep37(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCGTGA GCGCG
292	NLpep37(w/o Met)	A.A.	GVTGWRLCKRVSA
293	NLpep38(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCCAGA GCGCG
294	NLpep38(w/o Met)	A.A.	GVTGWRLCKRQSA
295	NLpep39(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCGAGA GCGCG
296	NLpep39(w/o Met)	A.A.	GVTGWRLCKRESA
297	NLpep40(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCCGGA GCGCG
298	NLpep40(w/o Met)	A.A.	GVTGWRLCKRRSA
299	NLpep41(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCTTCAG CGCG

10

20

30

300	NLpep41(w/o Met)	A.A.	GVTGWRLCKRFSA
301	NLpep42(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CAAC
302	NLpep42(w/o Met)	A.A.	GVTGWRLCKRISN
303	NLpep43(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CACC
304	NLpep43(w/o Met)	A.A.	GVTGWRLCKRIST
305	NLpep44(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CCGG
306	NLpep44(w/o Met)	A.A.	GVTGWRLCKRISR
307	NLpep45(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CCTG
308	NLpep45(w/o Met)	A.A.	GVTGWRLCKRISL
309	NLpep46(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGAG
310	NLpep46(w/o Met)	A.A.	GVTGWRLCKRISE
311	NLpep47(w/o Met)	N.A.	GGAGTGACCGGCTTCCGGCTGTGCAAGCGCATTAG CGCG
312	NLpep47(w/o Met)	A.A.	GVTGFRLCKRISA
313	NLpep48(w/o Met)	N.A.	GGAGTGACCGGCTACCGGCTGTGCAAGCGCATTAG CGCG
314	NLpep48(w/o Met)	A.A.	GVTGYRLCKRISA
315	NLpep49(w/o Met)	N.A.	GGAGTGACCGGCGAGCGGCTGTGCAAGCGCATTAG CGCG
316	NLpep49(w/o Met)	A.A.	GVTGERLCKRISA
317	NLpep50(w/o Met)	N.A.	CAGGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
318	NLpep50(w/o Met)	A.A.	QVTGWRLCKRISA
319	NLpep51(w/o Met)	N.A.	ACCGTGACCGGCTGGCGGCTGTGCAAGCGCATTAG CGCG

10

20

30

320	NLpep51(w/o Met)	A.A.	TVTGWRLCKRISA
321	NLpep52(w/o Met)	N.A.	GGAGTGGAGGGCTGGCGGCTGTGCAAGCGCATTAG CGCG
322	NLpep52(w/o Met)	A.A.	GVEGWRLCKRISA
323	NLpep53(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTTCAAGCGCATTAG CGCG
324	NLpep53(w/o Met)	A.A.	GVTGWRLFKRISA
325	NLpep54(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTACAAGCGCATTAG CGCG
326	NLpep54(w/o Met)	A.A.	GVTGWRLYKRISA
327	NLpep55(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGAGCAAGCGCATTAG CGCG
328	NLpep55(w/o Met)	A.A.	GVTGWRLSKRISA
329	NLpep56(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGGGCAAGCGCATT GCGCG
330	NLpep56(w/o Met)	A.A.	GVTGWRLHKRISA
331	NLpep57(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGATGAAGCGCATTAG CGCG
332	NLpep57(w/o Met)	A.A.	GVTGWRLMKRISA
333	NLpep58(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGGCCAAGCGCATT GCGCG
334	NLpep58(w/o Met)	A.A.	GVTGWRLAKRISA
335	NLpep59(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGCAGAAGCGCATTAG CGCG
336	NLpep59(w/o Met)	A.A.	GVTGWRLQKRISA
337	NLpep60(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGCTGAAGCGCATTAG CGCG
338	NLpep60(w/o Met)	A.A.	GVTGWRLLRISA
339	NLpep61(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGAAGAAGCGCATTAG CGCG

10

20

30

340	NLpep61(w/o Met)	A.A.	GVTGWRLKKRISA
341	NLpep62(w/o Met)	N.A.	AACCACACCGGCTGGCGGCTGAACAAGAAGGTGAG CAAC
342	NLpep62(w/o Met)	A.A.	NITGWRLNKKVSN
343	NLpep63(w/o Met)	N.A.	AACCACACCGGCTACCGGCTGAACAAGAAGGTGAG CAAC
344	NLpep63(w/o Met)	A.A.	NITGYRLNKKVSN
345	NLpep64(w/o Met)	N.A.	TGCGTGACCGGCTGGCGGCTGTTCAAGCGCATTAG CGCG
346	NLpep64(w/o Met)	A.A.	CVTGWRLFKRISA
347	NLpep65(w/o Met)	N.A.	CCCGTGACCGGCTGGCGGCTGTTCAAGCGCATTAG CGCG
348	NLpep65(w/o Met)	A.A.	PVTGWRLFKRISA
349	NLpep66(w/o Met)	N.A.	AACCACACCGGCTACCGGCTGTTCAAGAAGGTGAG CAAC
350	NLpep66(w/o Met)	A.A.	NITGYRLFKKVS
351	NLpep67(w/o Met)	N.A.	AACGTGACCGGCTACCGGCTGTTCAAGAAGGTGAG CAAC
352	NLpep67(w/o Met)	A.A.	NVTGYRLFKKVS
353	NLpep68(w/o Met)	N.A.	AACGTGACCGGCTGGCGGCTGTTCAAGAAGGTGAG CAAC
354	NLpep68(w/o Met)	A.A.	NVTGWRLFKKVS
355	NLpep69(w/o Met)	N.A.	AACGTGACCGGCTGGCGGCTGTTCAAGAAGATTAG CAAC
356	NLpep69(w/o Met)	A.A.	NVTGWRLFKKIS
357	NLpep70(w/o Met)	N.A.	AACGTGACCGGCTGGCGGCTGTTCAAGCGCATTAG CAAC
358	NLpep70(w/o Met)	A.A.	NVTGWRLFKRIS
359	NLpep71(w/o Met)	N.A.	GGAGTGACCGGCTGGCGGCTGTTCAAGCGCATTAG CAAC

10

20

30

360	NLpep71(w/o Met)	A.A.	GVTGWRLFKRISN
361	NLpep72(w/o Met)	N.A.	AACGTGACCGGCTGGCGGCTGTTCAAGCGCATTAG CAAC
362	NLpep72(w/o Met)	A.A.	NVTGWRLFERISN
363	NLpep73(w/o Met)	N.A.	AACGTGACCGGCTGGCGGCTGTTCAAGCGCATTCT GAAC
364	NLpep73(w/o Met)	A.A.	NVTGWRLFKRILN
365	NLpep74(w/o Met)	N.A.	AACGTGACCGGCTGGCGGCTGTTCAAGCGCATTAG CGCG
366	NLpep74(w/o Met)	A.A.	NVTGWRLFKRISA
367	NLpep75(w/o Met)	N.A.	AACGTGACCGGCTGGCGGCTGTTCAAGAAAGATTAG CAAC
368	NLpep75(w/o Met)	A.A.	NVTGWRLFEKISN
369	NLpep76(w/o Met)	N.A.	AACGTGAGCGGCTGGCGGCTGTTCAAGAAAGATTAG CAAC
370	NLpep76(w/o Met)	A.A.	NVSGWRLFEKISN
371	NLpep77(w/o Met)	N.A.	GTGACCGGCTGGCGGCTGTTCAAGAAGATTAGCAA C
372	NLpep77(w/o Met)	A.A.	VTGWRLFKKISN
373	NLpep78(w/o Met)	N.A.	AACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAG CAAC
374	NLpep78(w/o Met)	A.A.	NVSGWRLFKKISN
375	NLpep79(w/o Met)	N.A.	AACGTGACCGGCTACCGGCTGTTCAAGAAGATTAG CAAC
376	NLpep79(w/o Met)	A.A.	NVTGYRLFKKISN
377	NLpep80(w/o Met)	N.A.	GTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAA C
378	NLpep80(w/o Met)	A.A.	VSGWRLFKKISN
379	NLpep81(w/o Met)	N.A.	AGCGGCTGGCGGCTGTTCAAGAAGATTAGCAAC

10

20

30

380	NLpep81(w/o Met)	A.A.	SGWRLFKKISN
381	NLpep82(w/o Met)	N.A.	GGCTGGCGGCTGTTCAAGAAGATTAGCAAC
382	NLpep82(w/o Met)	A.A.	GWRLFKKISN
383	NLpep83(w/o Met)	N.A.	AACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAG C
384	NLpep83(w/o Met)	A.A.	NVSGWRLFKKIS
385	NLpep84(w/o Met)	N.A.	AACGTGAGCGGCTGGCGGCTGTTCAAGAAGATT
386	NLpep84(w/o Met)	A.A.	NVSGWRLFKKI
387	NLpep85(w/o Met)	N.A.	AACGTGAGCGGCTGGCGGCTGTTCAAGAAG
388	NLpep85(w/o Met)	A.A.	NVSGWRLFKK
389	NLpep86(w/o Met)	N.A.	GTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGC
390	NLpep86(w/o Met)	A.A.	VSGWRLFKKIS
391	NLpep87(w/o Met)	N.A.	AGCGGCTGGCGGCTGTTCAAGAAGATT
392	NLpep87(w/o Met)	A.A.	SGWRLFKKI
393	NLpep88(w/o Met)	N.A.	AACGTGAGCGGCTGGGGCCTGTTCAAGAAGATTAG CAAC
394	NLpep88(w/o Met)	A.A.	NVSGWGLFKKISN
395	NLpep89(w/o Met)	N.A.	CCCGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAG CAAC
396	NLpep89(w/o Met)	A.A.	PVSGWRLFKKISN
397	NLpep90(w/o Met)	N.A.	AACCCCGTGAGCGGCTGGCGGCTGTTCAAGAAGAT TAGCAAC
398	NLpep90(w/o Met)	A.A.	NPVSGWRLFKKISN
399	NLpep91(w/o Met)	N.A.	ATCAACCCCGTGAGCGGCTGGCGGCTGTTCAAGAA GATTAGCAAC

10

20

30

400	NLpep91(w/o Met)	A.A.	INPVSGWRLFKKISN
401	NLpep92(w/o Met)	N.A.	ACCATCAACCCCGTGAGCGGCTGGCGGCTGTTCAA GAAGATTAGCAAC
402	NLpep92(w/o Met)	A.A.	TINPVSGWRLFKKISN
403	NLpep93(w/o Met)	N.A.	GTGACCATCAACCCCGTGAGCGGCTGGCGGCTGTT CAAGAAGATTAGCAAC
404	NLpep93(w/o Met)	A.A.	VTINPVSGWRLFKKISN
405	NLpep94(w/o Met)	N.A.	CGGGTGACCATCAACCCCGTGAGCGGCTGGCGGC TGTTCAAGAAGATTAGCAAC
406	NLpep94(w/o Met)	A.A.	RVTINPVSGWRLFKKISN
407	NLpep95(w/o Met)	N.A.	AGCGGCTGGCGGCTGCTGAAGAAGATT
408	NLpep95(w/o Met)	A.A.	SGWRLKKI
409	NLpep96(w/o Met)	N.A.	ACCGGCTACCGGCTGCTGAAGAAGATT
410	NLpep96(w/o Met)	A.A.	TGYRLKKI
411	NLpep97(w/o Met)	N.A.	AGCGGCTGGCGGCTGTTCAAGAAG
412	NLpep97(w/o Met)	A.A.	SGWRLFKK
413	NLpep98(w/o Met)	N.A.	GTGACCGGCTACCGGCTGTTCAAGAAGATTAGC
414	NLpep98(w/o Met)	A.A.	VTGYRLFKKIS
415	NLpep99(w/o Met)	N.A.	GTGACCGGCTACCGGCTGTTCGAGAAGATTAGC
416	NLpep99(w/o Met)	A.A.	VTGYRLFEEKIS
417	NLpep100(w/o Met )	N.A.	GTGACCGGCTACCGGCTGTTCGAGCAGATTAGC
418	NLpep100(w/o Met )	A.A.	VTGYRLFEEKIS
419	NLpep101 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTCGAGAAGGAGAGC

10

20

30

420	NLpep101 (w/o Me t)	A.A.	VTGYRLFEEKES
421	NLpep102 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTTCGAGCAGGAGAGC
422	NLpep102 (w/o Me t)	A.A.	VTGYRLFEEQES
423	NLpep103 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTTCGAGCAGGAGCTG
424	NLpep103 (w/o Me t)	A.A.	VTGYRLFEEQEL
425	NLpep104 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTTCGAGAAGATTAGC
426	NLpep104 (w/o Me t)	A.A.	VEGYRLFEEKIS
427	NLpep105 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTTCGAGCAGATTAGC
428	NLpep105 (w/o Me t)	A.A.	VEGYRLFEEQIS
429	NLpep106 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTTCGAGAAGGAGAGC
430	NLpep106 (w/o Me t)	A.A.	VEGYRLFEEKES
431	NLpep107 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTTCGAGCAGGAGAGC
432	NLpep107 (w/o Me t)	A.A.	VEGYRLFEEQES
433	NLpep108 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTTCGAGCAGGAGCTG
434	NLpep108 (w/o Me t)	A.A.	VEGYRLFEEQEL
435	NLpep109 (w/o Me t)	N.A.	ATTAGCGGCTGGCGGCTGATGAAGAACATTAGC
436	NLpep109 (w/o Me t)	A.A.	ISGWRLMKNIS
437	NLpep110 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTCAAGAAGATTAGC
438	NLpep110 (w/o Me t)	A.A.	VEGYRLFKKIS

10

20

30

40

2162	NLpep111 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTTCGAGGAGATCAG C
2163	NLpep111 (w/ Met )	A.A.	MVTGYRLFEEIS
2164	NLpep112 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTTCGAGGAGGCCAG C
2165	NLpep112 (w/ Met )	A.A.	MVTGYRLFEEAS
2166	NLpep113 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTTCGAGGAGGAGAG C
2167	NLpep113 (w/ Met )	A.A.	MVTGYRLFEEES
2168	NLpep114 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTTCGAGGAGATCCT G
2169	NLpep114 (w/ Met )	A.A.	MVTGYRLFEEIL
2170	NLpep115 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTTCGAGGAGGCCCT G
2171	NLpep115 (w/ Met )	A.A.	MVTGYRLFEEAL
2172	NLpep116 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTTCGAGGAGGAGCT G
2173	NLpep116 (w/ Met )	A.A.	MVTGYRLFEEEL
2174	NLpep117 (w/ Met )	N.A.	ATGGTGGAGGGCTACCGGCTGTTTCGAGGAGATCAG C
2175	NLpep117 (w/ Met )	A.A.	MVEGYRLFEEIS
2176	NLpep118 (w/ Met )	N.A.	ATGGTGGAGGGCTACCGGCTGTTTCGAGGAGGCCAG C
2177	NLpep118 (w/ Met )	A.A.	MVEGYRLFEEAS
2178	NLpep119 (w/ Met )	N.A.	ATGGTGGAGGGCTACCGGCTGTTTCGAGGAGGAGAG C
2179	NLpep119 (w/ Met )	A.A.	MVEGYRLFEEES

10

20

30

40

2180	NLpep120 (w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAGGAGATCCT G
2181	NLpep120 (w/ Met)	A.A.	MVEGYRLFEEIL
2182	NLpep121 (w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAGGAGGCCCT G
2183	NLpep121 (w/ Met)	A.A.	MVEGYRLFEEAL
2184	NLpep122 (w/ Met)	N.A.	ATGGTGGAGGGCTACCGGCTGTTCGAGGAGGAGCT G
2185	NLpep122 (w/ Met)	A.A.	MVEGYRLFEEEL
2186	NLpep123 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGTTCGAAGAAGATCCT G
2187	NLpep123 (w/ Met)	A.A.	MVTGYRLFKKIL
2188	NLpep124 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGATGAAGAAGATCCT G
2189	NLpep124 (w/ Met)	A.A.	MVTGYRLMKKIL
2190	NLpep125 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGCACAAGAAGATCCT G
2191	NLpep125 (w/ Met)	A.A.	MVTGYRLHKKIL
2192	NLpep126 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGCTGAAGAAGATCCT G
2193	NLpep126 (w/ Met)	A.A.	MVTGYRLKKIL
2194	NLpep127 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGAGCAAGAAGATCCT G
2195	NLpep127 (w/ Met)	A.A.	MVTGYRLSKKIL
2196	NLpep128 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGTTCGAGAAGATCCT G
2197	NLpep128 (w/ Met)	A.A.	MVTGYRLFEEKIL
2198	NLpep129(w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGATGGAGAAGATCCT G
2199	NLpep129(w/ Met)	A.A.	MVTGYRLMEKIL

10

20

30

40

2200	NLpep130 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGCACGAGAAGATCCTG
2201	NLpep130 (w/ Met)	A.A.	MVTGYRLHEKIL
2202	NLpep131 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGCTGGAGAAGATCCTG
2203	NLpep131 (w/ Met)	A.A.	MVTGYRLLEKIL
2204	NLpep132 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGAGCGAGAAGATCCTG
2205	NLpep132 (w/ Met)	A.A.	MVTGYRLSEKIL
2206	NLpep133 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGATGGAGGAGATCCTG
2207	NLpep133 (w/ Met)	A.A.	MVTGYRLMEEIL
2208	NLpep134(w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGCACGAGGAGATCCTG
2209	NLpep134(w/ Met)	A.A.	MVTGYRLHEEIL
2210	NLpep135 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGCTGGAGGAGATCCTG
2211	NLpep135 (w/ Met)	A.A.	MVTGYRLLEEIL
2212	NLpep136 (w/ Met)	N.A.	ATGGTGACCGGCTACCGGCTGAGCGAGGAGATCCTG
2213	NLpep136 (w/ Met)	A.A.	MVTGYRLSEEIL
2214	NLpep137(w/ Met)	N.A.	ATGGTGAGCGGCTACCGGCTGTTCGAGGAGATCCTG
2215	NLpep137(w/ Met)	A.A.	MVSGYRLFEEIL
2216	NLpep138(w/ Met)	N.A.	ATGGTGACCGGCTGGCGGCTGTTCGAGGAGATCCTG
2217	NLpep138(w/ Met)	A.A.	MVTGWRLFEEIL
2218	NLpep139 (w/ Met)	N.A.	ATGGTGAGCGGCTGGCGGCTGTTCGAGGAGATCCTG
2219	NLpep139 (w/ Met)	A.A.	MVSGWRLFEEIL

10

20

30

40

2220	NLpep140 (w/ Met )	N.A.	ATGAACGTGACCGGCTACCGGCTGTTTCGAGGAGAT CCTG
2221	NLpep140 (w/ Met )	A.A.	MNVTGYRLFEEIL
2222	NLpep141 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTTCGAGGAGATCCT GAAC
2223	NLpep141 (w/ Met )	A.A.	MVTGYRLFEEILN
2224	NLpep142 (w/ Met )	N.A.	ATGAACGTGACCGGCTACCGGCTGTTTCGAGGAGAT CCTGAAC
2225	NLpep142 (w/ Met )	A.A.	MNVTGYRLFEEILN
2226	NLpep143 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTTCGAGGAGATC
2227	NLpep143 (w/ Met )	A.A.	MVTGYRLFEEI
2228	NLpep144 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTCCAGAAGATCAG C
2229	NLpep144 (w/ Met )	A.A.	MVTGYRLFQKIS
2230	NLpep145 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTCAAGAAGATCAG CAAC
2231	NLpep145 (w/ Met )	A.A.	MVTGYRLFKKISN
2232	NLpep146 (w/ Met )	N.A.	ATGGTGACCGGCTACCGGCTGTTCAAGAAGATCAG C
2233	NLpep146 (w/ Met )	A.A.	MVTGYRLFKKIS
2234	NLpep147 (w/ Met )	A.A.	MVSGWRLFKKISA
2235	NLpep148 (w/ Met )	A.A.	MGVSGWRLFKKIS
2236	NLpep149 (w/ Met )	A.A.	MSVSGWRLFKKISN
2237	NLpep150 (w/ Met )	A.A.	MSVSGWRLFKKISA
2238	NLpep151 (w/ Met )	A.A.	MNSVSGWRLFKKISA
2239	NLpep152 (w/ Met )	A.A.	MNSVSGWRLFKKISN

10

20

30

40

2240	NLpep153 (w/ Met )	A.A.	MSNVSGWRLFKKIS
2241	NLpep154 (w/ Met )	A.A.	MSGVSGWRLFKKIS
2242	NLpep155 (w/ Met )	A.A.	MNSNVSGWRLFKKIS
2243	NLpep156 (w/ Met )	A.A.	MNSGVSGWRLFKKIS
2244	NLpep157 (w/ Met )	A.A.	MSVSGWRLFKKIS
2245	NLpep158 (w/ Met )	A.A.	MNSVSGWRLFKKIS
2246	NLpep159 (w/ Met )	A.A.	MSNVSGWRLFKKISN
2247	NLpep160 (w/ Met )	A.A.	MNSNVSGWRLFKKISN
2248	NLpep161 (w/ Met )	A.A.	MGWRLFKK
2249	NLpep162(w/ Met)	A.A.	MGWALFKK
2250	NLpep163 (w/ Met )	A.A.	MVTGWALFEEIL
2251	NLpep164 (w/ Met )	A.A.	MVTGYALFQEIL
2252	NLpep165 (w/ Met )	A.A.	MVTGYALFEQIL
2253	NLpep166 (w/ Met )	A.A.	MVTGYALFEEIL
2254	NLpep167 (w/ Met )	N.A.	ATGGTGTCCGGCTGGGCACTGTTCAAGAAAATTC C
2255	NLpep167 (w/ Met )	A.A.	MVSGWALFKKIS
2256	NLpep168 (w/ Met )	A.A.	MVSGWKLFKKIS
2257	NLpep169 (w/ Met )	N.A.	ATGGTGTCCGGCTGGCAGCTGTTCAAGAAAATTC C
2258	NLpep169 (w/ Met )	A.A.	MVSGWQLFKKIS
2259	NLpep170 (w/ Met )	A.A.	MVSGWELFKKIS

10

20

30

40

2260	NLpep171 (w/ Met)	N.A.	ATGGTGTCCGGCTGGCTGCTGTTCAAGAAAATTC
2261	NLpep171 (w/ Met)	A.A.	MVSGWLLFKKIS
2262	NLpep172(w/ Met)	N.A.	ATGGTGTCCGGCTGGGTGCTGTTCAAGAAAATTC
2263	NLpep172(w/ Met)	A.A.	MVSGWVLFKKIS
2264	NLpep111 (w/o Met)	N.A.	GTGACCGGCTACCGGCTGTTCGAGGAGATCAGC
2265	NLpep111 (w/o Met)	A.A.	VTGYRLFEEIS
2266	NLpep112 (w/o Met)	N.A.	GTGACCGGCTACCGGCTGTTCGAGGAGGCCAGC
2267	NLpep112 (w/o Met)	A.A.	VTGYRLFEEAS
2268	NLpep113 (w/o Met)	N.A.	GTGACCGGCTACCGGCTGTTCGAGGAGGAGAGC
2269	NLpep113 (w/o Met)	A.A.	VTGYRLFEEES
2270	NLpep114 (w/o Met)	N.A.	GTGACCGGCTACCGGCTGTTCGAGGAGATCCTG
2271	NLpep114 (w/o Met)	A.A.	VTGYRLFEEIL
2272	NLpep115 (w/o Met)	N.A.	GTGACCGGCTACCGGCTGTTCGAGGAGGCCCTG
2273	NLpep115 (w/o Met)	A.A.	VTGYRLFEEAL
2274	NLpep116 (w/o Met)	N.A.	GTGACCGGCTACCGGCTGTTCGAGGAGGAGCTG
2275	NLpep116 (w/o Met)	A.A.	VTGYRLFEEEL
2276	NLpep117 (w/o Met)	N.A.	GTGGAGGGCTACCGGCTGTTCGAGGAGATCAGC
2277	NLpep117 (w/o Met)	A.A.	VEGYRLFEEIS
2278	NLpep118 (w/o Met)	N.A.	GTGGAGGGCTACCGGCTGTTCGAGGAGGCCAGC
2279	NLpep118 (w/o Met)	A.A.	VEGYRLFEEAS

10

20

30

40

2280	NLpep119 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTCGAGGAGGAGAGC
2281	NLpep119 (w/o Me t)	A.A.	VEGYRLFEEES
2282	NLpep120 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTCGAGGAGATCCTG
2283	NLpep120 (w/o Me t)	A.A.	VEGYRLFEEIL
2284	NLpep121 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTCGAGGAGGCCCTG
2285	NLpep121 (w/o Me t)	A.A.	VEGYRLFEEAL
2286	NLpep122 (w/o Me t)	N.A.	GTGGAGGGCTACCGGCTGTTCGAGGAGGAGCTG
2287	NLpep122 (w/o Me t)	A.A.	VEGYRLFEEEL
2288	NLpep123 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTCAAGAAGATCCTG
2289	NLpep123 (w/o Me t)	A.A.	VTGYRLFKKIL
2290	NLpep124 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGATGAAGAAGATCCTG
2291	NLpep124 (w/o Me t)	A.A.	VTGYRLMKKIL
2292	NLpep125 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGCACAAGAAGATCCTG
2293	NLpep125 (w/o Me t)	A.A.	VTGYRLHKKIL
2294	NLpep126 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGCTGAAGAAGATCCTG
2295	NLpep126 (w/o Me t)	A.A.	VTGYRLKKIL
2296	NLpep127 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGAGCAAGAAGATCCTG
2297	NLpep127 (w/o Me t)	A.A.	VTGYRLSKKIL
2298	NLpep128 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTCGAGAAGATCCTG
2299	NLpep128 (w/o Me t)	A.A.	VTGYRLFEEKIL

10

20

30

40

2300	NLpep129(w/o Met )	N.A.	GTGACCGGCTACCGGCTGATGGAGAAGATCCTG
2301	NLpep129(w/o Met )	A.A.	VTGYRLMEKIL
2302	NLpep130 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGCACGAGAAGATCCTG
2303	NLpep130 (w/o Me t)	A.A.	VTGYRLHEKIL
2304	NLpep131 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGCTGGAGAAGATCCTG
2305	NLpep131 (w/o Me t)	A.A.	VTGYRLLEKIL
2306	NLpep132 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGAGCGAGAAGATCCTG
2307	NLpep132 (w/o Me t)	A.A.	VTGYRLSEKIL
2308	NLpep133 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGATGGAGGAGATCCTG
2309	NLpep133 (w/o Me t)	A.A.	VTGYRLMEEIL
2310	NLpep134(w/o Met )	N.A.	GTGACCGGCTACCGGCTGCACGAGGAGATCCTG
2311	NLpep134(w/o Met )	A.A.	VTGYRLHEEIL
2312	NLpep135 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGCTGGAGGAGATCCTG
2313	NLpep135 (w/o Me t)	A.A.	VTGYRLLEEIL
2314	NLpep136 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGAGCGAGGAGATCCTG
2315	NLpep136 (w/o Me t)	A.A.	VTGYRLSEEIL
2316	NLpep137(w/o Met )	N.A.	GTGAGCGGCTACCGGCTGTTCGAGGAGATCCTG
2317	NLpep137(w/o Met )	A.A.	VSGYRLFEEIL
2318	NLpep138(w/o Met )	N.A.	GTGACCGGCTGGCGGCTGTTCGAGGAGATCCTG
2319	NLpep138(w/o Met )	A.A.	VTGWRLFEEIL

10

20

30

40

2320	NLpep139 (w/o Me t)	N.A.	GTGAGCGGCTGGCGGCTGTTTCGAGGAGATCCTG
2321	NLpep139 (w/o Me t)	A.A.	VSGWRLFEEIL
2322	NLpep140 (w/o Me t)	N.A.	AACGTGACCGGCTACCGGCTGTTTCGAGGAGATCCT G
2323	NLpep140 (w/o Me t)	A.A.	NVTGYRLFEEIL
2324	NLpep141 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTTCGAGGAGATCCTGAA C
2325	NLpep141 (w/o Me t)	A.A.	VTGYRLFEEILN
2326	NLpep142 (w/o Me t)	N.A.	AACGTGACCGGCTACCGGCTGTTTCGAGGAGATCCT GAAC
2327	NLpep142 (w/o Me t)	A.A.	NVTGYRLFEEILN
2328	NLpep143 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTTCGAGGAGATC
2329	NLpep143 (w/o Me t)	A.A.	VTGYRLFEEI
2330	NLpep144 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTCCAGAAGATCAGC
2331	NLpep144 (w/o Me t)	A.A.	VTGYRLFQKIS
2332	NLpep145 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTCAAGAAGATCAGCAA C
2333	NLpep145 (w/o Me t)	A.A.	VTGYRLFKKISN
2334	NLpep146 (w/o Me t)	N.A.	GTGACCGGCTACCGGCTGTTCAAGAAGATCAGC
2335	NLpep146 (w/o Me t)	A.A.	VTGYRLFKKIS
2336	NLpep147 (w/o Me t)	A.A.	VSGWRLFKKISA
2337	NLpep148 (w/o Me t)	A.A.	GVSGWRLFKKIS
2338	NLpep149 (w/o Me t)	A.A.	SVSGWRLFKKISN
2339	NLpep150 (w/o Me t)	A.A.	SVSGWRLFKKISA

10

20

30

40

2340	NLpep151 (w/o Me t)	AA	NSVSGWRLFKKISA
2341	NLpep152 (w/o Me t)	AA	NSVSGWRLFKKISN
2342	NLpep153 (w/o Me t)	AA	SNVSGWRLFKKIS
2343	NLpep154 (w/o Me t)	AA	SGVSGWRLFKKIS
2344	NLpep155 (w/o Me t)	AA	NSNVSGWRLFKKIS
2345	NLpep156 (w/o Me t)	AA	NSGVSGWRLFKKIS
2346	NLpep157 (w/o Me t)	AA	SVSGWRLFKKIS
2347	NLpep158 (w/o Me t)	AA	NSVSGWRLFKKIS
2348	NLpep159 (w/o Me t)	AA	SNVSGWRLFKKISN
2349	NLpep160 (w/o Me t)	AA	NSNVSGWRLFKKISN
2350	NLpep161 (w/o Me t)	AA	GWRLFKK
2351	NLpep162(w/o Met )	AA	GWALFKK
2352	NLpep163 (w/o Me t)	AA	VTGWALFEEIL
2353	NLpep164 (w/o Me t)	AA	VTGYALFQEIL
2354	NLpep165 (w/o Me t)	AA	VTGYALFEQIL
2355	NLpep166 (w/o Me t)	AA	VTGYALFEEIL
2356	NLpep167 (w/o Me t)	NA	GTGTCCGGCTGGGCACTGTTCAAGAAAATTTCC
2357	NLpep167 (w/o Me t)	AA	VSGWALFKKIS
2358	NLpep168 (w/o Me t)	AA	VSGWKLFKKIS
2359	NLpep169 (w/o Me t)	NA	GTGTCCGGCTGGCAGCTGTTCAAGAAAATTTCC

10

20

30

40

2360	NLpep169 (w/o Met)	A.A.	VSGWQLFKKIS
2361	NLpep170 (w/o Met)	A.A.	VSGWELFKKIS
2362	NLpep171 (w/o Met)	N.A.	GTGTCCGGCTGGCTGCTGTTCAAGAAAATTTCC
2363	NLpep171 (w/o Met)	A.A.	VSGWLLFKKIS
2364	NLpep172(w/o Met)	N.A.	GTGTCCGGCTGGGTGCTGTTCAAGAAAATTTCC
2365	NLpep172(w/o Met)	A.A.	VSGWVLFKKIS

10

## 【 0 0 8 4 】

特定の実施形態では、表 1 のペプチドが提供される（たとえば、内部タグまたは内部タグの構造相補体として）。一部の実施形態では、内部タグまたは構造相補体は G V T G W R L C K R I L A（配列番号 2）及び／または表 1 でリストにしたペプチドのいずれかとの単一アミノ酸の差異を含む。一部の実施形態では、内部タグまたは構造相補体は G V T G W R L C K R I L A（配列番号 2）及び／または表 1 でリストにしたペプチドのいずれかとの 2 以上（たとえば、2、3、4、5、6、7、8、8、9、10 等）のアミノ酸の差異を含む。一部の実施形態では、配列番号 3～438 及び 2162～2365 のアミノ酸配列の 1 つを含む内部タグまたは構造相補体が提供される。一部の実施形態では、1 以上の付加、置換及び／または欠失を伴う配列番号 3～438 及び 2162～2365 のアミノ酸配列の 1 つを含む内部タグまたは構造相補体が提供される。一部の実施形態では、内部タグ、構造相補体、またはその一部は配列番号 3～438 及び 2162～2365 のアミノ酸配列の 1 以上との 70% を超える（たとえば、71%、75%、80%、85%、90%、95%、99%）配列同一性を含む。一部の実施形態では、配列番号 3～438 及び 2162～2365 の配列をコードする核酸の 1 つを含む核酸が提供される。一部の実施形態では、1 以上の付加、置換及び／または欠失を伴う配列番号 3～438 及び 2162～2365 の核酸配列の 1 つを含む核酸が提供される。一部の実施形態では、核酸またはその一部は配列番号 3～438 及び 2162～2365 の核酸配列の 1 以上との 70% を超える（たとえば、71%、75%、80%、85%、90%、95%、99%）配列同一性を含む。一部の実施形態では、配列番号 3～438 及び 2162～2365 のアミノ酸配列の 1 つをコードする核酸が提供される。一部の実施形態では、1 以上の付加、置換及び／または欠失を伴う配列番号 3～438 及び 2162～2365 のアミノ酸配列の 1 つをコードする核酸が提供される。一部の実施形態では、配列番号 3～438 及び 2162～2365 のアミノ酸配列の 1 以上との 70% を超える（たとえば、71%、75%、80%、85%、90%、95%、99%）配列同一性を持つアミノ酸をコードする核酸が提供される。

20

30

40

## 【 0 0 8 5 】

特定の実施形態では、表 1 の核酸が提供される。一部の実施形態では、表 1 のペプチドをコードする核酸が提供される。一部の実施形態では、別のポリペプチド配列に挿入された表 1 のペプチドをコードする核酸が提供される。一部の実施形態では、本発明の核酸は、M G V T G W R L C K R I L A（配列番号 2）及び／または表 1 でリストにしたペプチド（たとえば、ポリペプチド配列に挿入された）のいずれかとの単一アミノ酸の差異を含むペプチドをコードする。一部の実施形態では、核酸は、M G V T G W R L C K R I L A（配列番号 2）及び／または表 1 でリストにしたペプチド（たとえば、ポリペプチド配列に挿入された）のいずれかとの 2 以上（たとえば、2、3、4、5、6、7、8、8、9、10 等）のアミノ酸の差異を含むペプチドをコードする。一部の実施形態では、表 1 に

50

おける核酸の１つの配列（たとえば、ポリペプチド配列に挿入される）を含む核酸が提供される。一部の実施形態では、１以上の付加、置換及び／または欠失を伴う表１の核酸の１つを含む核酸が提供される。一部の実施形態では、核酸またはその一部は、表１の核酸の１以上（たとえば、ポリペプチド配列に挿入される）との７０％を超える（たとえば、７１％、７５％、８０％、８５％、９０％、９５％、９９％）配列同一性を含む。

#### 【００８６】

一部の実施形態では、本明細書に記載されている実施形態で使用する内部タグ及び／または内部タグの構造相補体には、配列番号４４０に由来する１以上のアミノ酸の置換、欠失または付加を伴うポリペプチドが含まれる。一部の実施形態では、本明細書で提供されるのは、表２のアミノ酸配列、及び／または表２の核酸配列を含む核酸を含む内部タグまたは内部タグの構造相補体である。

#### 【００８７】

表２．例となる内部タグ及び／または構造相補体のポリペプチド配列

配列 番号	ポリマ ー	ID	配列 番号	ポリマ ー	ID	配列 番号	ポリマ ー	ID
441	N.A.	R11N	727	N.A.	5A2+V58P	1013	N.A.	5P D6 (-152-157)
442	A.A	R11N	728	A.A	5A2+V58P	1014	A.A	5P D6 (-152-157)
443	N.A.	T13I	729	N.A.	5A2+V58Q	1015	N.A.	5P D7 (-151-157)
444	A.A	T13I	730	A.A	5A2+V58Q	1016	A.A	5P D7 (-151-157)
445	N.A.	G15S	731	N.A.	5A2+V58R	1017	N.A.	5P +F31A
446	A.A	G15S	732	A.A	5A2+V58R	1018	A.A	5P +F31A
447	N.A.	L18Q	733	N.A.	5A2+V58S	1019	N.A.	5P+F31C
448	A.A	L18Q	734	A.A	5A2+V58S	1020	A.A	5P+F31C
449	N.A.	Q20K	735	N.A.	5A2+V58T	1021	N.A.	5P+F31D
450	A.A	Q20K	736	A.A	5A2+V58T	1022	A.A	5P+F31D
451	N.A.	V27M	737	N.A.	5A2+V58W	1023	N.A.	5P+F31E
452	A.A	V27M	738	A.A	5A2+V58W	1024	A.A	5P+F31E
453	N.A.	F31I	739	N.A.	5A2+V58Y	1025	N.A.	5P+F31G
454	A.A	F31I	740	A.A	5A2+V58Y	1026	A.A	5P+F31G

455	N.A.	F31L	741	N.A.	5A2+A67C	1027	N.A.	5P+F31H
456	A.A	F31L	742	A.A	5A2+A67C	1028	A.A	5P+F31H
457	N.A.	F31V	743	N.A.	5A2+A67D	1029	N.A.	5P+F31I
458	A.A	F31V	744	A.A	5A2+A67D	1030	A.A	5P+F31I
459	N.A.	Q32R	745	N.A.	5A2+A67E	1031	N.A.	5P+F31K
460	A.A	Q32R	746	A.A	5A2+A67E	1032	A.A	5P+F31K
461	N.A.	N33K	747	N.A.	5A2+A67F	1033	N.A.	5P+F31L
462	A.A	N33K	748	A.A	5A2+A67F	1034	A.A	5P+F31L
463	N.A.	N33R	749	N.A.	5A2+A67G	1035	N.A.	5P+F31M
464	A.A	N33R	750	A.A	5A2+A67G	1036	A.A	5P+F31M
465	N.A.	I56N	751	N.A.	5A2+A67H	1037	N.A.	5P+F31N
466	A.A	I56N	752	A.A	5A2+A67H	1038	A.A	5P+F31N
467	N.A.	V58A	753	N.A.	5A2+A67I	1039	N.A.	5P+F31P
468	A.A	V58A	754	A.A	5A2+A67I	1040	A.A	5P+F31P
469	N.A.	I59T	755	N.A.	5A2+A67K	1041	N.A.	5P+F31Q
470	A.A	I59T	756	A.A	5A2+A67K	1042	A.A	5P+F31Q
471	N.A.	G67S	757	N.A.	5A2+A67L	1043	N.A.	5P+F31R
472	A.A	G67S	758	A.A	5A2+A67L	1044	A.A	5P+F31R
473	N.A.	G67D	759	N.A.	5A2+A67M	1045	N.A.	5P+F31S
474	A.A	G67D	760	A.A	5A2+A67M	1046	A.A	5P+F31S
475	N.A.	K75E	761	N.A.	5A2+A67N	1047	N.A.	5P+F31T

10

20

30

40

476	A.A	K75E	762	A.A	5A2+A67N	1048	A.A	5P+F31T
477	N.A.	M106V	763	N.A.	5A2+A67P	1049	N.A.	5P+F31V
478	A.A	M106V	764	A.A	5A2+A67P	1050	A.A	5P+F31V
479	N.A.	M106I	765	N.A.	5A2+A67Q	1051	N.A.	5P+F31W
480	A.A	M106I	766	A.A	5A2+A67Q	1052	A.A	5P+F31W
481	N.A.	D108N	767	N.A.	5A2+A67R	1053	N.A.	5P+F31Y
482	A.A	D108N	768	A.A	5A2+A67R	1054	A.A	5P+F31Y
483	N.A.	R112Q	769	N.A.	5A2+A67S	1055	N.A.	5P +L46A
484	A.A	R112Q	770	A.A	5A2+A67S	1056	A.A	5P +L46A
485	N.A.	N144T	771	N.A.	5A2+A67T	1057	N.A.	5P+L46C
486	A.A	N144T	772	A.A	5A2+A67T	1058	A.A	5P+L46C
487	N.A.	L149M	773	N.A.	5A2+A67V	1059	N.A.	5P+L46D
488	A.A	L149M	774	A.A	5A2+A67V	1060	A.A	5P+L46D
489	N.A.	N156D	775	N.A.	5A2+A67W	1061	N.A.	5P+L46E
490	A.A	N156D	776	A.A	5A2+A67W	1062	A.A	5P+L46E
491	N.A.	N156S	777	N.A.	5A2+A67Y	1063	N.A.	5P+L46F
492	A.A	N156S	778	A.A	5A2+A67Y	1064	A.A	5P+L46F
493	N.A.	V157D	779	N.A.	5A2+M106A	1065	N.A.	5P+L46G
494	A.A	V157D	780	A.A	5A2+M106A	1066	A.A	5P+L46G
495	N.A.	V157S	781	N.A.	5A2+M106C	1067	N.A.	5P+L46H

10

20

30

40

496	A.A	V157S	782	A.A	5A2+M106C	1068	A.A	5P+L46H
497	N.A.	G8A	783	N.A.	5A2+M106D	1069	N.A.	5P+L46I
498	A.A	G8A	784	A.A	5A2+M106D	1070	A.A	5P+L46I
499	N.A.	G15A	785	N.A.	5A2+M106E	1071	N.A.	5P+L46K
500	A.A	G15A	786	A.A	5A2+M106E	1072	A.A	5P+L46K
501	N.A.	G25A	787	N.A.	5A2+M106F	1073	N.A.	5P+L46M
502	A.A	G25A	788	A.A	5A2+M106F	1074	A.A	5P+L46M
503	N.A.	G26A	789	N.A.	5A2+M106G	1075	N.A.	5P+L46N
504	A.A	G26A	790	A.A	5A2+M106G	1076	A.A	5P+L46N
505	N.A.	G35A	791	N.A.	5A2+M106H	1077	N.A.	5P+L46P
506	A.A	G35A	792	A.A	5A2+M106H	1078	A.A	5P+L46P
507	N.A.	G48A	793	N.A.	5A2+M106I	1079	N.A.	5P+L46Q
508	A.A	G48A	794	A.A	5A2+M106I	1080	A.A	5P+L46Q
509	N.A.	G51A	795	N.A.	5A2+M106K	1081	N.A.	5P+L46R
510	A.A	G51A	796	A.A	5A2+M106K	1082	A.A	5P+L46R
511	N.A.	G64A	797	N.A.	5A2+M106L	1083	N.A.	5P+L46S
512	A.A	G64A	798	A.A	5A2+M106L	1084	A.A	5P+L46S
513	N.A.	G67A	799	N.A.	5A2+M106N	1085	N.A.	5P+L46T
514	A.A	G67A	800	A.A	5A2+M106N	1086	A.A	5P+L46T
515	N.A.	G71A	801	N.A.	5A2+M106P	1087	N.A.	5P+L46V

10

20

30

40

516	A.A	G71A	802	A.A	5A2+M106P	1088	A.A	5P+L46V
517	N.A.	G95A	803	N.A.	5A2+M106Q	1089	N.A.	5P+L46W
518	A.A	G95A	804	A.A	5A2+M106Q	1090	A.A	5P+L46W
519	N.A.	G101A	805	N.A.	5A2+M106R	1091	N.A.	5P+L46Y
520	A.A	G101A	806	A.A	5A2+M106R	1092	A.A	5P+L46Y
521	N.A.	G111A	807	N.A.	5A2+M106S	1093	N.A.	5P +N108A
522	A.A	G111A	808	A.A	5A2+M106S	1094	A.A	5P +N108A
523	N.A.	G116A	809	N.A.	5A2+M106T	1095	N.A.	5P+N108C
524	A.A	G116A	810	A.A	5A2+M106T	1096	A.A	5P+N108C
525	N.A.	G122A	811	N.A.	5A2+M106V	1097	N.A.	5P+N108D
526	A.A	G122A	812	A.A	5A2+M106V	1098	A.A	5P+N108D
527	N.A.	G129A	813	N.A.	5A2+M106W	1099	N.A.	5P+N108E
528	A.A	G129A	814	A.A	5A2+M106W	1100	A.A	5P+N108E
529	N.A.	G134A	815	N.A.	5A2+M106Y	1101	N.A.	5P+N108F
530	A.A	G134A	816	A.A	5A2+M106Y	1102	A.A	5P+N108F
531	N.A.	G147A	817	N.A.	5A2+L149A	1103	N.A.	5P+N108G
532	A.A	G147A	818	A.A	5A2+L149A	1104	A.A	5P+N108G
533	N.A.	I54A	819	N.A.	5A2+L149C	1105	N.A.	5P+N108H
534	A.A	I54A	820	A.A	5A2+L149C	1106	A.A	5P+N108H
535	N.A.	5A1 (G15A/ D19A/G35A /G51A/G67 A)	821	N.A.	5A2+L149D	1107	N.A.	5P+N108I

10

20

30

40

536	A.A	5A1 (G15A/ D19A/G35A/ /G51A/G67 A)	822	A.A	5A2+L149D	1108	A.A	5P+N108I
537	N.A.	4A1 (G15A/ G35A/G67A/ /G71A)	823	N.A.	5A2+L149E	1109	N.A.	5P+N108K
538	A.A	4A1 (G15A/ G35A/G67A/ /G71A)	824	A.A	5A2+L149E	1110	A.A	5P+N108K
539	N.A.	5A2 (G15A/ G35A/G51A/ /G67A/G71 A)	825	N.A.	5A2+L149F	1111	N.A.	5P+N108L
540	A.A	5A2 (G15A/ G35A/G51A/ /G67A/G71 A)	826	A.A	5A2+L149F	1112	A.A	5P+N108L
541	N.A.	5A2+A15G	827	N.A.	5A2+L149G	1113	N.A.	5P+N108M
542	A.A	5A2+A15G	828	A.A	5A2+L149G	1114	A.A	5P+N108M
543	N.A.	5A2+A35G	829	N.A.	5A2+L149H	1115	N.A.	5P+N108P
544	A.A	5A2+A35G	830	A.A	5A2+L149H	1116	A.A	5P+N108P
545	N.A.	5A2+A51G	831	N.A.	5A2+L149I	1117	N.A.	5P+N108Q
546	A.A	5A2+A51G	832	A.A	5A2+L149I	1118	A.A	5P+N108Q
547	N.A.	5A2+A67G	833	N.A.	5A2+L149K	1119	N.A.	5P+N108R
548	A.A	5A2+A67G	834	A.A	5A2+L149K	1120	A.A	5P+N108R
549	N.A.	5A2+A71G	835	N.A.	5A2+L149M	1121	N.A.	5P+N108S
550	A.A	5A2+A71G	836	A.A	5A2+L149M	1122	A.A	5P+N108S

10

20

30

40

551	N.A.	5A2+R11A	837	N.A.	5A2+L149N	1123	N.A.	5P+N108T
552	A.A	5A2+R11A	838	A.A	5A2+L149N	1124	A.A	5P+N108T
553	N.A.	5A2+R11C	839	N.A.	5A2+L149P	1125	N.A.	5P+N108V
554	A.A	5A2+R11C	840	A.A	5A2+L149P	1126	A.A	5P+N108V
555	N.A.	5A2+R11D	841	N.A.	5A2+L149Q	1127	N.A.	5P+N108W
556	A.A	5A2+R11D	842	A.A	5A2+L149Q	1128	A.A	5P+N108W
557	N.A.	5A2+R11E	843	N.A.	5A2+L149R	1129	N.A.	5P+N108Y
558	A.A	5A2+R11E	844	A.A	5A2+L149R	1130	A.A	5P+N108Y
559	N.A.	5A2+R11F	845	N.A.	5A2+L149S	1131	N.A.	5P +T144A
560	A.A	5A2+R11F	846	A.A	5A2+L149S	1132	A.A	5P +T144A
561	N.A.	5A2+R11G	847	N.A.	5A2+L149T	1133	N.A.	5P+T144C
562	A.A	5A2+R11G	848	A.A	5A2+L149T	1134	A.A	5P+T144C
563	N.A.	5A2+R11H	849	N.A.	5A2+L149V	1135	N.A.	5P+T144D
564	A.A	5A2+R11H	850	A.A	5A2+L149V	1136	A.A	5P+T144D
565	N.A.	5A2+R11I	851	N.A.	5A2+L149W	1137	N.A.	5P+T144E
566	A.A	5A2+R11I	852	A.A	5A2+L149W	1138	A.A	5P+T144E
567	N.A.	5A2+R11K	853	N.A.	5A2+L149Y	1139	N.A.	5P+T144F
568	A.A	5A2+R11K	854	A.A	5A2+L149Y	1140	A.A	5P+T144F
569	N.A.	5A2+R11L	855	N.A.	5A2+V157A	1141	N.A.	5P+T144G
570	A.A	5A2+R11L	856	A.A	5A2+V157A	1142	A.A	5P+T144G

10

20

30

40

571	N.A.	5A2+R11M	857	N.A.	5A2+V157C	1143	N.A.	5P+T144H
572	A.A	5A2+R11M	858	A.A	5A2+V157C	1144	A.A	5P+T144H
573	N.A.	5A2+R11N	859	N.A.	5A2+V157D	1145	N.A.	5P+T144I
574	A.A	5A2+R11N	860	A.A	5A2+V157D	1146	A.A	5P+T144I
575	N.A.	5A2+R11P	861	N.A.	5A2+V157E	1147	N.A.	5P+T144K
576	A.A	5A2+R11P	862	A.A	5A2+V157E	1148	A.A	5P+T144K
577	N.A.	5A2+R11Q	863	N.A.	5A2+V157F	1149	N.A.	5P+T144L
578	A.A	5A2+R11Q	864	A.A	5A2+V157F	1150	A.A	5P+T144L
579	N.A.	5A2+R11S	865	N.A.	5A2+V157G	1151	N.A.	5P+T144M
580	A.A	5A2+R11S	866	A.A	5A2+V157G	1152	A.A	5P+T144M
581	N.A.	5A2+R11T	867	N.A.	5A2+V157H	1153	N.A.	5P+T144N
582	A.A	5A2+R11T	868	A.A	5A2+V157H	1154	A.A	5P+T144N
583	N.A.	5A2+R11V	869	N.A.	5A2+V157I	1155	N.A.	5P+T144P
584	A.A	5A2+R11V	870	A.A	5A2+V157I	1156	A.A	5P+T144P
585	N.A.	5A2+R11W	871	N.A.	5A2+V157K	1157	N.A.	5P+T144Q
586	A.A	5A2+R11W	872	A.A	5A2+V157K	1158	A.A	5P+T144Q
587	N.A.	5A2+R11Y	873	N.A.	5A2+V157L	1159	N.A.	5P+T144R
588	A.A	5A2+R11Y	874	A.A	5A2+V157L	1160	A.A	5P+T144R
589	N.A.	5A2+A15C	875	N.A.	5A2+V157M	1161	N.A.	5P+T144S
590	A.A	5A2+A15C	876	A.A	5A2+V157M	1440	A.A	5P+T144S

10

20

30

40

591	N.A.	5A2+A15D	877	N.A.	5A2+V157N	1163	N.A.	5P+T144V
592	A.A	5A2+A15D	878	A.A	5A2+V157N	1164	A.A	5P+T144V
593	N.A.	5A2+A15E	879	N.A.	5A2+V157P	1165	N.A.	5P+T144W
594	A.A	5A2+A15E	880	A.A	5A2+V157P	1166	A.A	5P+T144W
595	N.A.	5A2+A15F	881	N.A.	5A2+V157Q	1167	N.A.	5P+T144Y
596	A.A	5A2+A15F	882	A.A	5A2+V157Q	1168	A.A	5P+T144Y
597	N.A.	5A2+A15G	883	N.A.	5A2+V157R	1169	N.A.	5P +P157A
598	A.A	5A2+A15G	884	A.A	5A2+V157R	1170	A.A	5P +P157A
599	N.A.	5A2+A15H	885	N.A.	5A2+V157S	1171	N.A.	5P+P157C
600	A.A	5A2+A15H	886	A.A	5A2+V157S	1172	A.A	5P+P157C
601	N.A.	5A2+A15I	887	N.A.	5A2+V157T	1173	N.A.	5P+P157D
602	A.A	5A2+A15I	888	A.A	5A2+V157T	1174	A.A	5P+P157D
603	N.A.	5A2+A15K	889	N.A.	5A2+V157W	1175	N.A.	5P+P157E
604	A.A	5A2+A15K	890	A.A	5A2+V157W	1176	A.A	5P+P157E
605	N.A.	5A2+A15L	891	N.A.	5A2+V157Y	1177	N.A.	5P+P157F
606	A.A	5A2+A15L	892	A.A	5A2+V157Y	1178	A.A	5P+P157F
607	N.A.	5A2+A15M	893	N.A.	5A2+Q20K	1179	N.A.	5P+P157G
608	A.A	5A2+A15M	894	A.A	5A2+Q20K	1180	A.A	5P+P157G
609	N.A.	5A2+A15N	895	N.A.	5A2+V27M	1181	N.A.	5P+P157H
610	A.A	5A2+A15N	896	A.A	5A2+V27M	1182	A.A	5P+P157H

10

20

30

40

611	N.A.	5A2+A15P	897	N.A.	5A2+N33K	1183	N.A.	5P+P157I
612	A.A	5A2+A15P	898	A.A	5A2+N33K	1184	A.A	5P+P157I
613	N.A.	5A2+A15Q	899	N.A.	5A2+V38I	1185	N.A.	5P+P157K
614	A.A	5A2+A15Q	900	A.A	5A2+V38I	1186	A.A	5P+P157K
615	N.A.	5A2+A15R	901	N.A.	5A2+I56N	1187	N.A.	5P+P157L
616	A.A	5A2+A15R	902	A.A	5A2+I56N	1188	A.A	5P+P157L
617	N.A.	5A2+A15S	903	N.A.	5A2+D108N	1189	N.A.	5P+P157M
618	A.A	5A2+A15S	904	A.A	5A2+D108N	1190	A.A	5P+P157M
619	N.A.	5A2+A15T	905	N.A.	5A2+N144T	1191	N.A.	5P+P157N
620	A.A	5A2+A15T	906	A.A	5A2+N144T	1192	A.A	5P+P157N
621	N.A.	5A2+A15V	907	N.A.	5A2+V27M+A 35G	1193	N.A.	5P+P157Q
622	A.A	5A2+A15V	908	A.A	5A2+V27M+A 35G	1194	A.A	5P+P157Q
623	N.A.	5A2+A15W	909	N.A.	5A2+A71G+K 75E	1195	N.A.	5P+P157R
624	A.A	5A2+A15W	910	A.A	5A2+A71G+K 75E	1196	A.A	5P+P157R
625	N.A.	5A2+A15Y	911	N.A.	5A2+R11E+L1 49M	1197	N.A.	5P+P157S
626	A.A	5A2+A15Y	912	A.A	5A2+R11E+L1 49M	1198	A.A	5P+P157S
627	N.A.	5A2+L18A	913	N.A.	5A2+R11E+V 157P	1199	N.A.	5P+P157T
628	A.A	5A2+L18A	914	A.A	5A2+R11E+V 157P	1200	A.A	5P+P157T
629	N.A.	5A2+L18C	915	N.A.	5A2+D108N+ N144T	1201	N.A.	5P+P157V
630	A.A	5A2+L18C	916	A.A	5A2+D108N+ N144T	1202	A.A	5P+P157V

10

20

30

40

631	N.A.	5A2+L18D	917	N.A.	5A2+L149M+V157D	1203	N.A.	5P+P157W
632	A.A	5A2+L18D	918	A.A	5A2+L149M+V157D	1204	A.A	5P+P157W
633	N.A.	5A2+L18E	919	N.A.	5A2+L149M+V157P	1205	N.A.	5P+P157Y
634	A.A	5A2+L18E	920	A.A	5A2+L149M+V157P	1206	A.A	5P+P157Y
635	N.A.	5A2+L18F	921	N.A.	3P (5A2+R11E+L149M+V157P)	1207	N.A.	5P+I107L
636	A.A	5A2+L18F	922	A.A	3P (5A2+R11E+L149M+V157P)	1208	A.A	5P+I107L
637	N.A.	5A2+L18G	923	N.A.	3P+D108N	1209	N.A.	5P +K75E
638	A.A	5A2+L18G	924	A.A	3P+D108N	1210	A.A	5P +K75E
639	N.A.	5A2+L18H	925	N.A.	3P+N144T	1211	N.A.	5P +K123E+N156D
640	A.A	5A2+L18H	926	A.A	3P+N144T	1212	A.A	5P +K123E+N156D
641	N.A.	5A2+L18I	927	N.A.	3E (5A2+R11E+L149M+V157E)	1213	N.A.	5P +I76V
642	A.A	5A2+L18I	928	A.A	3E (5A2+R11E+L149M+V157E)	1214	A.A	5P +I76V
643	N.A.	5A2+L18K	929	N.A.	3E+D108N	1215	N.A.	5P +G48D+H57R+L92M+I99V

10

20

30

40

644	A.A	5A2+L18K	930	A.A	3E+D108N	1216	A.A	5P +G48D+H57R+ L92M+I99V
645	N.A.	5A2+L18M	931	N.A.	3E+N144T	1217	N.A.	5P +F31L+V36A+I 99V
646	A.A	5A2+L18M	932	A.A	3E+N144T	1218	A.A	5P +F31L+V36A+I 99V
647	N.A.	5A2+L18N	933	N.A.	5P (3P+D108 N+N144T)	1219	N.A.	5P+F31L+H93P
648	A.A	5A2+L18N	934	A.A	5P (3P+D108 N+N144T)	1220	A.A	5P+F31L+H93P
649	N.A.	5A2+L18P	935	N.A.	6P (5P+I56N)	1221	N.A.	5P+V90A
650	A.A	5A2+L18P	936	A.A	6P (5P+I56N)	1222	A.A	5P+V90A
651	N.A.	5A2+L18Q	937	N.A.	5E (3E+D108 N+N144T)	1223	N.A.	5P+I44V
652	A.A	5A2+L18Q	938	A.A	5E (3E+D108 N+N144T)	1224	A.A	5P+I44V
653	N.A.	5A2+L18R	939	N.A.	6E (5E+I56N)	1225	N.A.	5P+L46R+H86Q+M 106V
654	A.A	5A2+L18R	940	A.A	6E (5E+I56N)	1226	A.A	5P+L46R+H86Q+M 106V

10

20

30

40

655	N.A.	5A2+L18S	941	N.A.	NLpoly1 (5A2+R11N+A15S+L18Q+F31I+V58A+A67D+M106V+L149M+V157D)	1227	N.A.	5P+R141H
656	A.A	5A2+L18S	942	A.A	NLpoly1 (5A2+R11N+A15S+L18Q+F31I+V58A+A67D+M106V+L149M+V157D)	1228	A.A	5P+R141H
657	N.A.	5A2+L18T	943	N.A.	NLpoly2 (5A2+A15S+L18Q+F31I+V58A+A67D+M106V+L149M+V157D)	1229	N.A.	5P+N33D+V58A
658	A.A	5A2+L18T	944	A.A	NLpoly2 (5A2+A15S+L18Q+F31I+V58A+A67D+M106V+L149M+V157D)	1230	A.A	5P+N33D+V58A
659	N.A.	5A2+L18V	945	N.A.	NLpoly3 (5A2+R11N+L18Q+F31I+V58A+A67D+M106V+L149M+V157D)	1231	N.A.	5P+I56N+P157H
660	A.A	5A2+L18V	946	A.A	NLpoly3 (5A2+R11N+L18Q+F31I+V58A+A67D+M106V+L149M+V157D)	1232	A.A	5P+I56N+P157H

10

20

30

40

661	N.A.	5A2+L18W	947	N.A.	NLpoly4 (5A2+R11N+A15S+F31I+V58A+A67D+M106V+L149M+V157D)	1233	N.A.	5P+L46Q+P157H
662	A.A.	5A2+L18W	948	A.A.	NLpoly4 (5A2+R11N+A15S+F31I+V58A+A67D+M106V+L149M+V157D)	1234	A.A.	5P+L46Q+P157H
663	N.A.	5A2+L18Y	949	N.A.	NLpoly5 (5A2+R11N+A15S+L18Q+V58A+A67D+M106V+L149M+V157D)	1235	N.A.	5P+I59V
664	A.A.	5A2+L18Y	950	A.A.	NLpoly5 (5A2+R11N+A15S+L18Q+V58A+A67D+M106V+L149M+V157D)	1236	A.A.	5P+I59V
665	N.A.	5A2+F31A	951	N.A.	NLpoly6 (5A2+R11N+A15S+L18Q+F31I+A67D+M106V+L149M+V157D)	1237	N.A.	5P+A51T+E74K+P113L
666	A.A.	5A2+F31A	952	A.A.	NLpoly6 (5A2+R11N+A15S+L18Q+F31I+A67D+M106V+L149M+V157D)	1238	A.A.	5P+A51T+E74K+P113L

10

20

30

667	N.A.	5A2+F31C	953	N.A.	NLpoly7 (5A2+R11N+A15S+L18Q+F31I+V58A+M106V+L149M+V157D)	1239	N.A.	5P+V36A
668	A.A	5A2+F31C	954	A.A	NLpoly7 (5A2+R11N+A15S+L18Q+F31I+V58A+M106V+L149M+V157D)	1240	A.A	5P+V36A
669	N.A.	5A2+F31D	955	N.A.	NLpoly8 (5A2+R11N+A15S+L18Q+F31I+V58A+A67D+L149M+V157D)	1241	N.A.	5P+A51T
670	A.A	5A2+F31D	956	A.A	NLpoly8 (5A2+R11N+A15S+L18Q+F31I+V58A+A67D+L149M+V157D)	1242	A.A	5P+A51T
671	N.A.	5A2+F31E	957	N.A.	NLpoly9 (5A2+R11N+A15S+L18Q+F31I+V58A+A67D+M106V+V157D)	1243	N.A.	5P+H57R
672	A.A	5A2+F31E	958	A.A	NLpoly9 (5A2+R11N+A15S+L18Q+F31I+V58A+A67D+M106V+V157D)	1244	A.A	5P+H57R
673	N.A.	5A2+F31G	959	N.A.	NLpoly10 (5A2+R11N+A15S+L18Q+F31I+V58A+A67D+M106V+L149M)	1245	N.A.	5P+V58A

10

20

30

40

674	A.A	5A2+F31G	960	A.A	NLpoly10 (5A2+R11N+A15S+L18Q+F31I+V58A+A67D+M106V+L149M)	1246	A.A	5P+V58A
675	N.A.	5A2+F31H	961	N.A.	NLpoly11 (5A2+A15S+L18Q+M106V+L149M+V157D)	1247	N.A.	5P+E74K
676	A.A	5A2+F31H	962	A.A	NLpoly11 (5A2+A15S+L18Q+M106V+L149M+V157D)	1248	A.A	5P+E74K
677	N.A.	5A2+F31I	963	N.A.	NLpoly12 (5A2+A15S+L18Q+A67D+M106V+L149M+V157D)	1249	N.A.	5P+H86Q
678	A.A	5A2+F31I	964	A.A	NLpoly12 (5A2+A15S+L18Q+A67D+M106V+L149M+V157D)	1250	A.A	5P+H86Q
679	N.A.	5A2+F31K	965	N.A.	NLpoly13 (5A2+R11N+A15S+L18Q+M106V+L149M+V157D)	1251	N.A.	5P+H93P
680	A.A	5A2+F31K	966	A.A	NLpoly13 (5A2+R11N+A15S+L18Q+M106V+L149M+V157D)	1252	A.A	5P+H93P

10

20

30

681	N.A.	5A2+F31L	967	N.A.	5P+V	1253	N.A.	5P+I99V
682	A.A.	5A2+F31L	968	A.A.	5P+V	1254	A.A.	5P+I99V
683	N.A.	5A2+F31M	969	N.A.	5P+A	1255	N.A.	5P+K123E
684	A.A.	5A2+F31M	970	A.A.	5P+A	1256	A.A.	5P+K123E
685	N.A.	5A2+F31N	971	N.A.	5P+VT	1257	N.A.	5P+T128S
686	A.A.	5A2+F31N	972	A.A.	5P+VT	1258	A.A.	5P+T128S
687	N.A.	5A2+F31P	973	N.A.	5P+VA	1259	N.A.	5P+L142Q+T154N
688	A.A.	5A2+F31P	974	A.A.	5P+VA	1260	A.A.	5P+L142Q+T154N
689	N.A.	5A2+F31Q	975	N.A.	5P+AT	1261	N.A.	5P+H57Q
690	A.A.	5A2+F31Q	976	A.A.	5P+AT	1262	A.A.	5P+H57Q
691	N.A.	5A2+F31R	977	N.A.	5P+AA	1263	N.A.	5P+L92M
692	A.A.	5A2+F31R	978	A.A.	5P+AA	1264	A.A.	5P+L92M
693	N.A.	5A2+F31S	979	N.A.	5P+GG	1265	N.A.	5P+P113L
694	A.A.	5A2+F31S	980	A.A.	5P+GG	1266	A.A.	5P+P113L
695	N.A.	5A2+F31T	981	N.A.	5P+AA	1267	N.A.	5P+G48D
696	A.A.	5A2+F31T	982	A.A.	5P+AA	1268	A.A.	5P+G48D
697	N.A.	5A2+F31V	983	N.A.	5P+ATG	1269	N.A.	5P-B9 (-147-157)
698	A.A.	5A2+F31V	984	A.A.	5P+ATG	1270	A.A.	5P-B9 (-147-157)
699	N.A.	5A2+F31W	985	N.A.	5P+VTG	1271	N.A.	5P+L46R+P157S
700	A.A.	5A2+F31W	986	A.A.	5P+VTG	1272	A.A.	5P+L46R+P157S
701	N.A.	5A2+F31Y	987	N.A.	5P+VTA	1273	N.A.	5P+L46H+P157H
702	A.A.	5A2+F31Y	988	A.A.	5P+VTA	1274	A.A.	5P+L46H+P157H
703	N.A.	5A2+V58A	989	N.A.	5P+GTA	1275	N.A.	5P+L46R+H93P
704	A.A.	5A2+V58A	990	A.A.	5P+GTA	1276	A.A.	5P+L46R+H93P
705	N.A.	5A2+V58C	991	N.A.	5P+VTGW	1277	N.A.	5P+L46R+H93P+F31L
706	A.A.	5A2+V58C	992	A.A.	5P+VTGW	1278	A.A.	5P+L46R+H93P+F31L
707	N.A.	5A2+V58D	993	N.A.	5P+VTGWR	1279	N.A.	5P+L46R+H93P+K75E
708	A.A.	5A2+V58D	994	A.A.	5P+VTGWR	1280	A.A.	5P+L46R+H93P+K75E
709	N.A.	5A2+V58E	995	N.A.	5P+VTGWE	1281	N.A.	5P+L46R+H93P+I76V
710	A.A.	5A2+V58E	996	A.A.	5P+VTGWE	1282	A.A.	5P+L46R+H93P+I76V

10

20

30

40

711	N.A.	5A2+V58F	997	N.A.	5P+VTGWK	1283	N.A.	8S (5P+L46R+H93P+P157S+F31L)
712	A.A	5A2+V58F	998	A.A	5P+VTGWK	1284	A.A	8S (5P+L46R+H93P+P157S+F31L)
713	N.A.	5A2+V58G	999	N.A.	5P+VTGWQ	1285	N.A.	5P+L46R+H93P+P157S+K75E
714	A.A	5A2+V58G	1000	A.A	5P+VTGWQ	1286	A.A	5P+L46R+H93P+P157S+K75E
715	N.A.	5A2+V58H	1001	N.A.	5P+VTGWH	1287	N.A.	5P+L46R+H93P+P157S+I76V
716	A.A	5A2+V58H	1002	A.A	5P+VTGWH	1288	A.A	5P+L46R+H93P+P157S+I76V
717	N.A.	5A2+V58I	1003	N.A.	5P D1 (-157)	1289	N.A.	12S (8S+A51T+K75E+I76V+I107L)
718	A.A	5A2+V58I	1004	A.A	5P D1 (-157)	1290	A.A	12S (8S+A51T+K75E+I76V+I107L)
719	N.A.	5A2+V58K	1005	N.A.	5P D2 (-156-157)	1291	N.A.	11S (12-A51T)
720	A.A	5A2+V58K	1006	A.A	5P D2 (-156-157)	1292	A.A	11S (12-A51T)
721	N.A.	5A2+V58L	1007	N.A.	5P D3 (-155-157)	1293	N.A.	12S-K75E
722	A.A	5A2+V58L	1008	A.A	5P D3 (-155-157)	1294	A.A	12S-K75E
723	N.A.	5A2+V58M	1009	N.A.	5P D4 (-154-157)	1295	N.A.	12S-I76V
724	A.A	5A2+V58M	1010	A.A	5P D4 (-154-157)	1296	A.A	12S-I76V
725	N.A.	5A2+V58N	1011	N.A.	5P D5 (-153-157)	1297	N.A.	12S-I107L
726	A.A	5A2+V58N	1012	A.A	5P D5 (-153-157)	1298	A.A	12S-I107L

10

20

30

## 【 0 0 8 8 】

表 2 ( 配列番号 4 4 1 ~ 1 2 9 8 ) のポリペプチド及びコーディング核酸配列はすべて N 末端 M e t 残基 ( アミノ酸 ) または A T G 開始コドン ( 核酸 ) を含有する。一部の実施形態では、表 2 のポリペプチド及びコーディング核酸配列は N 末端 M e t 残基 ( アミノ酸 ) または A T G 開始コドン ( 核酸 ) を伴わずに提供される ( 配列番号 1 2 9 9 ~ 2 1 5 6 ) 。

40

## 【 0 0 8 9 】

特定の実施形態では、内部タグ及び／または構造相補体は配列番号 4 4 1 ~ 2 1 5 6 のアミノ酸ポリマーの 1 つを含む。一部の実施形態では、内部タグ及び／または構造相補体は配列番号 4 4 0 との単一のアミノ酸差異を含む。一部の実施形態では、内部タグ及び／または構造相補体は、配列番号 4 4 0 及び／または配列番号 4 4 1 ~ 2 1 5 6 のアミノ酸ポリマーのいずれかとの 2 以上 ( 2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30 . . . 35 . . . 40 . . . 45 . . . 50、以上 ) のアミノ酸差異を含む。一部の実施形態では、1 以上の付加、置換及び／または欠失を伴う配列

50

番号 4 4 1 ~ 2 1 5 6 のアミノ酸ポリマーの 1 つの配列を含む内部タグ及び / または構造相補体が提供される。一部の実施形態では、内部タグ及び / または構造相補体またはその一部は、配列番号 4 4 1 ~ 2 1 5 6 のアミノ酸ポリマーの 1 以上との 7 0 % を超える（たとえば、> 7 1 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 1 %、> 9 2 %、> 9 3 %、> 9 4 %、> 9 5 %、> 9 6 %、> 9 7 %、> 9 8 %、または > 9 9 %）配列同一性を含む。

#### 【 0 0 9 0 】

特定の実施形態では、表 2 の核酸が提供される。一部の実施形態では、表 2 のポリペプチド（たとえば、対象とするポリペプチドに挿入される）をコードする核酸が提供される。一部の実施形態では、本発明の核酸は、配列番号 4 4 0 及び / または配列番号 4 4 1 ~ 2 1 5 6 のアミノ酸ポリマーのいずれか（たとえば、対象とするポリペプチドに挿入された）との単一のアミノ酸差異を含むポリペプチドをコードする。一部の実施形態では、核酸は、配列番号 4 4 0 及び / または表 2 でリストにされたポリペプチドのいずれか（たとえば、対象とするポリペプチドに挿入される）との 2 以上（2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30...35...40...45...50、以上）のアミノ酸差異を含むポリペプチドをコードする。一部の実施形態では、配列番号 4 4 1 ~ 2 1 5 6 の核酸ポリマーの 1 つ（たとえば、対象とするポリペプチドに挿入される）の配列を含む核酸が提供される。一部の実施形態では、1 以上の付加、置換及び / または欠失を伴う配列番号 4 4 1 ~ 2 1 5 6 の核酸ポリマーの 1 つの配列を含む核酸が提供される。一部の実施形態では、核酸またはその一部は、配列番号 4 4 1 ~ 2 1 5 6 の核酸ポリマーの 1 以上（たとえば、対象とするポリペプチドに挿入される）との 7 0 % を超える（たとえば、> 7 1 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 1 %、> 9 2 %、> 9 3 %、> 9 4 %、> 9 5 %、> 9 6 %、> 9 7 %、> 9 8 %、または > 9 9 %）配列同一性を含む。一部の実施形態では、核酸またはその一部は、配列番号 4 4 1 ~ 2 1 5 6 のアミノ酸ポリマーの 1 以上（たとえば、対象とするポリペプチドに挿入される）との 7 0 % を超える（たとえば、> 7 1 %、> 7 5 %、> 8 0 %、> 8 5 %、> 9 0 %、> 9 1 %、> 9 2 %、> 9 3 %、> 9 4 %、> 9 5 %、> 9 6 %、> 9 7 %、> 9 8 %、または > 9 9 %）配列同一性を含むポリペプチドをコードする。一部の実施形態では、配列番号 4 4 1 ~ 2 1 5 6 のポリペプチドの 1 つ（たとえば、対象とするポリペプチドに挿入される）をコードする核酸が提供される。一部の実施形態では、1 以上の付加、置換及び / または欠失を伴う配列番号 4 4 1 ~ 2 1 5 6 のポリペプチドの 1 つ（たとえば、対象とするポリペプチドに挿入される）をコードする核酸が提供される。

#### 【 0 0 9 1 】

本発明は、基礎研究、医学研究、分子診断等を含む種々の分野で有用である組成物及び方法を提供する。本明細書に記載されている試薬及びアッセイは特定の適用に限定されず、有用な適用は本発明の範囲内にあると見なされるべきである。

#### 【 0 0 9 2 】

本発明の実施形態を利用する典型的な適用には、タンパク質 / タンパク質の相互作用（たとえば、ヘテロ二量体、ホモ二量体）（図 1 を参照）、タンパク質 / R N A の相互作用、タンパク質 / D N A の相互作用、タンパク質 / 小分子の相互作用、または分子実体の他の組み合わせのモニタリング / 検出が関与する。対象とするタンパク質は内部でタグ付けされ、対象とする第 2 の実体は構造相補体に連結される。特定のアッセイ条件下で検出可能なシグナルが生成されると、そのとき、対象とするタンパク質と対象とする実体との相互作用が推測される。そのようなアッセイは、好適な条件下（たとえば、試験管内、生体内、そのまま、動物全体等）で分子の相互作用をモニターするのに有用であり、たとえば、創薬、分子経路を解明すること、複合体構築の平衡または動態の態様を検討すること、高処理能力のスクリーニング、近接センサー等にて使用される。

#### 【 0 0 9 3 】

本発明の実施形態を利用する他の典型的な適用には、タンパク質またはポリペプチドの

検出または局在確認（たとえば、細胞での局在確認、細胞内局在確認、等）が関与する。試料における対象とするタンパク質は内部でタグ付けされ、構造相補体が試料に添加される。特定のアッセイ条件下で検出可能なシグナルが生成されると、そのとき、対象とするタンパク質の存在または位置が推測される。そのようなアッセイは好適な条件下（たとえば、試験管内、生体内、そのまま、動物全体等）でタンパク質を検出するまたは局在確認するのに有用であり、たとえば、創薬、分子経路を解明すること、複合体構築の平衡または動態の態様を検討すること、高処理能力のスクリーニング、近接センサー等にて使用される。

#### 【0094】

一部の実施形態では、既知の特徴（たとえば、スペクトルの特徴、対の相互親和性）の内部タグ及び構造相補体を用いて、対象とするタンパク質及び潜在的に会合する対象とする実体（タンパク質、核酸、小分子等）との親和性を解明し、またはその相互作用を理解する。他の実施形態では、よく特徴付けられた相互作用対を用いて内部タグ及び構造相補体の特徴（たとえば、スペクトルの特徴、対の相互親和性）を決定する。

#### 【0095】

本明細書に記載されている実施形態は薬物スクリーニング及び／または薬物開発で使用される。たとえば、小分子薬物または小分子のライブラリ全体（たとえば、構造相補体で標識された）の内部でタグ付けされた対象とする標的タンパク質（たとえば、治療用標的）との相互作用は、1以上の関連する条件（たとえば、生理的条件、病状等）下でモニターされる。他の実施形態では、2つの実体（たとえば、受容体とリガンド、タンパク質／タンパク質等）間の相互作用を高めるまたは阻害する小分子薬物または小分子のライブラリ全体の能力がアッセイされる。一部の実施形態では、高処理能力形式にて薬物スクリーニングの応用を行って、何万もの異なる分子の標的への結合の検出を可能にし、または他の実体の結合に対するこれら分子の効果を調べる。

#### 【0096】

一部の実施形態では、本発明は生きている生物（たとえば、細菌、酵母、真核生物、哺乳類、霊長類、ヒト等）及び／または細胞における分子相互作用の検出を提供する。一部の実施形態では、内部でタグ付けされたタンパク質と相補体で標識されたタンパク質が細胞にてまたは生物全体にて同時発現し、シグナルが検出され、相互作用複合体の形成と関連付けられる。一部の実施形態では、細胞は、ベクター（たとえば、内部でタグ付けされたタンパク質、相補体で標識されたタンパク質等をコードする）で一時的に及び／または安定的に形質転換される、または形質移入される。一部の実施形態では、本明細書に記載されているアッセイを実施するのに必要な成分（たとえば、内部でタグ付けされたタンパク質、相補体で標識されたタンパク質等）をコードする遺伝子導入生物が生成される。他の実施形態では、ベクターは生物全体に注入される。

#### 【0097】

本発明はまた、内部タグ及び構造相補体及びそれらから生じる生物発光複合体の設計及び／または最適化の方法を提供する。本明細書に記載されている実施形態に一致する非発光性の対／基及び／またはそのパネルの設計のための好適な方法は本発明の範囲内である。

#### 【実施例】

#### 【0098】

##### 実施例 1

本発明の実施形態の開発の間に実験を行って、実験モデルとしての非発光性ポリペプチドNLpoly11Sと高親和性非発光性ペプチドNLpep86との間での構造的な相補性を明らかにした。NLpep86の挿入のための標的としてHaloTagを選択した。HaloTag（登録商標）タンパク質の中での種々の部位（挿入部位：18/19、32/33、78/79、98/99）にて直列型のNLpep86（高親和性、配列GSSG-[VSGWRLFKKIS]-E-[VSGWRLFKKIS]-GSSG）を挿入することによって多数のHaloTag/NLpep86融合タンパク質を生成し

10

20

30

40

50

た。当初の実験は、NLpoly11Sと示したHaloTag/NLpep86融合タンパク質で一時的に形質移入したHeLa細胞で行った。結果は、NLpoly11SとHaloTag内の様々な位置に挿入したNLpep86との間で構造的な相補性を達成することが可能であることを示している(図4)。TMR-HTRリガンドで標識した細胞のBRET(HaloTagリガンドを結合する修飾されたHaloTagの能力を必要とする)(図4)または画像化(図5)によってHaloTagの機能を測定した。NLpep86の挿入はHaloTagの機能に適合することが明らかにされた(図5)。内部融合体を用いた構造的な相補性の観察された効率は、NLpoly11SとNLpep86を用いたN末端またはC末端のNLpep86のHaloTagへの融合に対して1~40%の間で変化する。

10

【0099】

## 実施例2

NLpep114で内部にタグ付けした標的タンパク質とNLpoly11Sでタグ付けしたプロテインGとを用いて試験抗体が主導するNANOLUCの生物発光相補性を明らかにする実験を実施した。

【0100】

## プロテインG-11Sの構築、発現及び精製

ATG-2071(NLpoly11Sでタグ付けしたプロテインG)プラスミド(配列番号2576):免疫グロブリンGに結合するプロテインG[Uniprot P19909]に由来するアミノ酸303~497を合成遺伝子(GenScript)から増幅し、6xHisタグを付加し、リンカーNLpoly11Sを含有するpF5K(Flexiベクター、CMVプロモータ)にクローニングした。次いで細菌発現のために、6xHis-プロテインG-NLpoly11S融合体をpF1A(Flexiベクター、T7プロモータ;Promega)にサブクローニングした。

20

【0101】

グルコース/ラムノース自己誘導法によって大腸菌の発現系にてNLpoly11Sでタグ付けしたプロテインGを産生させた。手短には、製造元が推奨するプロトコルを用いてプラスミドATG-2071で大腸菌KRX細胞(Promega)を形質転換し、抗生剤を含有するLB培地(50ml)にて振盪しながら(275rpm)37℃で17~22時間増殖させた。このスターター培養物を250mlの自己誘導培地(グルコースとラムノース(各0.05%)と抗生剤を伴ったLB培地)にて希釈し(1:100)、振盪しながら(275rpm)25℃で17~22時間増殖させた。遠心分離(5,000xgで4分)にて20分間)によって細胞を回収し、培地を取り除き、細菌細胞のペレットを-20℃で保存した。

30

【0102】

ペレットにした細胞を50mlの溶解緩衝液(100mMのHEPES(pH7.5), 500mMのNaCl, 10mMのイミダゾール, 0.5xのFastBreak(Promega), 1xプロテアーゼ阻害剤カクテル(Promega), 0.2mg/mlのリゾチーム(Sigma)及び250単位のRQ1 DNase(Promega))にて再浮遊させ、次いで時々混合しながら常温にて30分間インキュベートした。遠心分離(4分にて15,000xgで20分間)によって可溶性分画を分離し、出発緩衝液(50mMのHEPES(pH7.5), 500mMのNaCl, 10mMのイミダゾール)で平衡化した5mlのHisTrapカラム(Life Technologies)にかけた(1ml/分)。試料の適用後、4CVの出発緩衝液で樹脂を洗浄した。結合したタンパク質を20CVの線形勾配、0~100%の最終緩衝液(50mMのHEPES(pH7.5), 500mMのイミダゾール)によって溶出した。分画(2.5ml)をSDS/PAGEによって解析した。十分量の主要な41kDaのバンド(NLpoly11S/pG)と軽微な混入物を伴ったものをプールし、1xPBSに対して透析し、-20℃で保存した。

40

【0103】

50

## VEGFA構築物についての発現プラスミド構築の方法

NLpep114及び/またはFLAGオクタペプチド(双方とも合成遺伝子; Gene Dynamics)のいずれかに融合させたVEGFA-165をベクターpCIHN(Flexiベクター, CMVプロモータ; Promega)に移すことによってVEGFA構築物ATG-1915(配列番号2577)、-1917(配列番号2578)及び-1946(配列番号2579)を作った。このベクターはIL6分泌シグナルと共にN末端のHalotagを含有する。

### 【0104】

#### 一般的な細胞形質移入のプロトコール

1:10の質量比にて10ugの総DNAでNLpep114標的融合体構築物DNAをキャリアDNA(pGEM3Zf(-); Promega)にて希釈した。製造元(Promega)のプロトコールに従って、1:3(ugDNA/ulFuGENE)の比でDNA:FuGENE複合体を形成した。DMEM(Gibco)+10%FBS(Hyclone)にて $2 \times 10^5$ 個/mlの密度で浮遊させたHEK293T細胞(ATCC)の20部(容積/容積)と1部の形質移入複合体を混合した。細胞(50ul/ウェル)を96穴組織培養プレートに分配し、37/5%CO<sub>2</sub>で加湿したインキュベータにて18~24時間インキュベートした。

### 【0105】

#### 標的抗体が主導するNANOLUC生物発光相補性

上記で記載したような3種のNLpep114-VEGFAのDNAによってHEK293T細胞(ATCC)に形質移入し、一晚インキュベートした。培地を等量のopti-MEM(Gibco)で置き換えることによって同一条件下で細胞を4時間血清飢餓状態に置いた。1xPBS/0.1%BSA(Promega)におけるNLpoly11S(15)pGを0.5ug/mlの最終濃度(12nM)で加えた(25ul/ウェル)。1xPBS/0.1%BSA(Promega)における抗VEDF抗体(R&D Systems、#293)を0~0.73ug/ml(0~5.3nM)の最終濃度で加えた(25ul/ウェル)。LCS試薬(Promega、100ul/ウェル、10uMの最終濃度)の添加の後、IngenuityF500マイクロタイタープレートリーダー(Tacana)を用いてルシフェラーゼ活性を測定した。

### 【0106】

図6~11は、一緒にインキュベートすると、NLpep114-VEGF融合タンパク質とNLpoly11S-プロテインG融合タンパク質と未修飾の抗-VEGF抗体とが一緒になって活性のあるルシフェラーゼ複合体を形成することを明らかにしている。シグナルは測定可能ではあるが、抗体の非存在下では極度に低い。NLpoly11S/NLpep114相互作用の高いK<sub>D</sub>を考えると、促進性ではない相補性は使用される濃度では極度に低いはずである。このことは、NLpoly11S断片が若干低いレベルの残留ルシフェラーゼ活性を有することを示している。ATG-1915(HT-VEGF-114)については、抗VEGF抗体の濃度が0から0.73ug/ml(0~5.3nM)に上昇するにつれてこのバックグランドシグナルは400倍を超えて上昇した。

### 【0107】

NLpep114タグのコンテキストはシステム全体の性能に影響を及ぼす。C末端(外側)のNLpep114タグを伴うATG-1915は最も高いシグナルを有する。C末端FLAG配列(ATG-1946)の添加によってNLpep114タグがやや内部移行すると、この構築物は40%のシグナル喪失を被る。2つの大きなドメイン間に置かれると、シグナルは5倍低下する(ATG-1917)。総シグナルはNLpep114タグの位置で変化する一方で、算出されるEC<sub>50</sub>は一定のままであることを留意のこと。NLpep114タグのコンテキストは相補性のレベルを変えるが、NLpoly11S/NLpep114対の親和性を変えることはない。

### 【0108】

#### 実施例3

結合対 F K B P と F r b を用いて促進性の N A N O L U C 生物発光相補性を明らかにする実験を実施した ( 図 1 3 )。

#### 【 0 1 0 9 】

形質移入はすべて、プレートに播くことに先立って形質移入複合体を細胞の浮遊液と混合することにより逆形質移入として行った。手短には、500 u l の O p t i M E M と 5 u g の D N A と 1 5 u l の F u g e n e H D ( P r o m e g a ) を含有する形質移入ミックス ( 1 枚の 9 6 穴プレートに十分な ) を作った。相補性対 ( F r b - X / F K B P - Y ) の D N A は 1 : 1 の比だった。

#### 【 0 1 1 0 】

形質移入については、相補性対の D N A を希釈しないで、または 1 : 5 0 の希釈で使  
10  
した。キャリア D N A として p G E M 3 Z を用いて総 D N A 含量を 5 u g に合わせた。穏やかなボルテックスによって形質移入ミックスを混合し、使用に先立って室温で 5 ~ 1 0 分間インキュベートした。

#### 【 0 1 1 1 】

トリプシン処理によって細胞を回収し、洗浄し、D M E M + 1 0 % F B S にて  $2 \times 10^5$  個 / m l の濃度に希釈した。形質移入については、0 . 5 m l の形質移入ミックスを 1 0 m l の細胞浮遊液に加えた。次いで白色 9 6 穴組織培養プレート ( ウェル当たり 1 0 0 u l ) のウェルに細胞浮遊液を入れ、3 7 にて一晩インキュベートした。

#### 【 0 1 1 2 】

形質移入した細胞で 3 つの異なるアッセイを行った。  
20

a ) 単一濃度のラパマイシンを用いたエンドポイントアッセイ ( 図 1 4 及び 1 5 )

形質移入した細胞上の増殖培地 ( D M E M + 1 0 % F B S ) を吸引によって取り除き、ラパマイシン ( 1 m M ) とフリマジン ( 1 0 m M ) を含む 1 0 0 u l の O p t i M E M を加えた。細胞を室温で 1 0 分間インキュベートし、B M G C l a r i o s t a r または G l o m a x M u l t i プラスプレートリーダーで発光を読み取った。

b ) エンドポイントアッセイ : ラパマイシンの用量反応 ( 図 1 6 )

形質移入した細胞上の増殖培地 ( D M E M + 1 0 % F B S ) を吸引によって取り除き、ラパマイシンの連続希釈とフリマジン ( 1 0 m M ) を含む 1 0 0 u l の O p t i M E M を加えた。細胞を室温で 1 0 分間インキュベートし、B M G C l a r i o s t a r または G l o m a x M u l t i プラスプレートリーダーで発光を読み取った。  
30

c ) 動態アッセイ ( 図 1 7 及び 1 8 )

形質移入した細胞上の増殖培地 ( D M E M + 1 0 % F B S ) を吸引によって取り除き、フリマジン ( 1 0 m M ) を含む 5 0 u l の O p t i M E M を加えた。発光の検出は B M G C l a r i o s t a r プレートリーダーで開始し、ラパマイシン ( 1 m M ) とフリマジン ( 1 0 m M ) を含む 5 0 u l の O p t i M E M を細胞上に注入した。発光を連続的に読み取った。

#### 【 0 1 1 3 】

#### 実施例 4

内部の高親和性 N L p e p は種々の実施形態で使用される。N 末端も C 末端もタンパク質タグの連結について魅力的な点を表さない場合が生じる。たとえば、  
40

1 ) タンパク質の末端が所望の細胞局在に存在しない。たとえば、所与の膜タンパク質について、細胞外の部位でタグを有することが所望であってもよいが、双方の末端は細胞内にある。

2 ) タグの末端付加がタンパク質 / タンパク質の相互作用を妨害する。たとえば、多数の膜タンパク質 ( たとえば、A D R B 2 ) は、まさに C 末端に P D Z 結合モチーフを有する。C 末端タグの付加はこれらの相互作用を無効にし、適正なタンパク質の機能を変化させる。

3 ) タグにとって末端が許容するよりもタンパク質の所与の位置に空間的に接近して置かれることが望まれる。

4 ) N 末端でのタグの配置が適正なシグナル配列の機能及び切断を妨げる。  
50

5) 末端がすでに他のタグまたは融合タンパク質のために使用されている。

【0114】

#### 実施例 5

内部の高親和性 N L p e p は膜タンパク質の表面発現の測定に使用される。細胞表面に発現した所与のタンパク質の量を測定することが一般に望まれる。これは、

- ・受容体の活性化及び内部移行
- ・エンドソームからの受容体の再利用
- ・調節されたエキソサイトーシス
- ・タンパク質の輸送及び分泌

に関する検討を可能にする。

10

【0115】

一部の実施形態では、精製された N L p o l y、たとえば、N L p o l y 1 1 S、タンパク質に加えてフリマジン基質を細胞外媒体に加えることができるように以下の実験を構成した。原形質膜の細胞外側での高親和性 N L p e p 配列、たとえば、N L p e p 8 0 との相補性は自然発生的な相補性をもたらすことができ、表面上のタンパク質の量に直接比例する発光シグナルを生じる。

【0116】

a) 嚢胞性線維症膜貫通コンダクタンス調節因子 (C F T R) の F 5 0 8 d e l 変異は嚢胞性線維症の最も一般的な原因であり、タンパク質が原形質膜に正しく標的化されるのを妨げ、代わりにそれが分解される。表面への輸送の増加を促進することができる小分子シャペロンが特定されている。そのような小分子をスクリーニングするために、C F T R の表面発現を定量するための単純な発光アッセイを有することは有益である。

20

【0117】

高親和性の N L p e p によって C F T R にタグ付けすることは表面発現の定量を可能にするが、C F T R の N 末端及び C 末端は双方ともに細胞内にある。したがって、N L p e p タグをタンパク質の細胞外ループの 1 つに置く。一部の実施形態では、F l a g エピトープタグを残基 A s n 9 0 1 の後に予め置いて細胞外タグを提供する。C F T R の F 5 0 8 d e l 変異体における同じ位置での高親和性 N L p e p 配列の挿入は以下の 1 以上を可能にする。

1) 原形質膜でのタンパク質の量の単純な定量。化合物ライブラリ及び適正な輸送を促進することが知られている陽性対照化合物で細胞を処理し、N L p o l y、たとえば、N L p o l y 1 1 S とフリマジンの存在下で生細胞にて発光を測定する。

30

2) 細胞におけるタンパク質の総量を定量するために N L p o l y、たとえば、N L p o l y 1 1 S とフリマジンを含有する溶解試薬で細胞を処理する。タンパク質分解の低下は蛍光シグナルを増やすことになる。

3) 緩衝液中の N L p o l y、たとえば、N L p o l y 1 1 S とフリマジンのプロット膜への添加によってタンパク質プロットにおけるバンドのシフトとして、変異中に生じる C F T R のグリコシル化が容易に検出される。

【0118】

b) 原形質膜の内外での神経伝達物質受容体の輸送は厳しく調節されている。A M P A 受容体 (A M P A R) は迅速な興奮性シナプス伝達に介在し、シナプス強度はシナプス後膜における A M P A R の組成によって決定され、それは A M P A R サブユニットの調節された輸送によって制御される。A M P A 受容体の細胞外ループに高親和性 N L p e p を挿入することによってタンパク質のレベル及びエキソサイトーシスとエンドサイトーシスの動態の簡単な測定が可能になる。

40

【0119】

上記明細書で言及された出版物及び特許はすべて参照によって本明細書に組み入れられる。本発明の記載されている方法及び系の種々の改変及び変化は本発明の範囲及び精神から逸脱することなく当業者に明らかであろう。本発明は特定の実施形態と関連して記載されているが、請求されるような本発明はそのような特定の実施形態に過度に限定されるべ

50

きではないことが理解されるべきである。実際、関連する分野の当業者に自明である本発明を実施するための記載されている方法の種々の改変は本発明の範囲内にあるように意図されている。

本発明の好ましい態様は、下記の通りである。

〔１〕N末端断片と、C末端断片と、内部タグとを含むポリペプチドであって、前記内部タグが、対象とするタンパク質の中に挿入された配列番号２との１００％未満でかつ３０％を超える配列同一性を有するアミノ酸配列を含み、前記内部タグが配列番号４４０のポリペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、前記ポリペプチド。

〔２〕前記N末端断片及び前記C末端断片の両方が、少なくとも２アミノ酸の長さである、前記〔１〕に記載のポリペプチド。

〔３〕前記N末端断片及び／又は前記C末端断片が、少なくとも２０アミノ酸の長さである、前記〔２〕に記載のポリペプチド。

〔４〕前記内部タグが、配列番号２のペプチドに比べて１つ以上の形質の向上を示し、前記形質が、配列番号４４０のポリペプチドに対する親和性、発現、細胞内溶解性、細胞内安定性、及び配列番号４４０のポリペプチドと組み合わせたときの生物発光活性から選択される、前記〔１〕に記載のポリペプチド。

〔５〕前記内部タグの前記アミノ酸配列が、表１のペプチドから選択される、前記〔１〕に記載のポリペプチド。

〔６〕前記N末端断片と前記C末端断片とが、前記内部タグの非存在下で直接連結される場合、対象とする第１のタンパク質の配列を含む、前記〔１〕に記載のポリペプチド。

〔７〕前記内部タグが、リンカーペプチドによって前記N末端断片及び／又は前記C末端断片に接続される、前記〔１〕に記載のポリペプチド。

〔８〕前記〔１〕に記載のポリペプチドをコードする配列を含む、核酸。

〔９〕

（ａ）前記〔１〕に記載のポリペプチドと、

（ｂ）配列番号４４０との１００％未満でかつ３０％を超える配列同一性を有する相補体ポリペプチドを含む第２のポリペプチドと、  
を含む、生物発光複合体。

〔１０〕前記内部タグ及び前記相補体ポリペプチドが、互いに低親和性を有する、前記〔９〕に記載の生物発光複合体。

〔１１〕前記第２のポリペプチドが、対象とする第２のタンパク質との融合体である、前記〔１０〕に記載の生物発光複合体。

〔１２〕前記融合体が、内部融合体又は従来の融合体である、前記〔１１〕に記載の生物発光複合体。

〔１３〕前記対象とする第２のタンパク質が、前記N末端断片及び／又は前記C末端断片のすべて又は一部に対して高親和性を有する、前記〔１２〕に記載の生物発光複合体。

〔１４〕前記第２のポリペプチドが、対象とする分子に連結されている、前記〔１０〕に記載の生物発光複合体。

〔１５〕前記N末端断片及び／又は前記C末端断片のすべて又は一部が、前記対象とする分子に対して高親和性を有する、前記〔１４〕に記載の生物発光複合体。

〔１６〕セレンテラジン基質を更に含む、前記〔９〕に記載の生物発光複合体。

〔１７〕前記内部タグ及び前記相補体ポリペプチドが、互いに高親和性を有する、前記〔９〕に記載の生物発光複合体。

〔１８〕前記第２のポリペプチドが、融合ポリペプチドではないか、又は対象とする分子に連結されていない、前記〔９〕に記載の生物発光複合体。

〔１９〕前記相補体ポリペプチドの前記アミノ酸配列が、表２のペプチドから選択される、前記〔９〕に記載の生物発光複合体。

〔２０〕N末端断片と、C末端断片と、内部タグとを含むポリペプチドであって、前記内部タグが、対象とするタンパク質の中に挿入された配列番号４４０との１００％未満でか

10

20

30

40

50

つ 30 % を超える配列同一性を有するアミノ酸配列を含み、検出ペプチドが配列番号 2 のポリペプチドに接触すると、基質の存在下で、検出可能な生物発光シグナルが生成される、前記ポリペプチド。

〔 2 1 〕 前記 N 末端断片及び前記 C 末端断片の両方が、少なくとも 20 アミノ酸の長さである、前記〔 2 0 〕に記載のポリペプチド。

〔 2 2 〕 前記 N 末端断片及び / 又は前記 C 末端断片が、少なくとも 50 アミノ酸の長さである、前記〔 2 1 〕に記載のポリペプチド。

〔 2 3 〕 前記内部タグが、配列番号 440 のペプチドと比べて 1 つ以上の形質の向上を示し、前記形質が、配列番号 2 のポリペプチドに対する親和性、発現、細胞内溶解性、細胞内安定性、及び配列番号 2 のポリペプチドと組み合わせたときの生物発光活性から選択される、前記〔 2 0 〕に記載のポリペプチド。

〔 2 4 〕 前記内部タグの前記アミノ酸配列が、表 2 のペプチドから選択される、前記〔 2 0 〕に記載のポリペプチド。

〔 2 5 〕 前記 N 末端断片及び前記 C 末端断片が、前記内部タグの非存在下で直接連結される場合、対象とする第 1 のタンパク質の配列を含む、前記〔 2 0 〕に記載のポリペプチド。

〔 2 6 〕 前記〔 2 0 〕に記載のポリペプチドをコードする配列を含む、核酸。

〔 2 7 〕

( a ) 前記〔 2 0 〕に記載のポリペプチドと、

( b ) 配列番号 2 との 100 % 未満でかつ 30 % を超える配列同一性を有する相補体ペプチドと、  
を含む、生物発光複合体。

〔 2 8 〕 前記内部タグ及び前記相補体ペプチドが、互いに低親和性を有する、前記〔 2 7 〕に記載の生物発光複合体。

〔 2 9 〕 前記相補体ペプチドが、対象とする第 2 のタンパク質との融合体である、前記〔 2 8 〕に記載の生物発光複合体。

〔 3 0 〕 前記融合体が、内部融合体又は従来の融合体である、前記〔 2 9 〕に記載の生物発光複合体。

〔 3 1 〕 前記対象とする第 2 のタンパク質が、前記 N 末端断片及び / 又は前記 C 末端断片のすべて又は一部に対して高親和性を有する、前記〔 3 0 〕に記載の生物発光複合体。

〔 3 2 〕 前記相補体ペプチドが、対象とする分子に連結されている、前記〔 2 8 〕に記載の生物発光複合体。

〔 3 3 〕 前記 N 末端断片及び / 又は前記 C 末端断片のすべて又は一部が、前記対象とする分子に対して高親和性を有する、前記〔 2 7 〕に記載の生物発光複合体。

〔 3 4 〕 セレンテラジン基質を更に含む、前記〔 2 7 〕に記載の生物発光複合体。

〔 3 5 〕 前記内部タグ及び前記相補体ペプチドが、互いに高親和性を有する、前記〔 2 7 〕に記載の生物発光複合体。

〔 3 6 〕 前記相補体ペプチドが、表 1 の前記ペプチドから選択される、前記〔 2 7 〕に記載の生物発光複合体。

〔 3 7 〕 前記相補体ペプチドが、融合ポリペプチドではないか、又は対象とする分子に連結されていない、前記〔 9 〕に記載の生物発光複合体。

〔 3 8 〕 第 1 のアミノ酸配列と第 2 のアミノ酸配列との間での安定な相互作用の検出方法であって、

( a ) 内部タグが第 1 のアミノ酸配列の N 末端にも C 末端にもないように、内部タグを第 1 のアミノ酸配列に挿入することによって内部融合体を作り出す工程であって、前記内部タグが、配列番号 2 との 100 % 未満でかつ 30 % を超える配列同一性を有し、前記内部タグが配列番号 440 のポリペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

( b ) 第 2 のアミノ酸配列と相補体ポリペプチドとの第 2 の融合体を作り出す工程であって、前記相補体ポリペプチドが、配列番号 440 との 100 % 未満でかつ 30 % を超える

10

20

30

40

50

配列同一性を有し、前記相補体ポリペプチドが配列番号 2 のペプチドに接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

(c) 前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間で見込まれる安定な相互作用を生じさせる条件に、前記内部融合体と、第 2 の融合体と、セレンテラジン基質とを置く工程と、

(d) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの検出が、前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間での安定な相互作用を示す、工程と、

を含む、前記検出方法。

[ 39 ] 前記第 2 の融合体が、内部融合体又は従来の融合体である、前記 [ 38 ] に記載の方法。

[ 40 ] 前記内部融合体が、前記第 1 のアミノ酸配列と前記内部タグとをコードする第 1 の核酸配列から発現され、前記第 2 の融合体が、前記第 2 のアミノ酸配列と前記相補体ポリペプチドとをコードする第 2 の核酸配列から発現される、前記 [ 38 ] に記載の方法。

[ 41 ] 単一のベクターが、前記第 1 の核酸配列と前記第 2 の核酸配列とを含む、前記 [ 40 ] に記載の方法。

[ 42 ] 前記第 1 の核酸配列及び前記第 2 の核酸配列が、別個のベクターにある、前記 [ 40 ] に記載の方法。

[ 43 ] 工程 ( a ) 及び ( b ) が、細胞内で前記内部融合体と第 2 の融合体とを発現させる工程を含む、前記 [ 40 ] に記載の方法。

[ 44 ] 第 1 のアミノ酸配列と第 2 のアミノ酸配列との間での安定な相互作用の検出方法であって、

( a ) 内部タグが第 1 のアミノ酸配列の N 末端にも C 末端にもないように、内部タグを第 1 のアミノ酸配列に挿入することによって内部融合体を作り出す工程であって、前記内部タグが、配列番号 440 との 100 % 未満でかつ 30 % を超える配列同一性を有し、前記内部タグが配列番号 2 のペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

( b ) 第 2 のアミノ酸配列と相補体ペプチドとの第 2 の融合体を作り出す工程であって、前記相補体ペプチドが、配列番号 2 との 100 % 未満でかつ 30 % を超える配列同一性を有し、前記相補体ペプチドが配列番号 440 のポリペプチドに接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

( c ) 前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間で見込まれる安定な相互作用を生じさせる条件に、前記内部融合体と、第 2 の融合体と、セレンテラジン基質とを置く工程と、

( d ) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの検出が、前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間での安定な相互作用を示す、工程と、

を含む、前記検出方法。

[ 45 ] 前記第 2 の融合体が、内部融合体又は従来の融合体である、前記 [ 44 ] に記載の方法。

[ 46 ] 前記内部融合体が、前記第 1 のアミノ酸配列と前記内部タグとをコードする第 1 の核酸配列から発現され、前記第 2 の融合体が、前記第 2 のアミノ酸配列と前記相補体ペプチドとをコードする第 2 の核酸配列から発現される、前記 [ 44 ] に記載の方法。

[ 47 ] 単一のベクターが、前記第 1 の核酸配列と前記第 2 の核酸配列とを含む、前記 [ 46 ] に記載の方法。

[ 48 ] 前記第 1 の核酸配列及び前記第 2 の核酸配列が、別個のベクターにある、前記 [ 46 ] に記載の方法。

[ 49 ] 工程 ( a ) 及び ( b ) が、細胞内で前記内部融合体と第 2 の融合体とを発現させる工程を含む、前記 [ 46 ] に記載の方法。

[ 50 ] 試料における標的ポリペプチドの検出方法であって、

10

20

30

40

50

( a ) 内部タグが標的ポリペプチドのN末端にもC末端にもないように、内部タグを標的ポリペプチドに挿入することによって内部融合体を作り出す工程であって、前記内部タグが、配列番号440との100%未満でかつ30%を超える配列同一性を有し、前記内部タグが配列番号2のペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

( b ) 前記試料に

( i ) 配列番号2との100%未満でかつ30%を超える配列同一性を有する相補体ペプチドと、

( i i ) セレンテラジン基質と、

を加える工程と、

( c ) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの検出が、前記試料における前記標的ポリペプチドの存在を示す、工程と

、

を含む、前記検出方法。

[ 5 1 ] 前記試料が、細胞を含む、前記 [ 5 0 ] に記載の方法。

[ 5 2 ] 工程 ( a ) が、前記細胞内で前記内部融合体を発現させる工程を含む、前記 [ 5 1 ] に記載の方法。

[ 5 3 ] 工程 ( b ) ( i ) が、前記細胞中の前記相補体ペプチドを含む、前記 [ 5 2 ] に記載の方法。

[ 5 4 ] 試料における標的ポリペプチドの検出方法であって、

( a ) 内部タグが標的ポリペプチドのN末端にもC末端にもないように、内部タグを標的ポリペプチドに挿入することによって内部融合体を作り出す工程であって、前記内部タグが、配列番号2との100%未満でかつ30%を超える配列同一性を有し、前記内部タグが配列番号440のペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

( b ) 前記試料に

( i ) 配列番号440との100%未満でかつ30%を超える配列同一性を有する相補体ポリペプチドと、

( i i ) セレンテラジン基質と、

を加える工程と、

( c ) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの検出が、前記試料における前記標的ポリペプチドの存在を示す、工程と

、

を含む、前記検出方法。

[ 5 5 ] 前記試料が、細胞を含む、前記 [ 5 4 ] に記載の方法。

[ 5 6 ] 工程 ( a ) が、前記細胞内で前記内部融合体を発現させる工程を含む、前記 [ 5 5 ] に記載の方法。

[ 5 7 ] 工程 ( b ) ( i ) が、前記細胞中の前記相補体ポリペプチドを含む、前記 [ 5 5 ] に記載の方法。

[ 5 8 ] 検出試薬であって、

( a ) 配列番号440との100%未満でかつ30%を超える配列同一性を有するアミノ酸配列を含む相補体ポリペプチドであって、前記ポリペプチドが配列番号2のペプチドと接触すると、基質の存在下で、検出可能な生物発光シグナルが生成される、相補体ポリペプチドと、

( b ) 前記ポリペプチドと配列番号2のペプチドとによって生成される生物発光複合体のための基質と、

を含む、前記検出試薬。

[ 5 9 ] 検出試薬であって、

( a ) 配列番号2との100%未満でかつ30%を超える配列同一性を有するアミノ酸配列を含む相補体ペプチドであって、前記ペプチドが配列番号440のポリペプチドと接触

10

20

30

40

50

すると、基質の存在下で、検出可能な生物発光シグナルが生成される、相補体ペプチドと、

(b) 前記ペプチドと配列番号 440 のポリペプチドとによって生成される生物発光複体のための基質と、

を含む、前記検出試薬。

[60] 第 1 のアミノ酸配列と第 2 のアミノ酸配列との間での相互作用の、潜在的阻害剤による変化の検出方法であって、

(a) 内部タグが第 1 のアミノ酸配列の N 末端にも C 末端にもないように、内部タグを第 1 のアミノ酸配列に挿入することによって内部融合体を作り出す工程であって、前記内部タグが、配列番号 2 との 100 % 未満でかつ 30 % を超える配列同一性を有し、前記内部タグが配列番号 440 のポリペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

(b) 第 2 のアミノ酸配列と相補体ポリペプチドとの第 2 の融合体を作り出す工程であって、前記相補体ポリペプチドが、配列番号 440 との 100 % 未満でかつ 30 % を超える配列同一性を有し、前記相補体ポリペプチドが配列番号 2 のペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

(c) 前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間で見込まれる安定な相互作用を生じさせる条件に、前記内部融合体と、第 2 の融合体と、セレンテラジン基質とを置く工程と、

(d) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの検出が、前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間での安定な相互作用を示す、工程と、

(e) 前記内部融合体と第 2 の融合体とセレンテラジン基質とに潜在的阻害剤を加える工程と、

(f) 存在する場合には、放射された生物発光シグナルを検出する工程と、

(g) 工程 (d) と (f) の前記生物発光シグナルを比較する工程であって、工程 (d) から工程 (f) への生物発光シグナルの低下が、前記潜在的阻害剤による前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間での前記相互作用の阻害を示す、工程と、

を含む、前記検出方法。

[61] 工程 (a) 及び (b) が、細胞内で前記内部融合体と第 2 の融合体とを発現させる工程を含む、前記 [60] に記載の方法。

[62] 前記潜在的阻害剤が、ペプチド又は小分子である、前記 [60] に記載の方法。

[62] 第 1 のアミノ酸配列と第 2 のアミノ酸配列との間での相互作用の、潜在的阻害剤による変化の検出方法であって、

(a) 内部タグが第 1 のアミノ酸配列の N 末端にも C 末端にもないように、内部タグを第 1 のアミノ酸配列に挿入することによって内部融合体を作り出す工程であって、前記内部タグが、配列番号 440 との 100 % 未満でかつ 30 % を超える配列同一性を有し、前記内部タグが配列番号 2 のペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

(b) 第 2 のアミノ酸配列と相補体ポリペプチドとの第 2 の融合体を作り出す工程であって、前記相補体ポリペプチドが、配列番号 2 との 100 % 未満でかつ 30 % を超える配列同一性を有し、前記相補体ポリペプチドが配列番号 440 のペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、工程と、

(c) 前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間で見込まれる安定な相互作用を生じさせる条件に、前記内部融合体と、第 2 の融合体と、セレンテラジン基質とを置く工程と、

(d) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの検出が、前記第 1 のアミノ酸配列と前記第 2 のアミノ酸配列との間での安定な相互作用を示す、工程と、

(e) 前記内部融合体と第 2 の融合体とセレンテラジン基質とに潜在的阻害剤を加える工

10

20

30

40

50

程と、

(f) 存在する場合には、放射された生物発光シグナルを検出する工程と、  
(g) 工程(d)と(f)の前記生物発光シグナルを比較する工程であって、工程(d)から工程(f)への生物発光シグナルの低下が、前記潜在的阻害剤による前記第1のアミノ酸配列と前記第2のアミノ酸配列との間での前記相互作用の阻害を示す、工程と、を含む、前記検出方法。

[61] 工程(a)及び(b)が、細胞内で前記内部融合体と第2の融合体を発現させる工程を含む、前記[60]に記載の方法。

[62] 前記潜在的阻害剤が、ペプチド又は小分子である、前記[60]に記載の方法。

[63] 第1のアミノ酸配列の構造配座の決定方法であって、

(a) 内部タグが第1のアミノ酸配列のN末端にもC末端にもないように、内部タグを第1のアミノ酸配列に挿入することによって内部融合体を作り出す工程であって、前記内部タグが、配列番号2との100%未満でかつ30%を超える配列同一性を有し、前記内部タグが配列番号440のポリペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成され、前記第1のアミノ酸配列の第1の構造配座が、前記内部タグへのアクセスを妨害し、前記第1のアミノ酸配列の第2の構造配座が、前記内部タグへのアクセスを可能にする、工程と、

(b) 前記内部融合体と、(i) 配列番号440との100%未満でかつ30%を超える配列同一性を有する相補体ポリペプチド又は(ii) 第2のアミノ酸配列及び前記相補体ポリペプチドの第2の融合体のいずれかとをセレンテラジン基質の存在下に置く工程と、

(c) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの非存在が、前記第1のアミノ酸配列が前記第1の構造配座を採用していることを示し、前記生物発光シグナルの存在が、前記第1のアミノ酸配列が前記第2の構造配座を採用していることを示す、工程とを含む、前記決定方法。

[64] 工程(c)が、

(i) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの非存在が、前記第1のアミノ酸配列が前記第1の構造配座を採用していることを示す、工程と、

(ii) 前記第1のアミノ酸配列において配座変化を誘導する工程と、

(iii) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの存在が、前記第1のアミノ酸配列が前記第2の構造配座を採用していることを示す、工程とを含む、前記[63]に記載の方法。

[65] 配座変化を誘導する工程が、前記第1のアミノ酸配列の一部を切断するプロテアーゼを加えること、前記第1のアミノ酸配列に結合する作用剤の添加、及びアッセイ条件を変更することから選択される、前記[64]に記載の方法。

[66] 第1のアミノ酸配列の構造配座の決定方法であって、

(a) 内部タグが第1のアミノ酸配列のN末端にもC末端にもないように、内部タグを第1のアミノ酸配列に挿入することによって内部融合体を作り出す工程であって、前記内部タグが、配列番号440との100%未満でかつ30%を超える配列同一性を有し、前記内部タグが配列番号2のポリペプチドと接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成され、前記第1のアミノ酸配列の第1の構造配座が、前記内部タグへのアクセスを妨害し、前記第1のアミノ酸配列の第2の構造配座が、前記内部タグへのアクセスを可能にする、工程と、

(b) 前記内部融合体と、(i) 配列番号2との100%未満でかつ30%を超える配列同一性を有する相補体ペプチド又は(ii) 第2のアミノ酸配列及び前記相補体ペプチドの第2の融合体のいずれかとをセレンテラジン基質の存在下に置く工程と、

(c) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの非存在が、前記第1のアミノ酸配列が前記第1の構造配座を採用してい

10

20

30

40

50

ることを示し、前記生物発光シグナルの存在が、前記第1のアミノ酸配列が前記第2の構造配座を採用していることを示す、工程とを含む、前記決定方法。

〔67〕工程(c)が

(i) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの非存在が、前記第1のアミノ酸配列が前記第1の構造配座を採用していることを示す、工程と、

(ii) 前記第1のアミノ酸配列において配座変化を誘導する工程と、

(iii) 存在する場合には、放射された生物発光シグナルを検出する工程であって、前記生物発光シグナルの存在が、前記第1のアミノ酸配列が前記第2の構造配座を採用していることを示す、工程と

を含む、前記〔66〕に記載の方法。

〔68〕配座変化を誘導する工程が、前記第1のアミノ酸配列の一部を切断するプロテアーゼを加えること、前記第1のアミノ酸配列に結合する作用剤の添加、及びアッセイ条件を変更することから選択される、前記〔67〕に記載の方法。

〔69〕N末端断片と、C末端断片と、2つ以上の内部タグとを含むポリペプチドであって、前記内部タグが、対象とするタンパク質の中に挿入された配列番号2との100%未満でかつ30%を超える配列同一性を有するアミノ酸配列を含み、前記内部タグの1つ以上が配列番号440のポリペプチドに接触すると、セレンテラジン基質の存在下で、検出可能な生物発光シグナルが生成される、前記ポリペプチド。

〔70〕前記2つ以上の内部タグが、2つの内部タグである、前記〔69〕に記載のポリペプチド。

〔71〕前記2つ以上の内部タグが、互いに直接接続されている、前記〔69〕に記載のポリペプチド。

〔72〕前記2つ以上の内部タグが、1つ以上のリンカーによって分離されている、前記〔69〕に記載のポリペプチド。

〔73〕前記2つ以上の内部タグが、対象とするタンパク質又はポリペプチド内の単一の位置に挿入されている、前記〔69〕に記載のポリペプチド。

〔74〕前記2つ以上の内部タグが、対象とするタンパク質又はポリペプチド内の2つ以上の位置に挿入されている、前記〔69〕に記載のポリペプチド。

〔75〕前記2つ以上の内部タグが、同一のアミノ酸配列を含む、前記〔69〕に記載のポリペプチド。

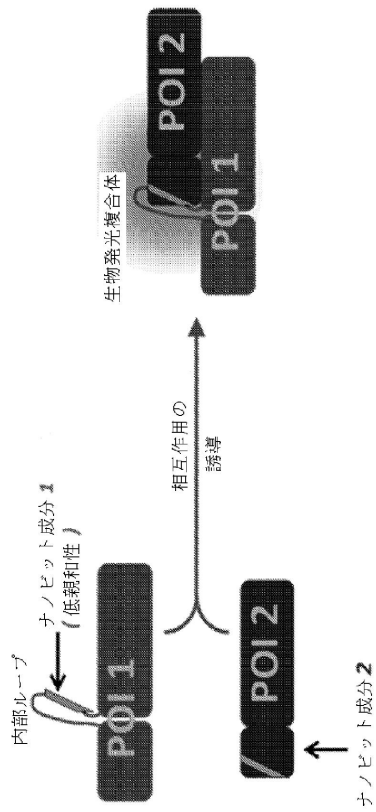
〔76〕2つ以上のまたは2つ以上の内部タグが、同一ではないアミノ酸配列を含む、前記〔69〕に記載のポリペプチド。

10

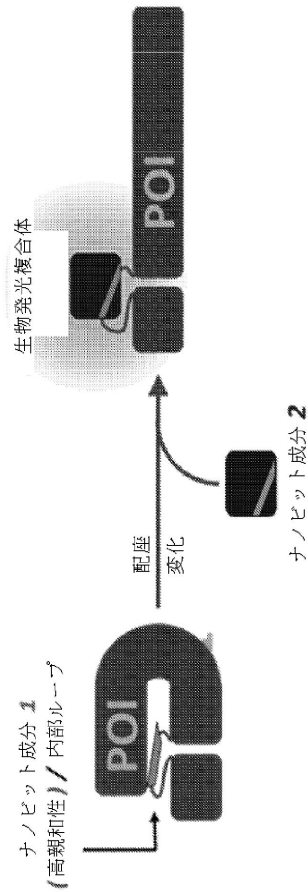
20

30

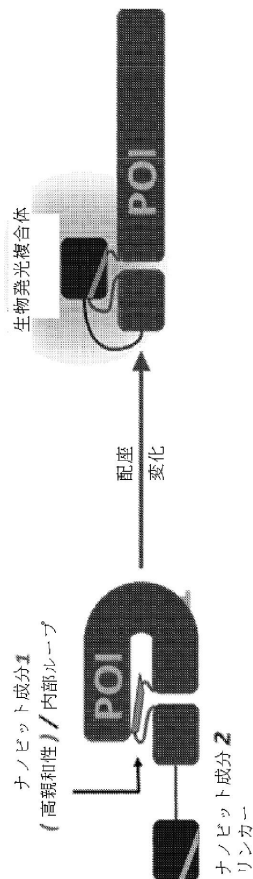
【図 1】



【図 2】

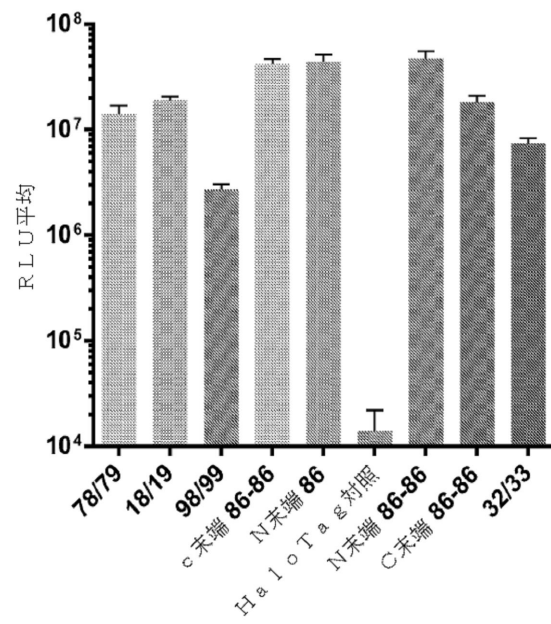


【図 3】



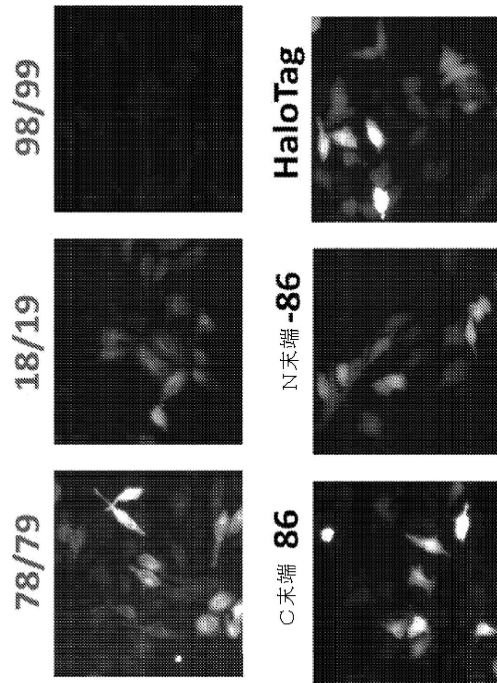
【図 4】

内部ペプチド 86 + F r b - 1 1 S の  
総発光 (G l o M a x)

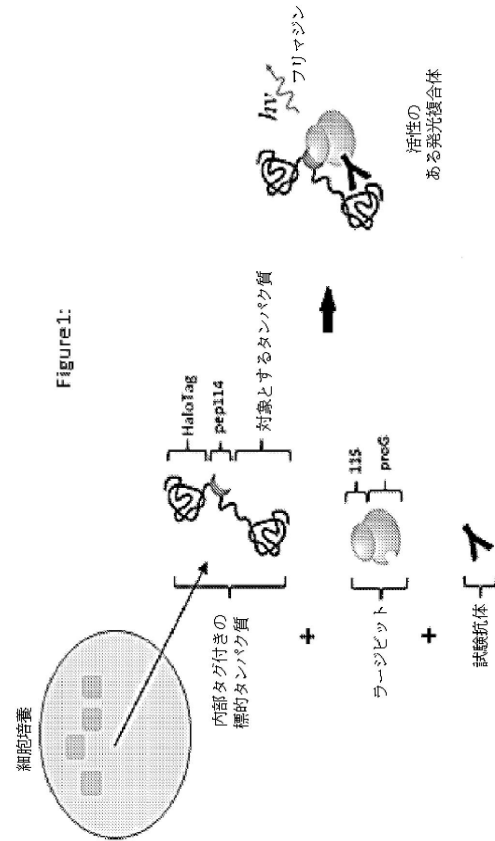


【図 5】

TMR-HaloTagリガンドで標識したHeLa細胞

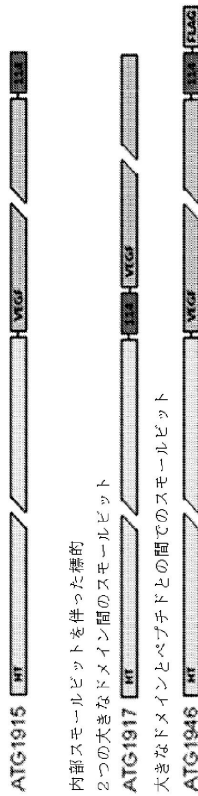


【図 6】

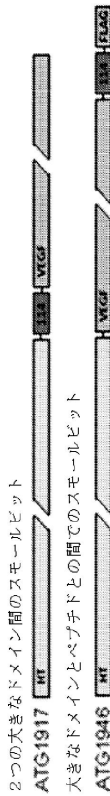


【図 7】

末端スモールビットを伴った標的



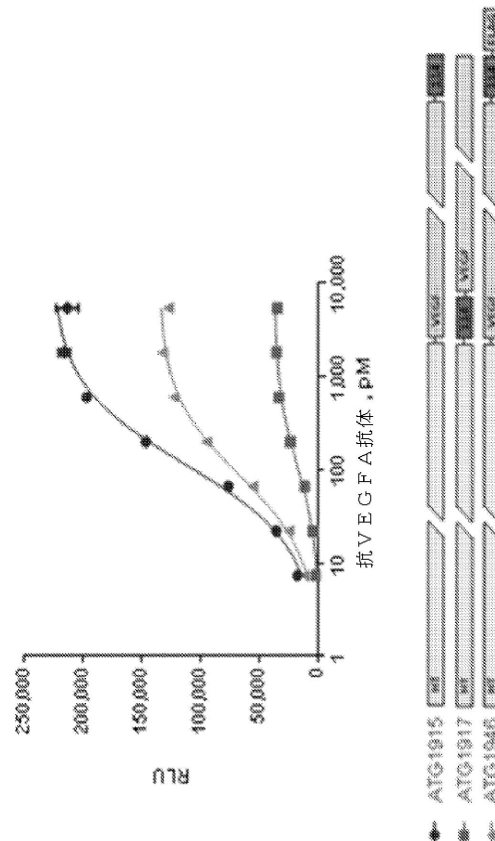
内部スモールビットを伴った標的



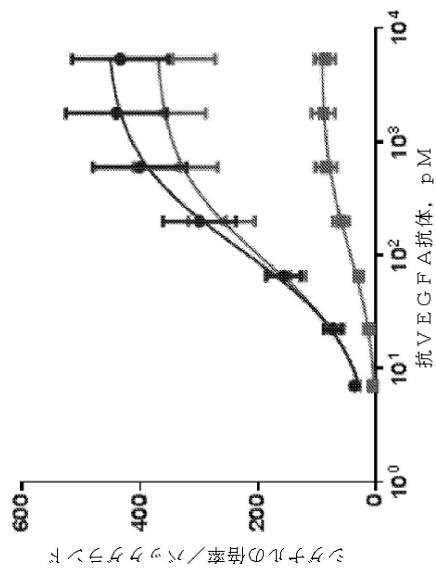
大きなドメインとペプチドとの間のスモールビット



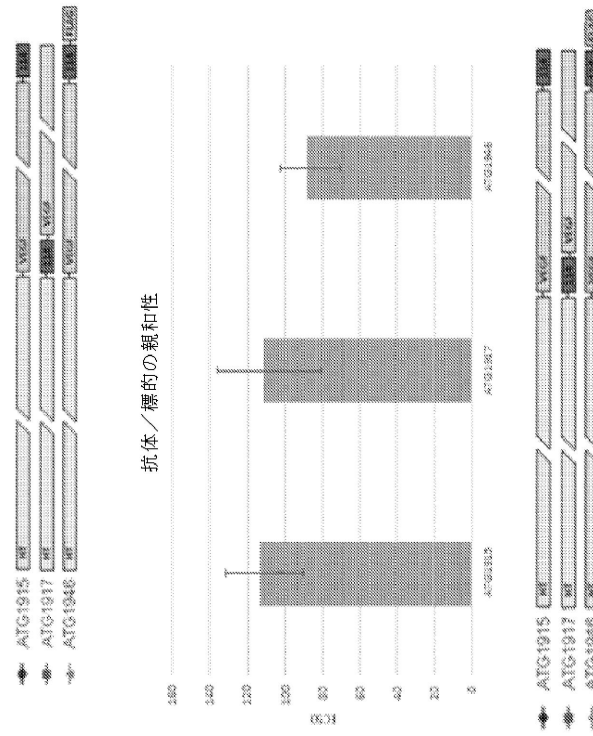
【図 8】



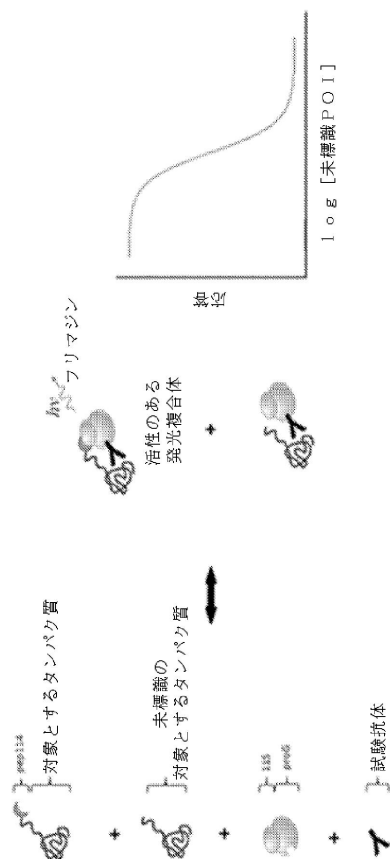
【図 9】



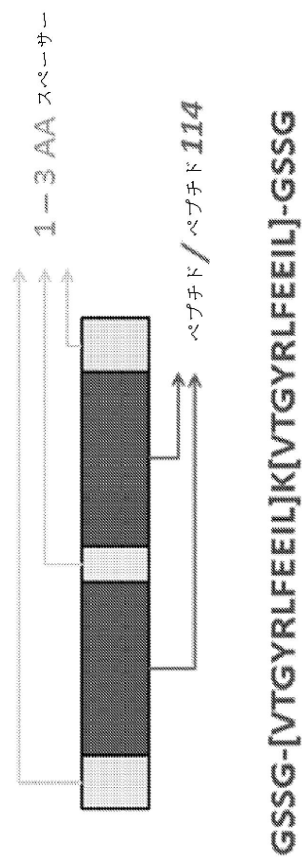
【図 10】



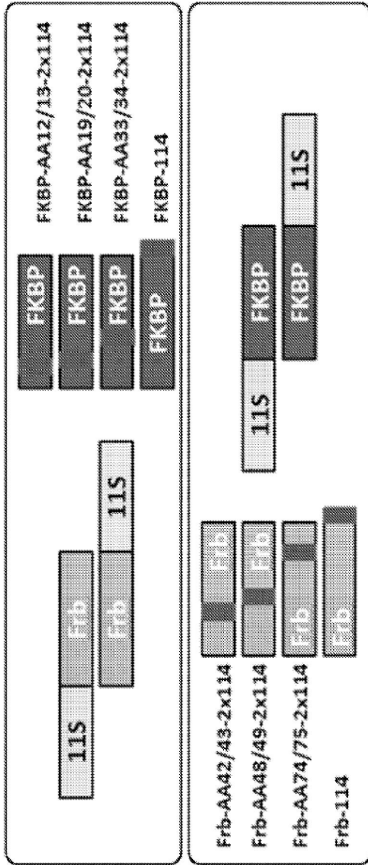
【図 11】



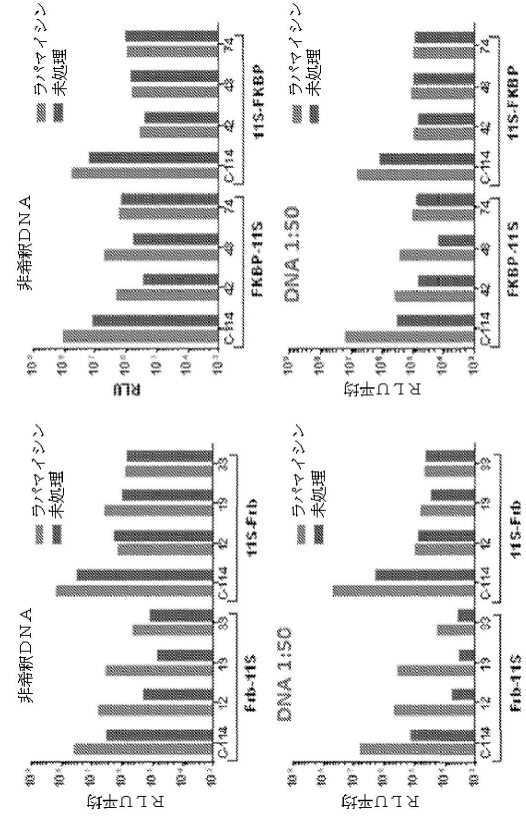
【図 12】



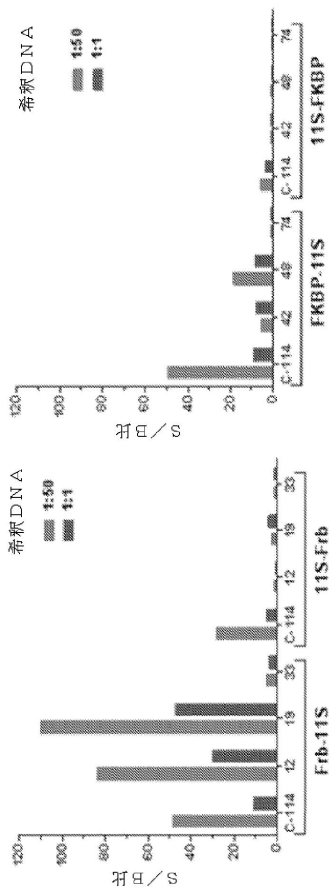
【 図 1 3 】



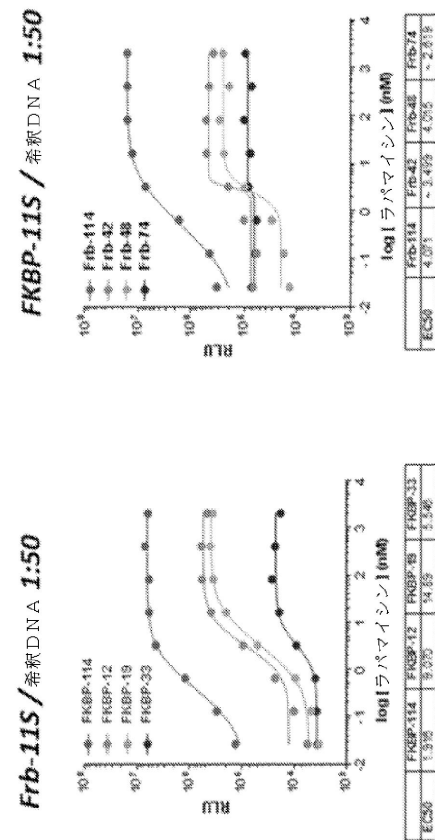
【 図 1 4 】



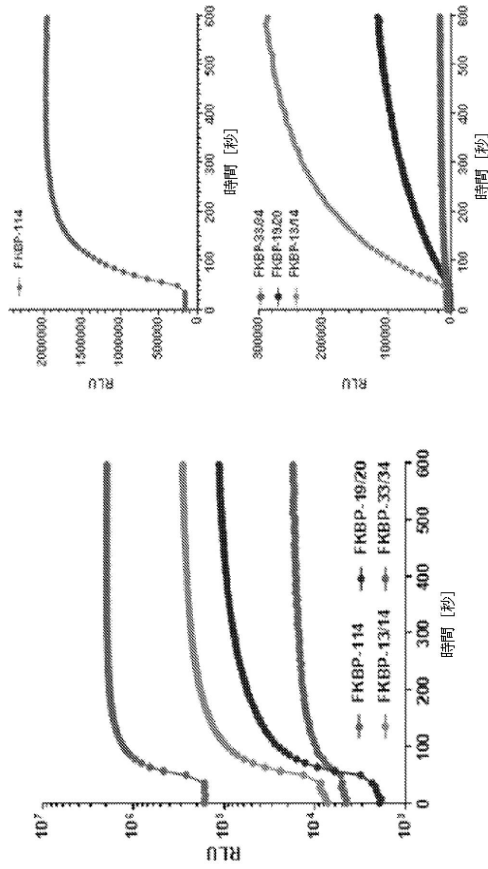
【 図 1 5 】



【 図 1 6 】



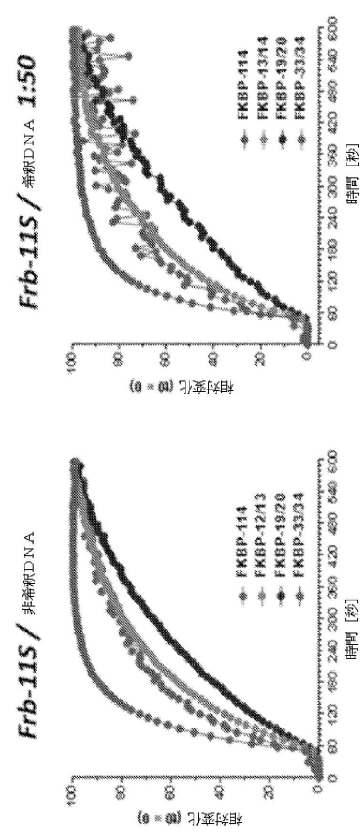
【図 17】



【配列表】

0006814132000001.app

【図 18】



## フロントページの続き

(51)Int.Cl. F I  
C 1 2 N 9/02 (2006.01) C 1 2 N 9/02

(74)代理人 100111796

弁理士 服部 博信

(72)発明者 ディクソン アンドリュー エス

アメリカ合衆国 ウィスコンシン州 5 3 7 1 1 マディソン ウッズ ホロー ロード 2 8 0  
0

(72)発明者 エンセル ランス ピー

アメリカ合衆国 ウィスコンシン州 5 3 7 1 1 マディソン ウッズ ホロー ロード 2 8 0  
0

(72)発明者 マヒレイド トーマス

アメリカ合衆国 ウィスコンシン州 5 3 7 1 1 マディソン ウッズ ホロー ロード 2 8 0  
0

(72)発明者 シュウィン マリー

アメリカ合衆国 ウィスコンシン州 5 3 7 1 1 マディソン ウッズ ホロー ロード 2 8 0  
0

(72)発明者 ウッド キース

アメリカ合衆国 ウィスコンシン州 5 3 7 1 1 マディソン ウッズ ホロー ロード 2 8 0  
0

(72)発明者 ウッド モニカ

アメリカ合衆国 ウィスコンシン州 5 3 7 1 1 マディソン ウッズ ホロー ロード 2 8 0  
0

(72)発明者 ツィーママン クリス

アメリカ合衆国 ウィスコンシン州 5 3 7 1 1 マディソン ウッズ ホロー ロード 2 8 0  
0

審査官 佐久 敬

(56)参考文献 特表2014-500011(JP,A)

国際公開第2007/041251(WO,A1)

特表2016-500263(JP,A)

特表2016-519065(JP,A)

(58)調査した分野(Int.Cl., DB名)

C 1 2 N

C 0 7 K

C 1 2 Q

C A p l u s / R E G I S T R Y ( S T N )

U n i P r o t / G e n e S e q