



(86) Date de dépôt PCT/PCT Filing Date: 2011/05/25
(87) Date publication PCT/PCT Publication Date: 2011/12/01
(45) Date de délivrance/Issue Date: 2018/10/23
(85) Entrée phase nationale/National Entry: 2012/11/08
(86) N° demande PCT/PCT Application No.: EP 2011/058591
(87) N° publication PCT/PCT Publication No.: 2011/147890
(30) Priorité/Priority: 2010/05/25 (GB1008682.5)

(51) Cl.Int./Int.Cl. *C07K 14/47* (2006.01),
C07K 5/10 (2006.01)
(72) Inventeurs/Inventors:
PARDON, ELS, BE;
STEYAERT, JAN, BE;
WYNS, LODE, BE
(73) Propriétaires/Owners:
VIB VZW, BE;
VRIJE UNIVERSITEIT BRUSSEL, BE
(74) Agent: FETHERSTONHAUGH & CO.

(54) Titre : MARQUEUR EPITOPIQUE POUR DES APPLICATIONS FONDEES SUR L'AFFINITE
(54) Title: EPITOPE TAG FOR AFFINITY-BASED APPLICATIONS

(57) **Abrégé/Abstract:**

The present invention relates to an epitope tag useful in affinity-based applications. The invention further includes fusion proteins, methods for preparing fusion proteins, nucleic acid molecules encoding these fusion proteins and recombinant host cells which contain these nucleic acid molecules. The invention also relates to nanobodies and other affinity ligands specifically recognizing said epitope tag, and uses thereof in affinity-based applications.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 December 2011 (01.12.2011)

(10) International Publication Number
WO 2011/147890 A1

(51) International Patent Classification:
C07K 14/47 (2006.01) *C07K 5/10* (2006.01)

(21) International Application Number:
PCT/EP2011/058591

(22) International Filing Date:
25 May 2011 (25.05.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1008682.5 25 May 2010 (25.05.2010) GB

(71) Applicants (for all designated States except US): **VIB VZW** [BE/BE]; Rijvisschestraat 120, B-9052 Gent (BE). **VRIJE UNIVERSITEIT BRUSSEL** [BE/BE]; Pleinlaan 2, B-1050 Brussel (BE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **PARDON, Els** [BE/BE]; Staatsbaan 85, B-3210 Lubbeek (BE). **STEYAERT, Jan** [BE/BE]; Beukenbosstraat 108, B-1652 Beersel (BE). **WYNS, Lode** [BE/BE]; Otto Veniusstraat 11, B-2000 Antwerpen (BE).

(74) Common Representative: **VIB VZW**; Rijvisschestraat 120, B-9052 Gent (BE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

— with sequence listing part of description (Rule 5.2(a))

(54) Title: EPITOPE TAG FOR AFFINITY-BASED APPLICATIONS

(57) Abstract: The present invention relates to an epitope tag useful in affinity-based applications. The invention further includes fusion proteins, methods for preparing fusion proteins, nucleic acid molecules encoding these fusion proteins and recombinant host cells which contain these nucleic acid molecules. The invention also relates to nanobodies and other affinity ligands specifically recognizing said epitope tag, and uses thereof in affinity-based applications.



WO 2011/147890 A1

EPITOPE TAG FOR AFFINITY-BASED APPLICATIONS

FIELD OF THE INVENTION

The present invention relates to an epitope tag useful in affinity-based applications. The invention further includes fusion proteins, methods for preparing fusion proteins, nucleic acid molecules
5 encoding these fusion proteins and recombinant host cells which contain these nucleic acid molecules. The invention also relates to nanobodies and other affinity ligands specifically recognizing said epitope tag, and uses thereof in affinity-based applications.

BACKGROUND OF THE INVENTION

Epitope tagging is a technique in which a known epitope tag (typically 6 to 30 amino acids) is fused to a
10 recombinant protein by placing sequence encoding the epitope within the same open reading frame of the protein by means of genetic engineering. By choosing an epitope tag for which an antibody is available, the technique makes it possible to detect tagged proteins for which no antibody is available. By selection of the appropriate epitope tag and antibody pair, it is possible to find a combination with properties that are suitable for the desired experimental application, such as Western blot analysis,
15 immunoprecipitation, immunochemistry, affinity purification, amongst others.

The first commercially available tags were originally designed for protein purification. Examples of these early commercial products include FLAG, 6 x His, and glutathione-S-transferase (GST) systems. The FLAG tagging system in its original version included the anti-FLAG M1 monoclonal antibody with calcium-dependent binding. FLAG-tagged proteins can be eluted from the M1 antibody with EDTA
20 (Hopp et al. 1988). Likewise, the 6xHis tag is used for purification of recombinant proteins by means of metal chelate chromatography (Hochuli et al. 1988). Similarly, GST-tagged proteins can be purified using glutathione agarose (Smith et al. 1988). In addition to the commercial tags, the development of other tags such as HA (Field et al. 1988) and c-myc (Evan et al. 1985) were reported. As research progressed and recombinant DNA technology evolved, the utility of epitope tags in the study of protein
25 interaction was recognized. For example, the anti-FLAG M2 monoclonal antibody was made available commercially (Brizzard et al. 1994) as were monoclonal antibodies for the 6xHis tag (Kaufmann et al. 2002), HA tag, and c-myc tag. Today, there are numerous types of tags with different features suited to diverse applications (Jarvik and Telmer 1998).

In addition to the traditional antibody and epitope combinations, other types of epitope tags have
30 been discovered. An early example is the protein A tag, which binds to IgG (Uhlén et al. 1983). Other examples include those based on interaction with streptavidin (Schmidt and Skerra 2007) and biotin

(Tucker and Grisshammer 1996), maltose-binding peptide (MBP tag) and maltose (di Guan et al. 1988), chitin-binding domain (CBD) and chitin (Chong et al. 1997), and the calmodulin-binding peptide that binds to calmodulin (Rigaut et al. 1999). Another type of tag is the S-peptide tag that binds to the S-protein derived from pancreatic RNase A (Hackbarth et al. 2004).

- 5 Further advancements in the field are the use of multiple epitope tags to increase signal strength and signal-to-noise ratio, e.g. multiple copies of FLAG, His, c-myc and HA tags, or the use of tandem copies of different tags. For example, in tandem affinity purification (TAP), used for the study of protein networks, an IgG-binding domain and calmodulin-binding peptide with an intervening tobacco etch virus (TEV) protease cleavage site are used (Rigaut et al. 1999). Recently, a shorter TAP tag has been
10 described, the SF-TAP tag (Gloeckner et al. 2007), in which two tandem copies of the Strep-tag II were combined with the FLAG tag.

While the use of epitope-tagging facilitates the study and characterization of newly discovered proteins, the technique does have limitations. The insertion of (an) epitope tag(s) can alter protein function, especially when a large tag is used (e.g. GST and MBP tag). Notably, the relatively short
15 epitope tags, such as FLAG, rarely affect the properties of the heterologous protein of interest and are very specific for their respective primary antibodies. However, antibody affinity chromatography often involves low or high pH elution which can irreversibly affect the properties of the fusion protein. It also employs resin which has limited reusability. Poly-Histidine, typically in the form of a hexa- His tag, appended to the N or C-terminus of a recombinant polypeptide, has been widely used to purify
20 recombinant proteins using immobilised metal ion chromatography (IMAC). However, only moderate purity from *Escherichia coli* extracts and relatively poor purification from yeast, *Drosophila* and HeLa extracts are retrieved using a (His)₆ tag (Lichty et al. 2005). Also, aspecific binding of contaminating proteins (e.g. *E. coli* SlyD, a prolyl isomerase) is often observed in IMAC preparations of heterologously expressed (His)₆-tagged proteins, due to histidine clusters or intrinsic metal binding sites on these
25 proteins. Further, few of the tags that have been described in literature have been tested in a high-throughput context. Peptide epitopes like the FLAG-tag, the calmodulin-binding peptide, the Strep-tag or Streptag II and the biotin acceptor peptide all exhibit a high degree of specificity for their cognate binding partners. However, the resins (immobilized proteins) that they interact with tend to be expensive, are easily fouled and have relatively low binding capacities, making them less than ideal for
30 high-throughput applications. Several other advantages and limitations of epitope tagging are listed in Jarvik and Telmer (1998), Kimple and Sondek (2004), Waugh (2005), Brizzard (2008).

Given that epitope tagging has the potential to make major contributions to the emerging fields of functional genomics and proteomics, there remains a need for alternative epitope tags. Indeed, future

applications in these fields will increasingly depend on the separate tagging of individual proteins, requiring multiple epitope tags that can be used in combination.

SUMMARY OF THE INVENTION

Proteins do not naturally lend themselves to high-throughput analysis because of their diverse physiochemical properties. Consequently, epitope tags have become indispensable tools for structural and functional proteomics initiatives. High-throughput protein production and analysis must, almost by definition, entail a generic strategy for purification. This is obviously a key attribute of epitope tags.

The small 4 amino acid tag of the present invention provide a good alternative for the His tag and other peptide tags that can be engineered into recombinant proteins. In particular, the C-terminal EPEA-tag (SEQ ID NO: 1) is short and is specifically recognized by a small single domain antibody fragment that can be immobilized to produce effective affinity resins to be used in the purification of tagged proteins. As exemplified further, the EPEA-tag (SEQ ID NO: 1) can also be used for the visualisation/identification of tagged proteins (ELISA, Western blot), for the copurification of protein complexes (Tap-tag applications), for (in vivo) imaging, for localizing the protein in the cell, for identifying interactions with other proteins/determining subunit structure, for determining the function of the protein and how it responds to changes in cell conditions, for determining what happens to the protein during the lifetime of the cell, for characterizing new proteins derived from cloned genes, and many others.

It is noted that WO 2006/045037 which is prior art, disclosed certain fragments of alphasynuclein protein.

In a first aspect, the epitope tag of the present invention consists of a 4 amino acid sequence, said sequence defined by EPEA (SEQ ID NO: 1), or by an amino acid sequence which has at least 75% sequence identity with SEQ ID NO: 1, or by an amino acid sequence wherein at least 50% or 75% of the amino acids have been substituted by a conservative amino acid.

In one embodiment, the present invention relates to a non-naturally occurring fusion protein comprising the epitope tag of the present invention and a target protein. Herein, the epitope tag is directly coupled to the C-terminal end of the target protein or via a linker.

10 In a further embodiment, the invention relates to an isolated nucleic acid sequence encoding the above described epitope tag or fusion protein. A vector comprising the above nucleic acid sequence and a host cell comprising said vector or said nucleic acid sequence forms also part of the present invention.

15 In a further aspect, the present invention relates to an affinity ligand specifically binding to the above described epitope tag. In particular, the affinity ligand can be an antibody or an antibody fragment or a nanobody. In specific embodiments, the invention also relates to an isolated nucleic acid sequence

encoding the herein described affinity ligands, as well as a vector comprising such a nucleic acid sequence and a host cell comprising said vector or said nucleic acid sequence.

In a further aspect, the present invention relates to a complex of (i) an affinity ligand as described herein, and (ii) a non-naturally occurring fusion protein as described herein, and (iii) optionally a ligand
5 bound to the target protein comprised in said non-naturally occurring fusion protein. The complex may be crystalline.

In another aspect, the invention relates to the use of the above epitope tag and/or the affinity ligand and/or the complex for affinity-based applications, such as affinity chromatography, affinity purification, immunoprecipitation, in-vivo imaging, tandem affinity purification, protein detection,
10 immunochemistry, surface-display, FRET-type applications, co-crystallization. Said protein detection can be performed in an ELISA detection assay or Western blot analysis. Said affinity purification can be performed on a column or on beads.

Also part of the invention is a method for purifying a non-naturally occurring fusion protein comprising the steps of

- 15 a) applying a solution containing a non-naturally occurring fusion protein as described herein to a solid support possessing an immobilized affinity ligand as described herein, and
 b) forming a complex between said immobilized affinity ligand and said fusion protein, and
 c) removing weakly bound or unbound molecules, and
 d) eluting the bound molecules.

20 The invention further also relates to a method for immobilizing and/or capturing a non-naturally occurring fusion protein on a solid support comprising the steps of:

- a) applying a solution containing a non-naturally occurring fusion protein to a solid support possessing an immobilized affinity ligand as described herein, and
 b) forming a complex between said immobilized affinity ligand and said fusion protein, and
25 c) removing weakly bound or unbound molecules.

Also part of the invention is a method for detecting a non-naturally occurring fusion protein in a sample by using the affinity ligand of the present invention.

In still another aspect, the invention provides a kit comprising an epitope tag as described herein, and/or a non-naturally occurring fusion protein as described herein, and/or a nucleic acid sequence as
30 described herein, and/or a vector as described herein, and/or an affinity ligand as described herein and/or a complex as described herein. The kit may be used for affinity-based applications, such as

affinity chromatography, affinity purification, immunoprecipitation, in-vivo imaging, tandem affinity purification, protein detection, immunochemistry, surface-display, FRET-type applications, co-crystallisation.

The invention as claimed relates to:

- 5 - a fusion protein comprising a target protein linked to a heterologous C-terminal epitope tag, wherein said C-terminal epitope tag is defined by SEQ ID NO: 1, and wherein said SEQ ID NO: 1 represents the last four C-terminal amino acid residues of said fusion protein, provided the target protein does not comprise fragments of alphasynuclein;
- 10 - a complex of (i) an antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1; and (ii) the fusion protein as described herein; and (iii) optionally a ligand bound to the target protein comprised in said fusion protein;
- a solid surface comprising an antibody fragment or single-domain antibody
15 specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1;
- use of an antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1 and/or the complex as described herein for affinity chromatography, affinity purification, immunoprecipitation, in-vivo imaging, tandem affinity purification, protein detection, immunochemistry,
20 surface-display, FRET-type applications and/or co-crystallisation;
- a method for purifying a fusion protein comprising the steps of a) applying a solution containing the fusion protein as described herein to a solid support possessing an immobilized antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1, and b) forming a complex between
25 said immobilized antibody fragment or single-domain antibody and said fusion protein, and c) removing weakly bound or unbound molecules, and d) eluting the bound molecules;

81584211

- a method for immobilizing and/or capturing a fusion protein on a solid support comprising the steps of: a) applying a solution containing the fusion protein as described herein to a solid support possessing an immobilized antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1, and b) forming a complex between said immobilized antibody fragment or single-domain antibody and said fusion protein, and c) removing weakly bound or unbound molecules;
- a method for detecting the fusion protein as described herein in a sample or composition comprising the step of using an antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1;
- a kit comprising the fusion protein as described herein, and/or the nucleic acid as described herein, and/or the vector as described herein, and/or an antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1, said antibody fragment or single-domain antibody being pre-immobilized to a non-soluble matrix or solid support, and/or the complex as described herein, and instructions for using said kit; and
- the fusion protein as described herein or the complex as described herein, for use in affinity chromatography, affinity purification, immunoprecipitation, in-vivo imaging, tandem affinity purification, protein detection, immunochemistry, surface-display, FRET-type applications and/or co-crystallisation.

DESCRIPTION OF THE DRAWINGS

Fig. 1: Amino acid sequences of the different nanobodies that specifically bind to α -synuclein

81584211

Fig. 2: Analysis of the interactions between α -synuclein and a series of α -synuclein-specific nanobodies by Chemical Shift Mapping using NMR spectroscopy. The top panel represents the HSQC spectra of N^{15} labeled α -synuclein obtained in the absence (red) in the presence (blue) of 1 equivalent of Nbsyn2, Nb85 or Nb87. The bottom panel represents the relative intensity changes upon binding of the different nanobodies. Relative intensities (I/I_0) were calculated for each amino acid from the proton-nitrogen cross-peaks of free α -synuclein (I_0) and nanobody bound α -synuclein (I) and plotted against the residue numbers.

Fig. 3. IMAC purification of EPEA-tagged GFP in complex with nanobodies Nb85, Nb87, Nb88 or NbSyn2. Purified complexes were analysed by SDS-PAGE analysis. Panel (A) lane 1 to 4: Nb85 in complex with CA1508, CA1509, CA1510 or CA1511; lane 5 to 8: Nb87 in complex with CA1508, CA1509, CA1510 or CA1511; lane 9: protein marker. Panel (B) lane 1: protein marker, lane 2 to 5: Nb88 in complex with CA1508, CA1509, CA1510 or CA1511; lane 6 to 9: NbSyn2 in complex with CA1508, CA1509, CA1510 or CA1511;

Fig. 4. Affinity purification of a tagged GFP (CA1508) using NbSyn2 coupled to a resin: Affinity purified tagged GFP was analysed by SDS-PAGE analysis, lane 1: protein marker; lane 2: wash with 50mM Tris HCl pH7.5 without salt; lane 3: elution with 50mM Tris HCl containing 1M NaCl pH7.5; lane 4: elution with 50mM TrisHCl containing 1M NaCl pH7.5; lane 5: elution with 50mM Tris HCl pH7.5 containing 0.5M adipic acid

Fig. 5. Affinity purification of a tagged GFP (CA1508) using NbSyn2 coupled to a resin: Affinity purified tagged GFP was analysed by SDS-PAGE analysis, lane 1: protein marker; lane 2: flow through; lane 3 and 4: wash with 50mM TrisHCl pH7.5; lane 5: elution containing 750 μ M peptide GYQDYEPEA; lane 6: elution with 50mM Tris HCl containing 2M NaCl pH7.5; lane 7: protein marker

81584211

Fig. 6. Structure of the single domain camelid antibody fragment Nbsyn2 in complex with the C-terminal peptide (GYQDYEPEA - SEQ ID NO: 2) of α synuclein. Panel (A) gives a schematic overview of the complex. The framework is represented in grey while the CDR1, CDR2 and CDR3 loops are in green,

red and yellow, respectively. The peptide is colored purple. Panel (B) depicts the molecular interactions between the nanobody and the peptide.

Fig. 7. Detailed view of the side-chain interactions between the residues of the peptide with NbSyn2. Residues of the peptide are represented as sticks and residues of NbSyn2 are drawn as lines. Atoms are color-coded: C: green; N: pink; O: yellow. A $2F_o - F_c$ map contoured at 1σ , is represented as a light blue mesh. The interactions for each side chain are represented in different panels (from left to right, top panels: Asp 135, Tyr 136, Glu 137; bottom panels: Pro 138, Glu 139, Ala 140). Note that the side-chain of Tyr 136 interacts with residues of a symmetry related NbSyn2 molecule.

Fig. 8. Affinity purification of a tagged nanobody (Nb458) using NbSyn2 coupled to a resin: Affinity purified tagged versions of Nb458 were analysed by SDS-PAGE analysis: lane 1: Nb458-GYQDYEPEA; lane 2: protein marker, lane 3: Flow trough; lane 4: Nb458-gaaEPEA ; lane 5: Nb458-gaaEPEA.

Fig. 9 Detection of EPEA-tagged nanobodies in ELISA using a selfmade NbSyn2 alkaline phosphatase conjugate. Nb458-GYQDYEPEA (CA2944), Nb458-gaaEPEA (CA2945), and two His-tagged nanobodies (Nb458 – SEQ ID NO: 66 and Nb196 – SEQ ID NO: 65) were coated in different wells of an ELISA plate. Selfmade NbSyn2 alkaline phosphatase conjugate was added. The ELISA was developed using p-Nitrophenyl phosphate as a substrate.

Fig. 10. Detection of EPEA-tagged proteins in Western Blot. GYQDYEPEA tagged GFP (lanes 1 & 4), α synuclein (lane 3) and a protein marker (lane 2) were separated on SDS-PAGE and transferred to a nitro-cellulose sheet. The western blot was developed using His-tagged NbSyn2 as the primary antibody. Nb190 is an His-tagged nanobody that recognizes an unrelated antigen.

Fig. 11. Detection of EPEA-tagged proteins in Western Blot. Nb458-gaaEPAA (lanes 1), Nb458-GYQDYEPEA (lane 2) and a protein marker (lane 3) were separated on SDS-PAGE and transferred to a nitro-cellulose sheet. The western blot was developed using a selfmade NbSyn2 alkaline phosphatase conjugate in combination with NBT/BCIP.

Fig. 12. High throughput screening of a collection of VHH's carrying the EPEA tag using ELISA. Antigen was coated to the wells of the rows marked + of an ELISA plate. 48 nanobodies, selected trough phage display were added to the coated wells and to control wells without antigen (marked -). Antigen specific nanobodies were detected using our selfmade NbSyn2 alkaline phosphatase conjugate and p-nitrophenyl phosphate as a substrate.

Fig. 13. Visualisation of the EPEA tagged HIV-Rev protein inside a HeLa cell with a NbSyn2-GFP fusion or NbSyn2-mKO1 . Upper panels: HeLa cell lines were transfected with NbSyn2-GFP separately and in

combination with HIV-Rev-GYQDYEPEA and HIV-Rev-EPEA, respectively. Lower panels: HeLa cell lines were transfected with NbSyn2-mKO1 separately and in combination with HIV-Rev-GYQDYEPEA and HIV-Rev-EPEA, respectively.

Fig. 14 Detection of EPEA-tagged nanobodies in ELISA using the biNbSyn2 alkaline phosphatase conjugate. Nb458-gaaEPEA (CA2945) was coated at different concentrations per well of an ELISA plate (0,1 µg or 0,01 µg/well). biNbSyn2 alkaline phosphatase conjugate or NbSyn2 alkaline phosphatase conjugate were added in different dilutions: 1/250, 1/500, 1/1000, 1/2000 and 1/4000.. The ELISA was developed using p-Nitrophenyl phosphate as a substrate. The OD405 value was measured (A) after 60min of incubation and (B) after an overnight incubation.

Fig. 15. Detection of a EPEA-tagged membrane protein CsgG in Western Blot. A protein marker (lane 2), clones 1 to 4 of the EPEA tagged CsgG (lanes 2 to 5) and a positive control (EPEA tagged Nanobody)(lane 6) were separated on SDS-PAGE and transferred to a nitro-cellulose sheet. The western blot was developed using the selfmade NbSyn2 alkaline phosphatase conjugate.

Fig. 16. Affinity purification of a tagged membrane protein (CsgG) using NbSyn2 coupled to a resin: Affinity purified material was analysed by SDS-PAGE analysis: lane 1: sample before purification, Nb458-GYQDYEPEA; lane 2: Flow through, lane 3: protein marker; lane 4: wash; lane 5: CsgGggEPEA.

DETAILED DESCRIPTION OF THE INVENTION

The present invention will be described with respect to particular embodiments and with reference to certain drawings but the invention is not limited thereto but only by the claims. Any reference signs in the claims shall not be construed as limiting the scope. The drawings described are only schematic and are non-limiting. In the drawings, the size of some of the elements may be exaggerated and not drawn on scale for illustrative purposes. Where the term "comprising" is used in the present description and claims, it does not exclude other elements or steps. Where an indefinite or definite article is used when referring to a singular noun e.g. "a" or "an", "the", this includes a plural of that noun unless something else is specifically stated. Furthermore, the terms first, second, third and the like in the description and in the claims, are used for distinguishing between similar elements and not necessarily for describing a sequential or chronological order. It is to be understood that the terms so used are interchangeable under appropriate circumstances and that the embodiments of the invention described herein are capable of operation in other sequences than described or illustrated herein.

Unless otherwise defined herein, scientific and technical terms and phrases used in connection with the present invention shall have the meanings that are commonly understood by those of ordinary skill in the art. Generally, nomenclatures used in connection with, and techniques of molecular and cellular biology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well known and commonly used in the art. The methods and techniques of the present invention are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. See, for example, Sambrook et al. Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2002).

In a first aspect of the invention, an epitope tag consisting of a 4 amino acid sequence has been identified, wherein said sequence is defined by the amino acid sequence EPEA (SEQ ID NO: 1) or by an amino acid sequence which has at least 25%, at least 50%, preferably at least 75% sequence identity with SEQ ID NO: 1, or by an amino acid sequence wherein at least 25%, at least 50% or at least 75% of the amino acids have been substituted by a conservative amino acid.

As used herein, the term “epitope tag”, or otherwise “affinity tag”, refers to a short amino acid sequence or peptide enabling a specific interaction with a protein or a ligand.

Conservative substitutions, as used herein, are those in which at least one residue in the amino acid sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with Table 1. Table 1 shows amino acids which can be substituted for an amino acid in a protein and which are typically regarded as conservative substitutions.

TABLE 1

Residue	Conservative Substitutions
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Gln	Asn
Cys	Ser
Glu	Asp
Gly	Pro

Residue	Conservative Substitutions
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln
Met	Leu, Ile
Phe	Met, Leu, Tyr
Ser	Thr, Gly
Thr	Ser, Val
Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

As used herein, the terms “identical” or percent “identity” in the context of two or more polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues that are the same (e.g., 75% identity over a specified region) when compared and aligned for maximum correspondence over a comparison window or designated region as measured using sequence comparison algorithms or by manual alignment and visual inspection. In particular, the identity exists over a region that is 4 amino acids in length.

In a second aspect, the epitope tag is linked to or coupled to a molecule, said combination being referred to herein as a “fusion molecule” or a “fusion construct”. Accordingly, the present invention relates to a fusion molecule comprising the epitope tag as described herein. Preferably, the affinity tag is linked to a protein of interest or otherwise “target protein”. The epitope tag can be linked directly or indirectly to said molecule. The tag can be linked to any site of the molecule, e.g. to or near the end or terminus of the molecule, to one or more internal sites, attached to a side chain, or to the amino-terminal amino acid (N-terminal) or to the carboxy-terminal amino acid (C-terminal). Also, more than one epitope tag can be linked to the molecule. In a preferred embodiment, the epitope tag is linked to the C-terminal end of a target protein by placing a DNA sequence encoding the epitope tag within the same open reading frame of the recombinant protein by means of genetic engineering, which is a well-known technique by the person skilled in the art.

The term “fusion molecule” or “fusion construct” as used herein generally denotes a composition that does not occur in nature, or otherwise a non-naturally occurring fusion molecule or fusion construct.

In the case of indirect linking, a suitable linker sequence is inserted between the desired molecule and the epitope tag. The epitope tag can then be removed chemically or enzymatically if a cleavage site is present in the linker sequence, using methods known in the art. Preferred linkers are peptides of 1 to 30 amino acids long and include, but are not limited to the peptides EEGEPK (Kjeldsen et al. in WO98/28429; SEQ ID NO: 13) or EEAEPK (Kjeldsen et al. in WO97/22706; SEQ ID NO: 14); the GGGGS (SEQ ID NO: 81) immunosilent linker; a protease cleavage site such as Factor Xa cleavage site having the sequence IEGR (SEQ ID NO: 15), the thrombin cleavage site having the sequence LVPR (SEQ ID NO: 16) or the enterokinase cleaving site having the sequence DDDDK (SEQ ID NO: 17). Other suitable linkers are carbohydrates, PEG based linkers, and other available in the art. Preferably, the amino acid linker sequence is short and does not interfere with the biological activity of the protein and the specific interaction with the antibody fragment. Non-limiting examples of suitable short linker sequences are as described in the Example section, e.g. GAA and GAGA (SEQ ID NO: 12), as well as in Table 2.

The term "molecule" as used herein refers to proteins, including antibodies and enzymes, peptides, nucleotides, lipids, carbohydrates, drugs and cofactors, or combinations thereof. The nature of the molecule is not critical to the invention. The molecule may be of varying length, size or molecular weight, and can have any activity known and/or desired by the skilled person.

In a particular embodiment, the molecule is a protein. More particular, the epitope tag of the invention is coupled to a protein, or a fragment thereof, said combination also being referred to as a fusion protein. In fact and as used herein, a "fusion protein" refers to a polypeptide which comprises the amalgamation of two amino acid sequences derived from heterogeneous sources, and as such can be referred to as a "non-naturally occurring fusion protein". The protein can have any activity known and/or desired by the skilled person, e.g. immunogenic activity, enzymatic activity, or binding activity. The protein can also be a structural protein. Specific proteins, including fragments thereof, which can be linked to the epitope tag of this invention, include for example hormones, enzymes, cytokines, intracellular signaling peptides, receptors, antibodies, growth factors, blood factors, vaccines, reporter proteins, synthetic peptides. Non-limiting examples of specific proteins are described in the Example section, e.g. fluorescent proteins, antibody fragments, viral regulatory proteins, membrane proteins, cell cycle proteins, amongst others. In a specific embodiment, the proteins are derived from bacteria, yeasts, viruses, mammals, plants, insects, and the like.

As used herein, the terms "polypeptide", "protein", "peptide" are used interchangeably and refer to a polymeric form of amino acids of any length, which can include coded and non-coded amino acids,

chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones.

As will be evident to one of ordinary skill in the art, various fragments of the herein described proteins may be combined. Said fragment(s) may be of varying length, although it is generally preferred that
5 these are at least 4 amino acids long, and up to the length of the entire protein.

The preparation of the fusion molecules of this invention can be carried out using standard recombinant DNA methods known in the art.

According to another embodiment, the present invention relates to an isolated nucleic acid fragment that encodes the epitope tag described herein, an isolated nucleic acid comprising said nucleic acid
10 fragment, an isolated nucleic acid encoding the fusion molecule of the invention, as well as to a composition comprising such a nucleic acid molecule. As used herein, the terms "nucleic acid", "polynucleotide", "polynucleic acid" are used interchangeably and refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or combinations and/or analogs thereof. The nucleic acid segments that encode the molecule and the epitope tag may be
15 contiguous, such that in the transcription and/or translation products of the coding segments, the segments are juxtaposed. In some embodiments, the coding sequences of the tag of the invention and a molecule may be separated by a linker encoding nucleic acid sequence, or by one or more sequences that are non-coding. Thus, the present invention encompasses nucleic acid molecules containing one or more intervening sequences (e.g. introns) that may be transcribed from a DNA molecule into an
20 RNA molecule and subsequently removed (e.g. by splicing) prior to translation of the RNA molecule into protein. Nucleic acid molecules of the invention may be synthesized in vitro, in vivo, or by the action of cell-free transcription. Preferably, a nucleotide sequence coding for the epitope tag is first synthesized and then linked to a nucleotide sequence coding for the desired molecule or protein. The thus-obtained hybrid gene can be incorporated into an expression or cloning vector using standard
25 methods. Vectors according to this aspect of the invention can be double-stranded or single-stranded and may be DNA, RNA, or DNA/RNA hybrid molecules, in any conformation including but not limited to linear, circular, coiled, supercoiled, torsional, nicked and the like. These vectors of the invention include but are not limited to plasmid vectors and viral vectors, such as a bacteriophage, baculo virus, retrovirus, lentivirus, adenovirus, vaccinia virus, semliki forest virus and adeno-associated virus vectors,
30 all of which are well-known and can be purchased from commercial sources. Any vector may be used to construct the fusion molecules used in the methods of the invention. In particular, vectors known in the art and those commercially available (and variants or derivatives thereof) may in accordance with

the invention be engineered to include one or more recombination sites for use in the methods of the invention. General classes of vectors of particular interest include prokaryotic and/or eukaryotic cloning vectors, expression vectors, fusion vectors, two-hybrid or reverse two- hybrid vectors, phage and ribosome display vectors, TAP vectors, shuttle vectors for use in different hosts, mutagenesis
5 vectors, transcription vectors, vectors for receiving large inserts and the like. Other vectors of interest include viral origin vectors (M 13 vectors, bacterial phage 8 vectors, adenovirus vectors, and retrovirus vectors), and high, low and adjustable copy number vectors. Most of the requisite methodology can be found in Ausubel et al. 2007.

Vector constructs prepared for introduction into a prokaryotic or eukaryotic host will typically
10 comprise a replication system recognized by the host, including the intended DNA fragment encoding the fusion molecule of the present invention, and will preferably also include transcription and translational initiation regulatory sequences operably linked to the molecule- encoding segment. Expression systems may include, for example, an origin of replication or autonomously replicating sequence (ARS) and expression control sequences, a promoter, an enhancer and necessary processing
15 information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, and mRNA stabilizing sequences. Signal peptides may also be included, where appropriate, from secreted polypeptides of the same or related species, which allow the protein to cross and/or lodge in cell membranes, or be secreted from the cell.

An appropriate promoter and other necessary vector sequences will be selected so as to be functional
20 in the host. Examples of workable combinations of cell lines and expression vectors are described in Sambrook et al. (1989), Ausubel et al. (Eds.) (2007), and Metzger et al. (1988). Many useful vectors for expression in bacteria, yeast, fungal, mammalian, insect, plant or other cells are well known in the art. In addition, the construct may be joined to an amplifiable gene (e.g., DHFR) so that multiple copies of the gene may be made. For appropriate enhancer and other expression control sequences, see also
25 Enhancers and Eukaryotic Gene Expression (1983) Cold Spring Harbor Press, N.Y. While such expression vectors may replicate autonomously, they may less preferably replicate by being inserted into the genome of the host cell.

Expression and cloning vectors will likely contain a selectable marker, that is, a gene encoding a protein necessary for the survival or growth of a host cell transformed with the vector. Although such a marker
30 gene may be carried on another polynucleotide sequence co-introduced into the host cell, it is most often contained on the cloning vector. Only those host cells into which the marker gene has been introduced will survive and/or grow under selective conditions. Typical selection genes encode

proteins that (a) confer resistance to antibiotics e.g., kanamycin, tetracycline, etc. or other toxic substances; (b) complement auxotrophic deficiencies; or (c) supply critical nutrients not available from complex media. The choice of the proper selectable marker will depend on the host cell; appropriate markers for different hosts are known in the art.

5 Recombinant host cells, in the present context, are those which have been genetically modified to contain an isolated DNA molecule of the instant invention. The DNA can be introduced by any means known to the art which are appropriate for the particular type of cell, including without limitation, transformation, lipofection, electroporation or viral mediated transduction. A DNA construct capable of enabling the expression of the fusion molecule of the invention can be easily prepared by the art-
10 known techniques such as cloning, hybridization screening and Polymerase Chain Reaction (PCR). Standard techniques for cloning, DNA isolation, amplification and purification, for enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like, and various separation techniques are those known and commonly employed by those skilled in the art. A number of standard techniques are described in Sambrook et al. (1989), Maniatis et al. (1982), Wu (ed.) (1993) and Ausubel
15 et al. (1992).

An "isolated polypeptide" or an "isolated polynucleotide", as used herein, refers to respectively an amino acid sequence or a polynucleotide sequence which is not naturally-occurring or no longer occurring in the natural environment wherein it was originally present.

In a further embodiment, the present invention encompasses host cells comprising one or more
20 nucleic acid molecules of the invention (e.g. a nucleic acid molecule encoding one or more fusion molecules of the invention). Representative host cells that may be used with the invention include, but are not limited to, bacterial cells, yeast cells, plant cells and animal cells. Bacterial host cells suitable for use with the invention include *Escherichia* spp. cells, *Bacillus* spp. cells, *Streptomyces* spp. cells, *Erwinia* spp. cells, *Klebsiella* spp. cells, *Serratia* spp. cells, *Pseudomonas* spp. cells, and *Salmonella* spp. cells.
25 Animal host cells suitable for use with the invention include insect cells (most particularly *Drosophila melanogaster* cells, *Spodoptera frugiperda* Sf9 and SfZ1 cells and *Trichoplusia* High-Five cells), nematode cells (particularly *C. elegans* cells), avian cells, amphibian cells (particularly *Xenopus laevis* cells), reptilian cells, and mammalian cells (most particularly derived from Chinese hamster (e.g. CHO), monkey (e.g. COS and Vero cells), baby hamster kidney (BHK), pig kidney (PK15), rabbit kidney 13 cells
30 (RK13), the human osteosarcoma cell line 143 B, the human cell line HeLa and human hepatoma cell lines like Hep G2). Yeast host cells suitable for use with the invention include species within *Saccharomyces*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia* (e.g. *Pichia pastoris*), *Hansenula* (e.g.

81584211

Hansenula polymorpha), Yarrowia, Schwaniomyces, Schizosaccharomyces, Zygosaccharomyces and the like. Saccharomyces cerevisiae, S. carlsbergensis and K. lactis are the most commonly used yeast hosts, and are convenient fungal hosts. The host cells may be provided in suspension or flask cultures, tissue cultures, organ cultures and the like. Alternatively the host cells may also be transgenic animals.

5 Methods for introducing the nucleic acid molecules and/or vectors of the invention into the host cells described herein, to produce host cells comprising one or more of the nucleic acid molecules and/or vectors of the invention, will be familiar to those of ordinary skill in the art. For instance, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells using well known techniques of infection, transduction, transfection, and transformation. The nucleic acid molecules and/or vectors of the invention may be introduced alone or in conjunction with other nucleic acid molecules and/or vectors. Alternatively, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells as a precipitate, such as a calcium phosphate precipitate, or in a complex with a lipid.

15 Electroporation also may be used to introduce the nucleic acid molecules and/or vectors of the invention into a host. Likewise, such molecules may be introduced into chemically competent cells such as E. coli. If the vector is a virus, it may be packaged in vitro or introduced into a packaging cell and the packaged virus may be transduced into cells. Hence, a wide variety of techniques suitable for introducing the nucleic acid molecules and/or vectors of the invention into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art. Such techniques are reviewed at length, for example, in Sambrook et al. (1989), Watson et al. (1992), and Winnacker (1987), which are illustrative of the many laboratory manuals that detail these techniques.

25 In a further aspect, the present invention relates to an affinity ligand specifically binding to the above described epitope tags. As used herein, the "affinity ligand" binds non-covalently to the tags of the present invention and can be any molecule and more particular a metal affinity ligand, an antibody, an antibody fragment, a nanobody, a small molecule or a synthetic affinity ligand. In a preferred embodiment, the affinity ligand is a nanobody specifically binding to the above described epitope tag. It should be clear that the specific binding of the affinity ligand to the epitope tag, when said epitope tag is directly or indirectly coupled to a target molecule, is particularly envisaged in the present invention.

30 The terms "specifically bind" and "specific binding" as used herein refer to the ability of an antibody, an antibody fragment or a nanobody to preferentially bind to a particular antigen that is present in a

homogeneous mixture of different antigens. In certain embodiments, a specific binding interaction will discriminate between desirable and undesirable antigens in a sample, in some embodiments more than about 10 to 100-fold or more (e.g., more than about 1000- or 10,000-fold). "Specificity" refers to the ability of an immunoglobulin or an immunoglobulin fragment, such as a nanobody, to bind preferentially to one antigenic target versus a different antigenic target and does not necessarily imply high affinity.

Antibodies specifically binding the herein described tag are also part of the invention. Antibodies may be polyclonal and/or monoclonal. As used herein, the term "antibody" (Ab) is meant to include whole antibodies, including single-chain whole antibodies, and antigen-binding fragments. In some embodiments, antigen-binding fragments may be mammalian antigen-binding antibody fragments that include, but are not limited to, Fab, Fab' and F(ab')₂, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a VL or VH domain, as well as nanobodies (see further herein). One or more of the epitope tags and/or fusion molecules of the invention may be used as immunogens to prepare polyclonal and/or monoclonal antibodies capable of binding the epitope tags and/or fusion molecules using techniques well known in the art (Harlow & Lane, 1988). In brief, antibodies are prepared by immunization of suitable subjects (e.g. mice, rats, rabbits, goats, etc.) with all or a part of the epitope tags and/or fusion molecules of the invention. If the epitope tag and/or fusion molecule, or a fragment thereof, is sufficiently immunogenic, it may be used to immunize the subject. If necessary or desired to increase immunogenicity, the epitope tag and/or fusion molecule, or fragment, may be conjugated to a suitable carrier molecule (e.g., BSA, KLH, and the like).

Thus, according to a specific embodiment, the present invention also encompasses a method for generating antibodies specifically binding to the above described epitope tags comprising the steps of:

- (i) immunizing an animal with an epitope tag according to the invention and/or with a (fusion) protein comprising the epitope tag according to the invention, and
- (ii) screening for antibodies specifically binding to said epitope tag.

Monoclonal antibodies can be prepared from the immune cells of animals (e.g. mice, rats, etc.) immunized with all or a portion of one or more epitope tags and/or fusion molecules of the invention using conventional procedures, such as those described by Kohler and Milstein (1975). Thus, the present invention provides monoclonal antibodies specific to the epitope tag and/or fusion molecule of the invention, as well as cell lines producing such monoclonal antibodies. Antibodies of the invention may be prepared from any animal origin including birds and mammals.

A nanobody (Nb) is the smallest functional fragment or single variable domain (V_HH) of a naturally occurring single-chain antibody and is known to the person skilled in the art. They are derived from heavy chain only antibodies, seen in camelids (Hamers-Casterman et al. 1993; Desmyter et al. 1996). In the family of "camelids" immunoglobulins devoid of light polypeptide chains are found. "Camelids" comprise old world camelids (*Camelus bactrianus* and *Camelus dromedarius*) and new world camelids (for example *Lama paccos*, *Lama glama*, *Lama guanicoe* and *Lama vicugna*). Said single variable domain heavy chain antibody is herein designated as a Nanobody or a V_HH antibody. NanobodyTM, NanobodiesTM and NanocloneTM are trademarks of Ablynx NV (Belgium).

The small size and unique biophysical properties of Nbs excel conventional antibody fragments for the recognition of uncommon or hidden epitopes and for binding into cavities or active sites of protein targets. Further, Nbs can be designed as bispecific and bivalent antibodies or attached to reporter molecules (Conrath et al. 2001). Nbs are stable, survive the gastro-intestinal system and can easily be manufactured. Therefore, Nbs can be used in many applications including drug discovery and therapy (Saerens et al. 2008).

As used herein, 'epitope' refers to an antigenic determinant of a polypeptide. An epitope could comprise 3 amino acids in a spatial conformation, which is unique to the epitope. Generally an epitope consists of at least 4, 5, 6, 7 such amino acids, and more usually, consists of at least 8, 9, 10 such amino acids. Methods of determining the spatial conformation of amino acids are known in the art, and include, for example, x-ray crystallography and 2- dimensional nuclear magnetic resonance.

The nanobodies of the invention generally comprise a single amino acid chain that can be considered to comprise 4 "framework sequences" or FR's and 3 "complementary determining regions" or CDR's. The term "complementary determining region" or "CDR" refers to variable regions in nanobodies and contains the amino acid sequences capable of specifically binding to antigenic targets. These CDR regions account for the basic specificity of the nanobody for a particular antigenic determinant structure. Such regions are also referred to as "hypervariable regions." The nanobodies have 3 CDR regions, each non-contiguous with the others (termed CDR1, CDR2, CDR3). Some preferred nanobodies of the present invention as well as their CDR's are as described herein (see Table 3 and 4). The delineation of the CDR sequences is based on the IMGT unique numbering system for V-domains and V-like domains (Lefranc et al. 2003).

In a particular embodiment, the nanobody of the invention is bivalent and formed by bonding, chemically or by recombinant DNA techniques, together two monovalent single domain of heavy chains. In another particular embodiment the nanobody of the invention is bi-specific and formed by

bonding together two variable domains of heavy chains, each with a different specificity. Similarly, polypeptides comprising multivalent or multi-specific nanobodies are included here as non-limiting examples.

Examples of the above nanobodies include, but are not limited to, nanobodies as defined by SEQ ID
5 NOs: 18-26 (see Fig. 1; Table 3). A non-limiting example of a bivalent nanobody is defined by SEQ ID
NO: 68. In another embodiment, the above nanobodies can comprise at least one of the
complementary determining regions (CDRs) with an amino acid sequence selected from SEQ ID NOs:
27-42. (see Fig. 1; Table 4). More specifically, the above nanobodies can be selected from the group
comprising SEQ ID NOs: 18-26 (see Fig. 1; Table 3).

10 It should be noted that the term nanobody as used herein in its broadest sense is not limited to a
specific biological source or to a specific method of preparation. For example, the nanobodies of the
invention can generally be obtained: (1) by isolating the V_HH domain of a naturally occurring heavy
chain antibody; (2) by expression of a nucleotide sequence encoding a naturally occurring V_HH domain;
(3) by "humanization" of a naturally occurring V_HH domain or by expression of a nucleic acid encoding
15 a such humanized V_HH domain; (4) by "camelization" of a naturally occurring VH domain from any
animal species, and in particular from a mammalian species, such as from a human being, or by
expression of a nucleic acid encoding such a camelized VH domain; (5) by "camelisation" of a "domain
antibody" or "Dab" as described in the art, or by expression of a nucleic acid encoding such a camelized
VH domain; (6) by using synthetic or semi-synthetic techniques for preparing proteins, polypeptides or
20 other amino acid sequences known per se; (7) by preparing a nucleic acid encoding a nanobody using
techniques for nucleic acid synthesis known per se, followed by expression of the nucleic acid thus
obtained; and/or (8) by any combination of one or more of the foregoing.

It is also within the scope of the invention to use natural or synthetic analogs, mutants, variants,
alleles, homologs and orthologs (herein collectively referred to as "analogs") of the antibodies of the
25 invention as defined herein. Thus, according to one embodiment of the invention, the term "antibody
of the invention" in its broadest sense also covers such analogs. Generally, in such analogs, one or
more amino acid residues may have been replaced, deleted and/or added, compared to the antibodies
of the invention as defined herein.

By means of non-limiting examples, a substitution may for example be a conservative substitution (as
30 described above, see also Table 1) and/or an amino acid residue may be replaced by another amino
acid residue that naturally occurs at the same position. Thus, any one or more substitutions, deletions
or insertions, or any combination thereof, that either improve the properties of the antibody of the

invention or that at least do not detract too much from the desired properties or from the balance or combination of desired properties of the antibody of the invention (i.e. to the extent that the antibody is no longer suited for its intended use) are included within the scope of the invention. A skilled person will generally be able to determine and select suitable substitutions, deletions or insertions, or suitable combinations of thereof, based on the disclosure herein and optionally after a limited degree of routine experimentation, which may for example involve introducing a limited number of possible substitutions and determining their influence on the properties of the antibodies thus obtained.

Yet another modification may comprise the introduction of one or more detectable labels or other signal-generating groups or moieties, depending on the intended use of the labelled antibody. Suitable labels and techniques for attaching, using and detecting them will be clear to the skilled person, and for example include, but are not limited to, fluorescent labels (such as fluorescein, isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde, and fluorescamine and fluorescent metals such as Eu or others metals from the lanthanide series), phosphorescent labels, chemiluminescent labels or bioluminescent labels (such as luminal, isoluminol, theromatic acridinium ester, imidazole, acridinium salts, oxalate ester, dioxetane or GFP and its analogs), radio-isotopes, metals, metals chelates or metallic cations or other metals or metallic cations that are particularly suited for use in vivo, in vitro or in situ imaging, as well as chromophores and enzymes (such as malate dehydrogenase, staphylococcal nuclease, delta- V- steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate dehydrogenase, triose phosphate isomerase, biotinavidin peroxidase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-VI-phosphate dehydrogenase, glucoamylase and acetylcholine esterase). Other suitable labels will be clear to the skilled person, and for example include moieties that can be detected using NMR or ESR spectroscopy. Such labelled antibodies of the invention may for example be used for in vitro, in vivo or in situ assays (including immunoassays known per se such as ELISA, RIA, EIA and other "sandwich assays", etc.) as well as in vivo imaging purposes, depending on the choice of the specific label. As will be clear to the skilled person, another modification may involve the introduction of a chelating group, for example to chelate one of the metals or metallic cations referred to above. Suitable chelating groups for example include, without limitation, diethylenetriaminepentaacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA). Yet another modification may comprise the introduction of a functional group that is one part of a specific binding pair, such as the biotin-(strept)avidin binding pair. Such a functional group may be used to link the antibody of the invention to another protein, polypeptide or chemical compound that is bound to the other half of the binding pair, i.e. through formation of the binding pair. For example, an antibody of the invention may be conjugated to biotin, and linked to another protein, polypeptide, compound or

carrier conjugated to avidin or streptavidin. For example, such a conjugated antibody may be used as a reporter, for example in a system where a detectable signal-producing agent is conjugated to avidin or streptavidin.

In still another embodiment, a nucleic acid sequence (as defined herein) encoding any of the antibodies, antibody fragments, and/or nanobodies as described herein is also aimed at. Further, a vector (as defined herein) comprising a nucleic acid sequence encoding any of the antibodies, antibody fragments, and/or nanobodies as described herein is also provided, as well as a host cell (as defined herein) comprising such a vector or such a nucleic acid.

Also provided in a further aspect of the present invention is a complex of (i) an affinity ligand as described above, and (ii) a non-naturally occurring fusion protein as described above, and (iii) optionally a ligand specifically bound to the target protein comprised in said non-naturally occurring fusion protein. In a specific embodiment, said complex is crystalline. As used herein, the term "ligand" means a molecule that specifically binds to a target protein. A ligand may be, for example a polypeptide, a lipid, a small molecule, an antibody or an antibody fragment (such as a nanobody). A ligand also includes a "native ligand" which is a ligand that is an endogenous, natural ligand for a native target protein.

As used herein, the term "complex" refers to a group of two or more associated molecules. In a preferred embodiment, a complex can be a protein complex. As used herein, the terms "protein complex" or "multiprotein complex" refer to a group of two or more associated polypeptide chains. Proteins in a protein complex are linked by non-covalent protein-protein. The quaternary structure is the structural arrangement of the associated folded proteins in the protein complex.

The epitope tags of the present invention, as well as the affinity ligands specifically binding to these epitope tags, as well as complexes of epitope tags and affinity ligands, and/or any combinations thereof, may serve any purpose including but not limited to:

- make a fusion molecule amenable to affinity purification methods,
- make a fusion molecule amenable to co-purification with an affinity ligand specifically binding to the epitope tag
- make a complex of proteins and/or nucleic acids that contains the fusion molecule amenable to affinity purification methods
- make a fusion molecule suitable for covalent or non-covalent immobilization,
- enable one to identify whether a fusion molecule is present in a sample or composition,
- enable one to localize a fusion molecule within a tissue
- localize a fusion molecule within a cell

- identify interactions of the fusion molecule with other proteins/protein complexes
- determine the function of a fusion protein and how it responds to changes in cell conditions
- determine what happens to the fusion protein during the lifetime of the cell
- characterize new proteins derived from cloned genes
- 5 – make a fusion molecule amenable to co-crystallisation with a nanobody specifically binding to the epitope tag

In particular, the epitope tags of the present invention, as well as the affinity ligands specifically binding to these epitope tags, as well as complexes of epitope tags and affinity ligands, and/or any combinations thereof, can be used for affinity-based applications, such as affinity chromatography, affinity purification, immunoprecipitation, in-vivo imaging, tandem affinity purification, protein detection, immunochemistry, surface-display, FRET-type applications, crystallography, and the like. The usefulness of the presently disclosed universal epitope tag in the above applications has been further illustrated in a non-limiting way in the Example section.

In a specific embodiment, the present invention relates to the use of any of the above described epitope tags and/or affinity ligands for the purification of fusion molecules comprising one or more epitope tags of the invention. As such, molecules can be purified using generalized protocols in contrast to highly customized procedures associated with conventional chromatography. Fusion molecules may be purified from the host cell or from the host cell culture medium into which they have been secreted. Typically, when purified from a host cell, the host cell is lysed using standard techniques (e.g., enzymatic digestion, sonication, French press, etc.) to form a lysate comprising the fusion molecule. Said fusion molecule may be purified from a lysate or from a host cell culture medium material by contacting the lysate or medium with a suitable chromatography medium under conditions suitable for binding of the fusion molecule to the chromatography medium. The lysate or culture medium may be contacted with a chromatography medium in either a batchwise technique (e.g. by mixing the chromatography medium with the lysate or culture medium) or column technique. The resin bound fusion molecule may be washed one or more times to remove any weakly bounded materials, i.e. materials that do not bind as tightly as the fusion molecule to the chromatography medium. The molecule may then be eluted from the medium by contacting the medium with a suitable elution buffer known to the skilled person. The elution of the fusion molecule from the column can be carried out at a constant pH or with linear or discontinuously falling or raising pH gradients. The optimal elution conditions depend on the amount and type of impurities which are present, the amount of material to be purified, the column dimensions, the chromatography resin used, etc. and are easily determined by routine experimentation on a case-by-case basis.

In a particular embodiment, the invention relates to a method for purifying a non-naturally occurring fusion molecule comprising the steps of

- a) applying a solution containing a molecule linked to the epitope tag as described herein to a solid support possessing an immobilized affinity ligand as described herein, and
- 5 b) forming a complex between said immobilized affinity ligand and said molecule, and
- c) removing weakly bound or unbound molecules, and
- d) eluting the bound molecule.

As discussed herein, a fusion molecule may comprise a cleavage site for a protease, for example, located between the tag of the invention and a molecule of interest. After elution from the chromatography medium or while still bound to the medium, a fusion molecule of the invention may
10 be contacted with a solution comprising a protease enzyme that cleaves at the cleavage site. Accordingly, in a specific embodiment, the above described purification method further comprises a step (e) wherein the epitope tag is removed.

The purification method may be any method known in the art. Suitable purification methods include
15 but are not limited to affinity chromatography (e.g. immunoaffinity chromatography, Immobilized Metal Ion Affinity Chromatography (IMAC)), metal-affinity precipitation, immobilized-metal-ion-affinity electrophoresis and Immunoprecipitation.

In case of affinity chromatography, the affinity medium contains a solid support (e.g. resin, magnetic bead,...) coupled to one or more of the following affinity ligands: Fe, Co, Ni, Cu, Zn, or Al charged IMAC
20 ligand, an antibody, an antibody fragment, a nanobody, a small molecule or a synthetic affinity ligand (e.g. aptamers or ligands derived using Versaffm™).

As is demonstrated herein, the use of the epitope tag and/or affinity ligand of the present invention for purification results in highly purified proteins with a good yield. The term "purity" or "purified" as applied to proteins herein implies that the desired protein preferably comprises at least 60%, more
25 preferably at least 70%, more preferably at least about 80%, still more preferably at least about 90%, and most preferably at least about 95% of the total protein component.

The elution step can be performed amongst others with a small compound, a peptide, using solutions of high ionic strength or by changing the pH. In a particular embodiment, the fusion molecule can be eluted from the affinity chromatography medium by excess amounts of the epitope tag or excess
30 amounts of the affinity ligand.

The present invention furthermore encompasses a method for immobilizing and/or capturing a non-naturally occurring fusion molecule on a solid support comprising the steps of:

- a) applying a solution containing a molecule linked to the epitope tag as described herein to a solid support possessing an immobilized affinity ligand,
- 5 b) forming a complex between said immobilized affinity ligand and said molecule, and
- c) removing weakly bounded molecules.

The method may be performed using immobilized elements and the immobilization may be carried out using a variety of immobilization means on various solid supports (e.g., columns, beads, adsorbents, nitrocellulose paper, etc.). The immobilization assay can, for example, be used to screen a sample or a composition for antibodies against the molecule linked to the tag, or can be part of a screening assay in order to screen large libraries of test compounds (e.g., drugs, new antimicrobials, etc.). The screening assay is preferably conducted in a microplate format. Any means or method of detection can be used. For example, the detection means might be a plate reader, a scintillation counter, a mass spectrometer or fluorometer.

15 The invention further relates to a method for identifying whether a molecule is present in a sample or composition. For example, the epitope tag of the present invention can be used in detection of a fusion molecule using the affinity ligand fused to a reporter molecule (radioactive label, alkaline fosphatase, GFP, peroxidases, and the like, see above). In a particular embodiment, the invention relates to the detection of a fusion protein according to the invention with antibodies specifically binding to the epitope tag. This embodiment can be useful in tissue or subcellular localization experiments, ELISA, 20 western blotting or other immuno-analytical methods.

In a particular embodiment, antibodies, including antibody fragments or nanobodies, may be used for the detection of the epitope tag in an immunoassay, such as ELISA, Western blot, radioimmunoassay, enzyme immunoassay, and may be used in immunocytochemistry. In some embodiments, an anti-tag antibody may be in solution and the tag to be recognized may be in solution (e.g. an immunoprecipitation) or may be on or attached to a solid surface (e.g. a Western blot). In other 25 embodiments, the antibody may be attached to a solid surface and the tag may be in solution (e.g. immuno affinity chromatography or ELISA).

Antibodies, including antibody fragments or nanobodies, to the tags and/or fusion molecules of the invention may be used to determine the presence, absence or amount of one or more molecules in a 30 sample. The amount of specifically bound tag and/or fusion molecule may be determined using an

antibody to which a reporter molecule is attached, such as a radioactive, a fluorescent, or an enzymatic label. Alternatively, a labeled secondary antibody (e.g. an antibody that recognizes the antibody that is specific to the polypeptide) may be used to detect a polypeptide-antibody complex between the specific antibody and the polypeptide.

- 5 It should be clear to the skilled person that the above described methods using the presently disclosed epitope tag can also be used in a high-throughput approach. Furthermore, combinatorial tagging also forms part of the present invention.

In another specific embodiment, the epitope tag of the present invention can also be useful in crystallography of proteins, where a short tag or no tag is preferred. Usually, tags are removed as it is a disadvantage for crystallization to have unstructured parts in the protein. Limited proteases are used to remove the flexible parts. Once a protein is tagged with an epitope tag of the present invention, an affinity ligand, in particular a nanobody, can bind the tag, making the whole complex more soluble, easier to crystallize and helping to solve the structure of the complex. It has been shown that Nb can help in crystallisation by fixating flexible loops or making crystal contacts (Loris *et al.*, 2003, Korotkov *et al.*, 2009).

In still another specific embodiment, the invention also provides for the use of the epitope tag and/or the affinity ligand of the present invention in a Tandem Affinity Purification (TAP) method in order to purify protein complexes under native conditions. The TAP tagging method (Rigaut *et al.*, 1999; Puig *et al.* 2001) can be applied to the yeast proteome very successfully, and has been adapted to various cells and organisms. However, whilst the method is intrinsically sound, it has a number of drawbacks, in particular, the size of the TAP tag can be in the region of 21 kDa and the larger a tag is, the greater the possibility that it will interfere with the function or structure of the polypeptide to be isolated, or alternatively be cleaved or digested by the polypeptide target. Commonly used tag combinations for TAP are well-known and summarized in Li (2010). The epitope tags of the present invention can be incorporated in the development of TAP tags for affinity-based applications.

In a further aspect, the invention also provides for a kit for use in any of the affinity-based applications as described herein. Such a kit may comprise, without the purpose of being limitative, one or more of the following elements:

- one or more (poly)peptides of the invention, optionally detectably labeled, and optionally associated with a non-soluble matrix or solid support

- one or more affinity ligands of the invention, optionally pre-immobilized to a non-soluble matrix or solid support
- nucleic acids of the invention
- vector constructs of the invention
- 5 - labeling reagents known in the art
- a set of instructions
- buffers, elution components, detection components known in the art

It is to be understood that although preferred embodiments, specific constructions and configurations, as well as materials, have been discussed herein for the methods and tools according to the present invention, various changes or modifications in form and detail may be made without departing from the scope and spirit of this invention. Therefore, the present invention is limited only by the claims attached herein.

TABLE 2. Epitope tags, peptides, linker sequences

Amino acid sequence	SEQ ID NO
EPEA	1
GYQDYPEA	2
DYPEA	3
YQDYPEA	4
QDYPEA	5
GYQDYPE	6
SEEGYQDYPEA	7
GAAEPEA	8
GAGAEPEA	9
GAGAGYQDYPEA	10
YEPEA	11
GAGA	12
EEGEPK	13
EEAEPK	14
IEGR	15
LVPR	16
DDDDK	17

TABLE 3: Nanobodies that specifically bind to α -synuclein

NAME (see Fig. 1)	No of AMINO ACIDS	SEQ ID NO
Nb85	127	18
Nb88	127	19
NbSyn2 PstI	126	20
NbSyn2	126	21
NbSyn1b	126	22
NbSyn1c	126	23
NbSyn1a	126	24
Nb86	124	25
Nb87	122	26

TABLE 4: CDRs of nanobodies that specifically bind to α -synuclein

Nanobody (see Fig. 1)	CDR1	CDR2	CDR3
Nb85	GLTYSNYC (SEQ ID NO 27)	ISTRGIKT (SEQ ID NO 28)	AAVIYPGYGDSCPWTTSVNY (SEQ ID NO 29)
Nb88	GYTYSRIC (SEQ ID NO 30)	ISTRGIKT (SEQ ID NO 28)	AAVIYPGYGDSCPWTTSVNY (SEQ ID NO 29)
NbSyn2 PstI	GIDSSSYC (SEQ ID NO 31)	INGLGGVKT(SEQ ID NO 32)	AAKFSPGYCGGSWSNFGY (SEQ ID NO 33)
NbSyn2	GIDSSSYC (SEQ ID NO 31)	INGLGGVKT(SEQ ID NO 32)	AAKFSPGYCGGSWSNFGY (SEQ ID NO 33)
NbSyn1b	GLNASSYC (SEQ ID NO 34)	INGNAGIKT(SEQ ID NO 35)	AAKSSPGYCGGNWDNFGY (SEQ ID NO 36)
NbSyn1c	GLNASSYC (SEQ ID NO 34)	INGNAGIKT(SEQ ID NO 35)	AAKSSPGYCGGNWDNFGY (SEQ ID NO 36)
NbSyn1a	GLNASSYC (SEQ ID NO 34)	INGNAGIKT(SEQ ID NO 35)	AAKSSPGYCGGNWDNFGY (SEQ ID NO 36)
Nb86	GYTFRGNR (SEQ ID NO 37)	INTGGVNT (SEQ ID NO 38)	AADLTGWRPVGFSGYNY (SEQ ID NO 39)
Nb87	GYSKY (SEQ ID NO 40)	IYRGDKIT (SEQ ID NO 41)	AARRVVADSPLLSKTYAY (SEQ ID NO 42)

EXAMPLES

Example 1. Isolation, expression and purification of nanobodies against α -synuclein

Camelid nanobodies e.g. NbSyn1a, NbSyn1b, NbSyn1c and NbSyn2, Nb85, Nb86, Nb87 and Nb88 were isolated from either a dromedary immunized with α -synuclein (Uniprot accession number P37840) or a
 5 dromedary immunized with two proteins: the α -synuclein Ala53Thr mutant and β -synuclein (Uniprot accession number is Q16143). α -Synuclein and β -synuclein share 85% identity in the N-terminal part of the protein, in the C-terminal part both proteins are less homologous ($\pm$ 40%). Both proteins share the same C-terminal EPEA sequence. All nanobodies were selected by phage-display, essentially according to published protocols (Lauwereys *et al.* 1998; Conrath *et al.* 2001) or as described in materials and
 10 methods. *In vitro* selection on immobilized α -synuclein only revealed one family of binders: NbSyn1a, NbSyn1b, and NbSyn1c. A second *in vitro* selection was performed on immobilized α -synuclein in the presence of 10 μ M of NbSyn1a. From this selection, we obtained NbSyn2 after 3 rounds of panning. From the dromedary immunized with two proteins (the α -synuclein A53T mutant and β -synuclein) one independent phage display library of 10⁷ independent clones was constructed. This library was panned
 15 in parallel against β -synuclein and the α -synuclein A53T mutant in two separate *in vitro* selections. Nb85, Nb86 and Nb87 were recovered from the beta-synuclein panning. Two of these binders (Nb85, Nb86) were also selected on Ala53Thr α -synuclein. One add binder Nb88 was also recovered from the Ala53Thr α -synuclein panning. It thus appears that Nb85 and Nb86 bind α -synuclein A53T as well as β -synuclein. Nb88 only differs from Nb85 by a few point mutations in framework 1 and CDR1 while the
 20 CDR2 and CDR3 are identical (Fig 1).

All nanobodies were recloned into pHEN6 to have six histidines at the C-terminus. The purification occurred using immobilized metal affinity chromatography (IMAC) and size exclusion chromatography essentially according to published protocols (Conrath *et al.* 2001).

Example 2. Epitope characterization of the nanobodies against synuclein

25 We used ¹⁵N-¹H heteronuclear single quantum correlation spectroscopy (HSQC) to map the epitope of NbSyn2, Nb85 and Nb87 on α -synuclein by observing perturbations to the backbone amide resonances. Titration of unlabeled Nb into ¹⁵N labeled α -synuclein resulted in both broadening and shifts of specific resonances in the HSQC spectrum of the latter.

The effect of binding of NbSyn2, NbSyn85 and Nb87 to the HSQC spectrum of labeled α -synuclein is
 30 shown in Fig. 2. The top panel represents the raw spectra obtained upon addition of Nb, where the red peaks correspond to the free and the blue peaks correspond to the signals of α -synuclein in complex with the different nanobodies. The lower panel of Figure 2 summarizes a titration experiment, using subequimolar to equimolar concentrations of Nb as the titrant. In the used

concentration range most of the affected HSQC cross-peaks of labeled α -synuclein either disappear completely (intermediate exchange) or reappear in a different region in the spectrum (slow exchange). This is highly consistent with the formation of a stable complex upon binding. When the relative intensity of the peaks of the free synuclein spectrum in the free and bound spectra is plotted for each amino acid along the sequence (bottom panel of Fig. 2), the C-terminal region is subjected to the greatest perturbations in terms of chemical shift or intensity upon binding of each of the nanobodies. These data indicated that all three Nb's bind to the C-terminal region of α -synuclein.

Example 3. Affinity measurements of nanobodies interacting with α -synuclein

Isothermal titration calorimetry (ITC) was used to monitor the binding of NbSyn2 to α -synuclein. Measurements were performed at temperatures between 20°C and 37°C and the data are consistent with a 1:1 bimolecular association between NbSyn2 and α -synuclein in each case, with K_d values of 106 ± 21 nM, 130 ± 23 nM, 99 ± 17 nM and 260 ± 69 nM at 20°C, 25°C, 30°C and at 37°C respectively (Table 5).

TABLE 5: Thermodynamic parameters of the NbSyn2: α -synuclein interaction as measured by Isothermal Calorimetry (ITC)

Ligand	T (K)	ΔH (kcal mol ⁻¹)	$T\Delta S$ (kcal mol ⁻¹)	ΔG (kcal mol ⁻¹)	Stoichiometry	K _d (nM)	ΔC_p (kcal mol ⁻¹ K ⁻¹) ^a
Peptide ^b	293.1	- 11.20 ± 0.24	2.31 ± 0.25	- 8.88 ± 0.07	0.98 ± 0.01	240 ± 26	- 0.22 ± 0.01
	298.1	- 12.44 ± 0.35	3.28 ± 0.36	- 9.16 ± 0.09	0.99 ± 0.01	190 ± 30	
	303.1	- 13.50 ± 0.76	4.72 ± 0.77	- 8.78 ± 0.14	0.94 ± 0.02	460 ± 110	
	310.1	- 14.89 ± 0.83	6.29 ± 0.84	- 8.61 ± 0.09	0.92 ± 0.03	850 ± 130	
Full-length α -synuclein	293.1	- 18.00 ± 0.53	8.65 ± 0.55	- 9.35 ± 0.12	0.99 ± 0.01	106 ± 21	- 0.23 ± 0.01
	298.1	- 19.05 ± 0.45	9.67 ± 0.46	- 9.39 ± 0.10	0.93 ± 0.01	130 ± 23	
	303.1	- 20.32 ± 0.42	10.61 ± 0.43	- 9.71 ± 0.10	0.87 ± 0.01	99 ± 17	
	310.1	- 21.79 ± 1.23	12.47 ± 1.24	- 9.33 ± 0.16	0.91 ± 0.03	260 ± 69	

^a The contribution of protonation to the value of ΔH was determined using a 10mM tris-HCL buffer 100mM NaCl at pH 7,4 and was found to be insignificant.

^b The sequence of the peptide, N-YEPEA-C (SEQ ID NO: 11), corresponds to the residues Y136-A140 of full-length α -synuclein.

A similar ITC measurement was performed as described in material and methods using NbSyn2 and a peptide fragment of α -synuclein, encompassing D135-A140 of the α -synuclein sequence (NH₂-DYEPEA-COOH). Similar binding characteristics were observed, indicating that this peptide contains the epitope of NbSyn2 on α -synuclein (see Example 4). We found that the corresponding affinities and ΔC_p value are very similar to those for the interaction with the full-length α -synuclein (Table 5). This finding

confirms that the ΔC_p value arises overwhelmingly from the direct binding of NbSyn2 to these residues located within the C-terminal peptide of α -synuclein, and not as a consequence of significant conformational changes within the remainder of the molecule.

Example 4. Crystallisation of the NbSyn2-Peptide complex

5 In another experiment, we crystallized NbSyn2 in complex with the peptide $\text{NH}_2\text{-GYQDYEEA-COOH}$, corresponding to residues 132-140 of α -synuclein (Fig. 6). In the X-ray structure of the complex we observe significant electron density only for the last six C-terminal residues of the peptide. The other residues of the peptide are not visible in the electron density, indicating that they are highly flexible and disordered in the complex and implying that they are not involved in specific interactions with
10 NbSyn2. The residue Asp135 only shows poor-defined electron density (Fig. 7 first top panel) and although there is clear density for Tyr136, this residue is involved in crystal packing contacts with a symmetry related complex (Fig. 7, second top panel), indicating that the conformation of the side chain of Tyr136 could be more flexible in solution.

Analysis of the structure at the atomic level reveals that the peptide binds to NbSyn2 in a pocket
15 formed by residues of the CDR3 and CDR2 loops of NbSyn2. The residues of CDR2 and CDR3 loops make contacts with residues Tyr136, Glu137, Pro 138, Glu139 and Ala140 of the peptide. The binding is primarily mediated through side-chain interactions and is mainly electrostatic in nature, although an important hydrophobic interaction occurs between Pro138 of the peptide and Tyr105 of NbSyn2, where the C_γ and C_δ of the cyclic aliphatic ring of Pro138 stack perpendicularly against the aromatic
20 ring of Tyr105. The C_β of Ala140 is deeply buried in a pocket formed between the CDR3 and CDR2 loops of the nanobody (Fig. 6b), while the carbonyl groups of the C-terminus of Ala140 make a salt-bridge with Arg50 of NbSyn2 (Fig. 7). The peptide main-chain atoms, apart from the carbonyl groups of Ala140 and the carbonyl oxygens of Pro138 and Glu137 -which make direct hydrogen bonds with residues of NbSyn2- are either exposed to solvent or make indirect contact with the nanobody through bridging
25 water molecules.

The determination of the crystal structure of NbSyn2 bound to a peptide encompassing the nine C-terminal residues of α -synuclein (PDB-code 2X6M) confirms that NbSyn2 binds the last four C-terminal residues of α synuclein and more specifically the C-terminal sequence EPEA (SEQ ID NO: 1) (Fig 6).

Example 5. Affinity purification of tagged GFP

30 The NMR experiments (Example 2) and the ICT data (Example 3) show that the nanobodies NbSyn2, Nb85 and Nb87 bind the C-terminal end of α -synuclein. The X-ray structure (Example 4) confirms that NbSyn2 interacts mainly to the last 4 amino acids. To further confirm that these nanobodies bind this C-terminal linear 4 amino acid epitope, we decided to construct different GFP-peptide hybrids. By

standard DNA cloning techniques we made 4 different constructs of a GFP (Green Fluorescent Protein) fused to different tags at the C-terminal end:

1. GFP-GYQDYEPEA (CA1508)
2. GFP-YQDYEPEA (CA1509)
- 5 3. GFP-QDYEPEA (CA1510)
4. GFP-GYQDYEPE (CA1511)

The four constructs were cloned into the pBADmycHis vector as described in material and methods.

All GFP hybrids were expressed in the cytoplasm of *E. coli*. No green fluorescence could be observed in cells transformed with construct CA1510, indicating that this fusion protein is poorly expressed.

10 However, functional GFP expression could be obtained for CA1508, CA1509 and CA1511, as indicated by the green fluorescence of the induced cells transformed with these constructs. Nanobodies Nb85, Nb87, Nb88 and NbSyn2 were expressed by standard procedures in the periplasm of *E.coli*. Next, periplasmatic fractions containing one of the nanobodies were mixed with cellular lysates containing one of the four tagged GFP proteins and incubated for one hour. Then, we copurified the different
15 GFP-fusions with the His-tagged nanobodies from these mixtures by IMAC. In this co-purification experiment high salt (1M NaCl) was added to avoid aspecific binding. Next, copurified Nb-GFP complexes were eluted from the column by imidazol and the eluted fractions were analyzed by SDS-PAGE (Fig. 3). Using Nb85 or NbSyn2, we were able to copurify the tagged GFPs encoded by CA1508 and CA1509 from crude cell lysate mixtures. The GFP construct where the last Alanine from the Tag
20 was missing (CA1511) could never be recovered indicating that this C-terminal Ala is crucial for the binding of these GFP fusions by Nb85 or NbSyn2.

In a second step an affinity column was prepared by immobilizing NbSyn2 on a sepharose matrix. NbSyn2 was coupled to a CNBr activated sepharose according to the manufactures instructions. In a separate experiment, GFP fused to the GYQDYEPEA tag (construct CA1508) was expressed in *E.coli*.

25 Cells were recovered by centrifugation and lysed. This crude lysate containing the tagged GFP (construct CA1508) was then mixed with agarose slurry containing the immobilized NbSyn2 and incubated for one hour at 4°C. Next all the beads were recovered on a filter in a small column and washed with 10 column volumes (cv) 50mM Tris pH 7.5 and with 10 cv of 50mM Tris/1M NaCl pH7.5 to get rid of aspecific binding and equilibrated with 50mM Tris pH7.5. We tried to elute of the
30 GYQDYEPEA tagged GFP with different buffers all containing 50mM TrisHCl and titrated to pH7.5: 0.5M citric acid, 0.5M N-L-a-aspartyl-L-phenylalanine-1-methylester (Aspartame), 0.5M aspartaat/tyrosine, 0.5M glutamic acid/tyrosine, 0.5M formic acid or 0.5M adipic acid,. Elution of the GFP fusion was monitored by measuring the fluorescence of the eluate: adipic acid as well as NaCl was found to elute

the tagged GFP from the affinity matrix. Eluates were next analysed on SDS-PAGE and confirmed this observation: one specific band with a molecular weight corresponding to the GFP fusion could be identified on SDS-PAGE (Fig. 4). In a separate experiment we trapped the GFP-GYQDYEEPA fusion from a crude lysate onto the NbSyn2 affinity column, washed it with 10cv of buffer and successfully eluted the tagged GFP with 750µM of the (synthesised) free NH₂-GYQDYEEPA-COOH peptide (Fig. 5). In this case we also identified one band with a molecular weight corresponding to the GFP fusion on SDS-PAGE. With a second high salt wash (2M NaCl) hardly any remaining protein could be eluted from the column indicating that the peptide quantitatively elutes GYQDYEEPA-tagged proteins with the synthetic GYQDYEEPA peptide.

In conclusion, using a NbSyn2 affinity column we were able to trap the GFP-GYQDYEEPA fusion (CA1508) from a cell lysate and to elute it from the affinity column with molar concentrations of NaCl or adipic acid or with millimolar concentrations of the synthetic NH₂-GYQDYEEPA-COOH peptide.

Example 6. Affinity purification of tagged nanobodies

As another example we also tagged a nanobody (that is specific for another protein) and purified it using a NbSyn2 affinity resin. To produce tagged nanobodies we constructed 2 new plasmids pMESy1 and pMESy2 that are described in detail in Material and Methods. In pMESy1 the original His-tag was replaced by the longer GYQDYEEPA. In pMESy2 we cloned the 4 amino acid EPEA tag behind a small GAA-linker. One VHH, Nb458 (recognizing an unrelated antigen or Aox1 of *Pichia pastoris*; SEQ ID NO: 66) was cloned in both vectors using the *Pst*I and *Eco*91I site. Both tagged versions of Nb458 were expressed in the periplasm of *E.coli*. Fresh periplasmic fractions were prepared and mixed with a slurry of NbSyn2 cross-linked to Sepharose-beads as described in example 4. The sepharose slurries were collected on the filter of small columns and washed with 50mM TrisHCl pH7.5. Both tagged nanobodies could successfully be eluted from these columns with 1M of NaCl (Fig. 8). In conclusion, using a NbSyn2 affinity column we were able to trap the Nb458-GYQDYEEPA fusion and the Nb458-gaaEPEA fusion from a cell lysate and were able to elute it from the affinity column with molar concentrations of NaCl.

Example 7: Application of the EPEA-tag in ELISA

First, we coated different wells of a 96-well plate with Nb458 tagged at its C-terminus with GYQDYEEPA and with Nb458 tagged with the gaaEPEA tag, respectively. Next, the plate was blocked with milk in PBS. Both immobilized EPEA-tagged Nb458 were detected using the His-tagged NbSyn2 (His-tag at the C-terminus of NbSyn2), followed by an antiHis mouse monoclonal and an anti mouse alkaline phosphatase conjugate. Significant alkaline phosphate activity could be detected only in those wells that were coated with Nb458 tagged at its C-terminus with GYQDYEEPA or with Nb458 tagged with the gaaEPEA tag as compared to the blank wells which were only blocked with milk. Next we coated different wells of an ELISA plate with (1) Nb458 with the hexa Histidine-epitope tag, (2) Nb458 with the GYQDYEEPA

tag and (3) Nb458 with the gaaEPEA tag. All wells were blocked with milk. In this experiment, the EPEA-tagged but not the His-tagged nanobody could be detected with a selfmade NbSyn2 alkaline phosphatase conjugate. A clear signal was obtained only for the GYQDYEPEA and the gaaEPEA tagged constructs (Fig. 9). It thus appears that the short EPEA tag can be used in combination with NbSyn2 to
5 detect recombinantly expressed proteins in ELISAs.

Example 8: Application of the EPEA-tag in Western Blot

We tested the EPEA-tag NbSyn2 combination in Western Blot analysis. In a first experiment GYQDYEPEA tagged GFP (CA1508) and α -synuclein were run on an SDS-PAGE and transferred to a
10 nitrocellulose sheet. Both blotted proteins could be visualized by using the His-tagged NbSyn2 as the primary antibody, a mouse monoclonal anti His-tag antibody as the secondary antibody and an anti-mouse alkaline phosphatase conjugate in combination with NBT and BCIP to develop the blot (Fig. 10). In a similar Western blot experiment, we could visualize blotted GYQDYEPEA tagged Nb458 (CA2944) and also Nb458-gaaEPEA (CA2945) in a single step with our selfmade NbSyn2-alkaline phosphatase
15 conjugate (NbSyn2-AP) in combination with NBT and BICP (Fig.11). It thus appears that the short EPEA tag can be used in combination with NbSyn2 to detect recombinantly expressed proteins in Western blots.

Example 9. In vitro detection and analysis of large protein libraries with the EPEA tag

It is clear from Example 7 that C-terminally EPEA tagged proteins can be detected in ELISA with NbSyn2-AP. To confirm that this tag also works in high throughput applications we used NbSyn2 to
20 screen large phage displayed libraries of EPEA-tagged VHH's. For this purpose, we cloned a VHH (nanobody) library (of an immunised llama) in phage display vector pMESy2. Two different *in vitro* selections by phage display were carried out to select subsets of nanobodies against two different
25 antigens. Screening of these subsets of clones for nanobodies that are specific for one antigen were performed by an ELISA as follows: Antigen was coated on 96-well plates, plates were blocked and each well was incubated with the periplasmatic fraction of another clone expressing a different VHH gene in fusion with the EPEA tag. Detection was performed with the NbSyn2 AP conjugate (Fig. 12). In both experiments, different specific binders were identified. It thus appears that the short EPEA tag can be
30 used in combination with NbSyn2 to detect recombinantly expressed proteins in a high-throughput analysis of phage display libraries and of libraries of recombinant proteins generated by other genomics or proteomics approaches.

Example 10. In vivo detection of EPEA tagged proteins with NbSyn2

In another set of experiments, we explored the use of the EPEA tag to detect/visualise the expression of recombinant proteins within a (living) eukaryotic cell. For this purpose, two different proteins (HIV-
 5 Rev and GFP) were EPEA tagged (Rev-GAGAGYQDYEEPA, Rev-GAGAEPEA, GFP-GYQDYEEPA and GFP-GAAEPEA) and cloned into a pcDNA transfection vector for transient expression in eukaryotic cells. To allow the detection of EPEA-tagged proteins NbSyn2 or NbSyn1a were fused to mKO1 or GFP and cloned into pcDNA to give NbSyn1a-mkO1, NbSyn2-mkO1, NbSyn1a-GFP and NbSyn2-GFP fusions. HeLa cells were transfected with these vectors or with combinations of these vectors to see whether
 10 we could detect the expression and specify the location of the EPEA-tagged proteins with GFP or mKO1 fused to NbSyn1a or NbSyn2. Fig 13 shows that the cells that are transfected with NbSyn2 or NbSyn1a fused to mKO1 or GFP colour the cell uniformly. Cells that are double transfected with Rev-GAGAGYQDYEEPA or Rev-GAGAEPEA and NbSyn2 or NbSyn1a fused to mKO1 or GFP have a deeper colour in the nucleoli. It is known that wild type Rev shuttles between the nucleoli and the cytoplasm
 15 of the cell with the majority of the protein localized in the nucleoli (Perkins *et al.*, 1989). It thus appears that the fluorescent nanobody fusions and the EPEA-tagged Rev protein colocalize in the nucleus of a double transfected cell, confirming that both the GYQDYEEPA and the EPEA tags are recognized *in vivo* NbSyn2 or NbSyn1a fusions, even in the nucleus of the cell.(Fig. 13).

Example 11. Improvement of the sensitivity of the in vitro detection using a bivalent nanobody

20 It is clear from Example 9 that the EPEA tag works in high throughput applications. To improve the sensitivity of the ELISA, we constructed a bivalent version of NbSyn2, named biNbSyn2 (SEQ ID NO: 68). For this purpose, we amplified the NbSyn2 gene and cloned it as a bivalent construct with a GGGGSGGGG (SEQ ID NO: 69) linker between both NbSyn2 sequences in the phage display vector pMES4. Expression, purification and conjugation of the biNbSyn2 to alkaline phosphatase was done as
 25 described in Material and Methods to the Examples. To compare the monovalent NbSyn2 with the bivalent “biNbSyn2” an ELISA was done as follows: antigen Nb458-gaaEPEA was coated on 96-well plates, at different concentrations, plates were blocked. Detection was performed either with the NbSyn2-AP conjugate or with the biNbSyn2-AP conjugate. From this experiment it is clear that the sensitivity of detection increases by using the biNbSyn2 (Fig. 14A, B).

Example 12. Detection and affinity purification of a tagged membrane protein

30 As another example we also tested whether tagged membrane proteins could be detected and purified by using affinity purification. Therefore, we tagged a membrane protein CsgG, checked its expression by Western blot and partially purified it using a NbSyn2 affinity resin. The *E. coli* UTI89 CsgG, cloned with its native RBS, behind the trc promotor in the pTrc99a (pMC2; Robinson *et al.*, 2006) was modified

to include a C-terminal EPEA affinity tag. Using the SLIM (Site-directed Ligation Independent Mutagenesis; Chiu et al. 2004) method, this C-terminal His tag was replaced with a sequence encoding GGEPEA (SEQ ID NO: 70) as described in detail in Material and Methods. A small scale culture of *E.coli* cells were grown, IPTG was added to induce the production of the EPEA tagged CsgG (SEQ ID NO: 72).

5 Western blot was done on 4 individual clones of this construct to prove the presence of the CsgG protein. A clear band running at the correct height could be detected (Fig 15). The plasmid was transformed to BL21 cells for a large scale production of the EPEA tagged CsgG. Cell membranes were prepared as described in the material an methods. Finally the membranes were solubilized overnight in 1% n-dodecyl beta-D-maltoside (DDM) and the solubilized material was run over a small column
10 containing the NbSyn2 cross-linked to Sepharose-beads as described in Example 4. The beads were washed with 25mM Tris pH8, 1mMDTT, 0.02% DDM, 1%C8E4, 5mM LDAO. A clear enrichment of the membrane protein could be seen after elution with the small synthetic NH₂-SEPEA-COOH peptide (SEQ ID NO: 71) demonstrating that capturing of EPEA tagged membrane proteins from 1% detergent solution is possible (Fig 16).

15

Materials and methods to the examples

All restriction enzymes, T4 DNA ligase were purchased from Fermentas.

Bacterial strains and plasmids

The *Escherichia coli* strains used were TG1: *supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5 (r_k⁻m_k⁻) [F' traD36 proAB lacI^qZDM15]*, Top10 [*mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 deoR recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (= Str^R) endA1 nupG*] (Invitrogen, Basel, Switzerland), WK6 [*Δ(lac-proAB), galE, strA/F' lacI^q, lacZΔM15, proA⁺B⁺*] (Zell & Fritz, 1987) and BL21 cells (Invitrogen).

Plasmids pHEN4 (Ghahroudi *et al.*, 1997), pHEN6 (Lauwereys *et al.*, 1998), pMES4 (genbank GQ907248), pBadMycHis (Invitrogen), pcDNA3.1. (Invitrogen) and pMC2 (Robinson et al. 2006) were
25 used.

Immunisation, isolation, expression and purification of nanobodies against α-synuclein

A dromedary (one) was 6 times immunized subcutaneously with α-synuclein during a six-week period. A second dromedary (two) was immunized with α-synuclein Ala53Thr and β-synuclein. From both animals total RNA was isolated from about 10⁷ lymphocytes as described by Chomczynski and Sacchi
30 (1987). First strand cDNA synthesis was prepared using a dN6 primer and the superscript RT according to the manufacturers (Invitrogen) instructions. From dromedary one the gene fragments encoding VHH genes were amplified by PCR using specific primers as described previously (Conrad et al., 2001), For

very slowly cooled down to room temperature. Hereafter the primers were diluted to a concentration of 1 μ M. Hundred ng or 0.034 pmol of purified pMES4 cut with *Bs*/WI and *Eco*91I was mixed with 1.7 pmol of primers and ligated ON with T4 DNA ligase. WK6 *E.coli* cells were transformed, analysed by PCR and sequenced.

5 NMR measurements

All NMR experiments were performed using Bruker Avance 500 MHz or 700 MHz spectrometers equipped with cryo-probes. All NMR data were subsequently processed using NMRpipe72 and the program SPARKY73 was used for analysis of the data.

10 ¹⁵N-¹H HSQC measurements with or without ¹⁴N NbSyn2 and assignment of the bound form of labeled α -synuclein

Standard ¹⁵N-¹H HSQC experiments were carried out on a Bruker Avance 700 MHz instrument equipped with a cryo-probe. Spectra of ¹⁵N labeled α -synuclein were recorded at different molar equivalents of unlabeled NbSyn2 (0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.5, 3.0, 3.25) by titration of small volumes of a concentrated stock solution of NbSyn2. All experiments were recorded in 10 mM phosphate buffer pH 7.4 and at 10°C. The amide resonances and C _{β} and C _{α} chemical shifts were determined previously (Dedmon et al., 2005a; 2005b) and used in the further analysis of the data. The chemical shifts of the α -synuclein resonances in complex with NbSyn2 were assigned using a series of standard 3D experiments: CBCA(CO)NH and HNCACB. Sample integrity checks were performed between and after each 3D experiment by recording ¹⁵N-¹H HSQC spectra.

20 ¹³C-¹⁵N CON measurements of ¹⁵N ¹³C labeled α -synuclein with or without ¹⁴N NbSyn2

¹³C-¹⁵N CON experiments, which correlate the ¹³C carbonyl chemical shift of residue n with the ¹⁵N chemical shift of residue n+1, were performed as described previously (Bermel *et al.*, 2006). The assignments of all crosspeaks in the spectra were obtained from Bermel *et al.*, 2006. In total two CON spectra were recorded for ¹⁵N ¹³C labeled α -synuclein free and after addition of one equivalent of ¹⁴N NbSyn2. The spectra were recorded on a 500 MHz Bruker Avance spectrometer equipped with a cryo-probe. The sample buffer was 10 mM phosphate buffer at pH 7.4 and the temperature was set at 283 K. The assignments of the CON crosspeaks of α -synuclein in the bound form were obtained by comparing the ¹⁵N chemical shifts for the corresponding peaks in the assigned ¹⁵N-¹H HSQC spectrum for bound α -synuclein.

¹⁵N-¹H HSQC measurements and assignment of ¹⁵N ¹³C labeled NbSyn2 bound to full-length α -synuclein and a peptide fragment of α -synuclein

We have reported the assignments of the backbone resonances of NbSyn2 elsewhere (Vuchelen *et al.*, 2009). The backbone assignments of samples of ¹⁵N ¹³C labelled NbSyn2 at 0.3 mM in 20 mM Phosphate buffer at pH 7.4 bound to α -synuclein were obtained by means of HNCA experiments using a Bruker Avance 700 MHz spectrometer equipped with a cryoprobe. Titrations of full-length ¹⁴N α -synuclein and a ¹⁴N 12-residue peptide, N-SEEGYQDYEP-EA-C (Genemed Synthesis, Inc, New York), were carried out with ¹⁵N labelled NbSyn2 at 0.3 mM in 20 mM phosphate buffer pH 7.4 at 298 and 283 K.

Isothermal calorimetry measurements.

Calorimetric measurements were recorded using a MCS-ITC calorimeter (MicroCal, LLC, Northampton, MA, USA). A solution of 273 μ l NbSyn2, at a concentration of 75 μ M was titrated in aliquots of 10 μ l into the calorimetric cell, containing 1.4 ml of 5 μ M α -synuclein. Both proteins were dialysed prior to measurements in exactly the same buffer containing 10 mM phosphate, 150 mM sodium chloride at a pH of 7.4. Each injection was performed every 4 min at the desired temperature. A titration of NbSyn2 in the sample cell containing only buffer was subtracted from the actual binding experiment before analysis. The thermodynamic analysis was performed using the Microcal analysis software (Origin 7.0) with a 1:1 binding model. The temperature dependence of ΔH , allowed the calculation of ΔC_p through the relationship $\Delta C_p = \delta\Delta H/\delta T$. The contribution of protonation to the ΔH was evaluated using a 10mM Tris-HCL buffer 100mM NaCl at pH 7.4.

Cloning, expression in *E.coli* and purification of tagged GFP

Four different tags were attached to the *gfp* gene by PCR using Takara Ex Taq polymerase (TAKARA BIO INC). The same specific primer EP124 (5' TAATCATGAGCAAAGGAGAAGAACTTTTCAC 3') (SEQ ID NO: 67) was used in all constructs, it anneals at the N-terminus of the gene. Different primers were used to fuse the coding sequence of different tags at the C-terminus: for GYQDYEP-EA: primer EP126 (5'-TTCGAATT-CATTACGCTTCCGGTTCATAATCCTGATATCCTTTGTAGAGCTCATCCATGC 3') (SEQ ID NO: 49) was used, for YQDYEP-EA we used primer EP127 (5'-TTCGAATTCATTACGCTTCCGGTTCATAATCCTGATATTTGTAGAGCT-CATCCATGCC-3') (SEQ ID NO: 50) and for QDYEP-EA primer EP128 (5'-TTCGAATTCATTACGCTTCCGGTTCATAATCCTGTTTGTAGAGCTCATCCATGCC-3') (SEQ ID NO: 51). Amplification conditions were 4min 94°C, 28 times (40sec 94°C, 40sec 50°C, 1min 72°C) and 10 min 72°C. The amplified *gfp* fusions coding for GFP fusion molecules with different epitope tags were cut with *Bsp*HI and *Eco*RI. The PCR products were ligated into the pBADmycHis vector, opened with *Nco*I and *Eco*RI. Finally the ligation product was

transformed into *E.coli* Top10. To construct GFP with the GYQDYEPE tag 1µl of the EP124-EP126 amplified GFP fragment was used as template for an extra PCR using primers EP124 and EP129 (5' TTCGAATTCATTATTCCGGTTCATAATCCTG 3') (SEQ ID NO: 52). The amplified PCR product was cloned to the pBADmycHis vector as written above.

- 5 To express the differently tagged GFP fusions, *E. coli* cell cultures were induced by the addition of 0.02% arabinose at a cell density of OD₆₀₀= 0.7 and were further incubated for overnight at 37°C. The cells from a 330ml LB culture were harvested by centrifugation (10 min, 4000rpm, 4 °C) and the pellet was resuspended in 20ml 50mM Tris-HCl pH 7.5 with 1 mM PMST. Cells were lysed by a C5 EmulsiFlex High Pressure Homogeniser 86-001 (AVESTIN, Germany). The insoluble material was
10 pelleted by centrifugation (30 min, 4000rpm, 4 C) and the supernatant (cytoplasmatic fraction) was collected.

To purify the tagged GFP in complex with the hexa histidines tagged (His-tagged) NbSyn2, a periplasmatic fraction of the His-tagged NbSyn2 was mixed with one cytoplasmatic fraction of each tagged GFP. An IMAC was done as described before. The NbSyn2-GFP tagged fusion complexes were
15 eluted with 1M imidazol.

Cloning, expression in *E.coli* and purification of VHH –EPEA tagged fusions

On 1ng of pMES4 a PCR was performed with Takara Ex Taq polymerase using primers EP138 (5'-ACCCAGGTCACCGTCTCCTCAGGGTATCAAGACTACGAACCTGAAGCCTAGTACCCGTACGACGT) (SEQ ID NO: 53) and FP (5'-CGCCAGGGTTTTCCAGTCACGAC-3') (SEQ ID NO: 54) to construct pMESy1 (CA2857) and
20 u s i n g p r i m e r s E P 1 3 9 (5 ' -ACCCAGGTCACCGTCTCCTCAGGGGCAGCGGAACCTGAAGCCTAGTACCCGTACGACGTTC-3') (SEQ ID NO: 55) and FP to construct pMESy2 (CA2741). Amplification conditions were 4min 94°C, 28 times (40sec 94°C, 40sec 55°C, 1min 72°C) and 10 min 72°C. Both PCR products and the pMES4 vector were cut with *Bam*HI and *Eco*91I and purified. Ligation with T4 DNA polymerase of 100ng vector to 50ng PCR product
25 was carried out overnight at 4°C. The mixture was transformed into WK6 cells and overnight grown at 37°C. Analysis was done by PCR using primers MP57 (5'-TTATGCTTCCGGCTCGTATG -3') (SEQ ID NO: 56) and GIII (5'- CCACAGACAGCCCTCATAG -3') (SEQ ID NO: 57) and *Spe*I digestion. Clones were sequenced and from the correct ones DNA was prepared (Qiagen).

The Nb458 was amplified using primer RP and primer EP101 (5'-TGAGGAGACGGTGACCTGGGT-3') (SEQ
30 ID NO: 58) with the same amplifications conditions as written above and cloned in the pMESy1 (CA2857) or the pMESy2 (CA2741) vector using the *Pst*I and the *Eco*9I restriction enzymes. Insert analysis was done by PCR with primers RP and GIII. Expression was done as described by Conrad *et al.* (2001). Nb458 in pMESy1 is fused to the sequence of the GYQDYEPEA tag and is referred to as CA2944 while in pMESy2, it has the gaaEPEA tag and is referred to as CA2945.

NbSyn2 affinity column purification

Purified His-tagged NbSyn2 was coupled to a CNBr activated sepharose (GE Health care Life science) following the manufacturer's instructions. The column was equilibrated with 50mM TrisHCl pH7.5. Lysate of tagged fusions (GFP or VHH) was brought onto the beads for 1 hour. The column was washed
 5 with 10 column volumes (cv) of 50mM TrisHCl pH7.5 and elution was preformed with 10cv of 50mM TrisHCl with extra compounds (see examples) all pH7.5. In the last experiments elution buffer was either 50mM TrisHCl pH7.5 with 1M or 2M NaCl, 50mM Tris-HCl pH7.5 with 0.5M adipic acid or 50mM Tris-HCl pH7.5 with 750μM of peptide (GYQDYEPEA).

Solid phase ELISA with His-tagged NbSyn2

10 Maxisorb 96-well plates (Nunc) were coated with Nb458-gaaEPEA or Nb458-GYQDYEPEA preparations overnight at 4°C at 2μg/ml in sodium bicarbonate buffer pH 8.2. Residual protein binding sites in the wells were blocked for two hours at room temperature with 2% milk in PBS. NbSyn2-His6tagged was added at 2μg/ml in PBS with 0.2% milk. Detection of NbSyn2 nanobodies was performed with a mouse anti-histidine-tag (Serotec). Subsequent detection of the mouse anti-his antibody was done with an
 15 alkaline phosphatase anti-mouse-IgG conjugate (Sigma). The absorption at 405 nm was measured 15min after adding the enzyme substrate p-nitrophenyl phosphate.

Solid phase ELISA, Western blot with NbSyn2 conjugated to Alkaline phosphatase

0.5mg of His-tagged NbSyn2 in PBS was conjugated to with 2.5mg of alkaline phosphatase in a final volumn of 0.5ml as described by Harlow & Lane (1988). Maxisorb 96-well plates (Nunc) were coated
 20 with protein preparations overnight at 4 °C at 2μg/ml in sodium bicarbonate buffer pH 8.2. SDS-PAGE and Western blot were done according to standard protocols. Residual protein binding sites in the wells or on the membranes were blocked for two hours at room temperature with 2% milk in PBS. Detection of the tags was performed with a dilution of 1/1000 NbSyn2-AP conjugate. For ELISA the absorption at 405 nm was measured 15 to 30min after adding the enzyme substrate p-nitrophenyl
 25 phosphate. For Western NBT/BCIP substrate (Roche) was used as described by the manufacturer.

Crystallisation of the NbSyn2-Peptide complex

Crystallization of the NbSyn2 complexed with a synthetic peptide (sequence NH₂-GYQDYEPEA-COOH) (BIO-SYNTHESIS, INC., Texas) was achieved using a crystallization robot Phoenix (Art Robbins
 30 Instruments) and a sitting drop vapor diffusion assay. 133 μl of 10 mg/ml of NbSyn2 was first mixed with 15 μl of peptide (at 20 mg/ml) before dispension in a JB Classic 1-4 (G10) screen (Jena Biosciences). 100 nl drops were mixed with 100 nl of the precipitant solution. Crystals were obtained under the following conditions: 25% w/v PEG 6000, 100mM HEPES, pH 7.5 and 100 mM LiCl. Crystals

were subsequently harvested and transferred in a cryo-protectant solution (25% w/v PEG 6000, 100mM HEPES, pH 7.5 and 100 mM LiCl and 10% PEG400) and flash-frozen in liquid nitrogen for data collection.

Data collection and structure solution

5 Diffraction patterns for frozen crystals of the NbSyn2-peptide complex were obtained at the X11 beam line equipped with a MAR555 detector, at the European Molecular Biology Laboratory (EMBL) using the Deutsches Elektronen Synchrotron (DESY Hamburg, Germany). Data were processed using Mosfilm65, and Pointless and Scala were used to determine the spacegroup and scale and merge the data respectively. Molecular replacement using the program Phaser66 and a model for NbSyn2,
10 namely pdb entry 1HCV67, was used to obtain the phase information associated with the structure factors. Model building and refinement were achieved using the programs Arp/Warp68 and Refmac569. The graphics program Coot70 was used to interpret the electron density maps and for refinement of the model. The change in solvent accessible surface area (ASA) upon the formation of the NbSyn2: α -synuclein complex was calculated using web based server PROTORP71.

15 Cloning of the long (GYQDYEPEA) and short (EPEA) tag for expression in HeLa cells

The two different tags GYQDYEPEA and EPEA were C-terminal fused to Rev by PCR using Takara Ex Taq polymerase (TAKARA BIO INC). A PCR was performed on pcDNARev-mKO1 kindly given to us by Dirk Daelemans with primer T7 (5'-ATTAATACGACTCACTATAGG-3') (SEQ ID NO: 59) and primer EP136 (5'-TCGAGCGGCCGCTTAGGCTTCAGGTTCTGATACCTGAAGCTCCTGCACCATTCTC-3') (SEQ ID NO:
20 60) to amplify the rev gene fused to the DNA sequence of GYQDYEPEA at the C-terminus. To fuse the EPEA tag to Rev primer T7 and EP176 (5'-TCGAGCGGCCGCTTAGGCTTCAGGTTCTGATACCTGAAGCTCCTGCACCATTCTC-3') (SEQ ID NO: 61) were used. Amplification conditions were 2min 94°C, 25 times (30sec 94°C, 30sec 52°C, 1min 72°C) and 10 min 72°C. The amplified DNA fragments were cut with KpnI and NotI and ligated into the pcDNA3.1 vector.
25 The ligation product was electroporated into *E.coli* Top10. Clones were analysed by PCR and sequence analysis. At the same time both tags were also fused to GFP as described above, using primer EP180 (5'-GCTTGGTACCTCATGAGCAAAGGAGAAGAAC-3') (SEQ ID NO: 62) and primer EP126 (5'-TTCGAATTCATTACGCT-TCCGGTTCATAATCCTGATATCCTTTGTAGAGCTCATCCATGC-3') (SEQ ID NO: 63) to fuse it to the GYQDYEPEA and primer EP180 and primer EP181 (5'-TTCGAATTCATTACGCTTCCGGTCT-
30 GCGGCTCCTTTGTAGA-GCTCATCCATGCC-3') (SEQ ID NO: 64) to fuse GFP to EPEA. The amplified DNA fragments were cut with KpnI and EcoRI and ligated into the pcDNA3.1 vector. The ligation product was then electroporated into *E.coli* Top10, clones were analysed by PCR and sequence analysis.

Transfection of HeLa cells

HeLa, human epithelial cells were maintained using standard procedures until confluent in DMEM medium with 10% Fetal calf serum. For transfection of plasmid DNA in HeLa, cells were harvested by a Trypsin/EDTA treatment, counted and were plated onto glass bottom microwell dishes (MatTek corporation). Transfections were performed according to manufacturer's manual: the 4×10^5 cells were seeded in a 35mm dish. 1.2µg DNA or in case 2 different plasmids are cotransfected twice 1µg is mixed with 4.5µl Transfect (Qiagen Valencia, California), dropped onto the cells and incubated overnight.

Confocal microscopy

Transfected HeLa cells were imaged with a laser scanning SP5 confocal microscope (Leica Microsystems) equipped with an AF 6000 microscope and an AOBs, using a HCX PL APO ×63 (NA 1.2) water immersion objective magnification. GFP was monitored with the Ar laser using the 488-nm line for excitation, and emission was detected between 492 and 570 nm. mKusabira orange fluorescent protein (mKO) was imaged using the DPSS-561 nm laser for excitation, and emission was detected between 570 and 675 nm.

Cloning, expression in *E.coli* and purification of bivalent biNbSyn2

On 2ng of the pHen6(c)NbSyn2 vector (EP358) two different PCRs were performed with kapa Taq polymerase using once primer SM17 (5'- CGCGGCCAGCCGGCCATGGCT-GATGTGCAGCTGGTG-GAGTCTGG-3') (SEQ ID NO: 73) and primer VHH1-reverse-5GS-BamHI (5'- AGTAGGATCCGCCACCTCCTGAGGAGACCGTGACCTGGGT-3') (SEQ ID NO: 74) and once primer VHH2-forward-4GS-BamHI (5'-TCTTGGATCCGGCGGAGGTAGTCAGGTGCAGCTGCAGGAGTCTGGGGGAGG-3') (SEQ ID NO: 75) and primer VHH2-reverse-TVSS-BstII (5'-TGAGGAGACGGTGACCTGGGT-3') (SEQ ID NO: 76) to amplify the NbSyn2 gene. Amplification conditions were 4min 94°C, 28 times (40sec 94°C, 40sec 55°C, 1min 72°C) and 10min 72°C. The PCR product generated by using primer SM17 and VHH1-reverse-5GS-BamHI was cut with *NcoI* and *BamHI*, the second PCR product generated by using primers VHH2-forward-4GS-BamHI and VHH2-reverse-TVSS-BstII was cut with *BamHI* and *Eco91I*. The pMES4 vector was cut with *NcoI* and *BamHI* and *Eco91I*. All digest were purified and the DNA concentration was measured. A three-point-ligation with T4 DNA polymerase of 100ng vector to 30ng of both PCR products was carried out in 10µl for 2h at room temperature. The mixture was diluted and cut with *XbaI* to eliminate the self-ligated vectors. The digest was purified and transformed into WK6 cells and overnight grown at 37°C. Analysis was done by PCR using primers MP57 (5'-TTATGCTTCCGGCTCGTATG-3') (SEQ ID NO: 56) and GIII (5'- CCACAGACAGCCCTCATAG -3') (SEQ ID NO: 57) and *BamHI* digestion. From the clones that had the correct PCR fragment size containing a *BamHI* site, the PCR fragment was

sequenced and from a correct one plasmid DNA was prepared (Qiagen) and referred to as pMES4biNbSyn2 (CA4394) .

Expression and purification was done as described by Conrad *et al.* (2001). After IMAC purification, the sample was concentrated and ran on preparative superdex75 size exclusion column to separate the bivalent biNbSyn2 from contaminants or degradation products.

Solid phase ELISA with biNbSyn2 conjugated to alkaline phosphatase

0.66mg of His-tagged biNbSyn2 in PBS was conjugated to 1.7mg of alkaline phosphatase, in an parallel experiment 0.33mg of the monovalent NbSyn2 was conjugated to 1.7mg of alkaline phosphatase. In both cases the final volume was of 0.33ml and the conjugation occurred as described by Harlow & Lane (1988). Maxisorb 96-well plates (Nunc) were coated with 100µl protein preparations overnight at 4 °C in sodium bicarbonate buffer pH 8.2 at a concentration of 1µg/ml or 0.1µg/ml. Residual protein binding sites in the wells were blocked for two hours at room temperature with 2% milk in PBS. Detection of the tags was performed with a dilution of 1/250 to 1/4000 of both NbSyn2-AP conjugate and biNbSyn2-AP conjugate. For ELISA the absorption at 405 nm was measured 60min after adding the enzyme substrate p-nitrophenyl phosphate and after overnight incubation.

Cloning, expression in *E.coli* of the CsgG membrane protein and its detection/purification

To replace the C-terminal His tag of CsgG with a sequence encoding GGEPEA the pMC2 plasmid was used in the SLIM method (Chiu et al. 2004) with the following primers: 5'-GGCGGCGAACCGGAAGCGTGACTGCAGGCATGCAAGCTTGGCT-3' (SEQ ID NO: 77), 5'-TGACTGCAGGC-ATGCAAGCTTGGCT-3' (SEQ ID NO: 78), 5'-CGCTTCCGGTTCGCCGCCATGGTGATGGTGATGGTGGGATTCC-3' (SEQ ID NO: 79), 5'- ATGGTGATGGTGATGGTGGGATTCC 3'(SEQ ID NO: 80). PCR and sequence analysis confirmed the correct replacement. To confirm the expression of the CsgG protein 4 different clones of the same construct were picked and a 25ml *E.coli* culture was grown and induced. Cells were harvested and separated by SDS-PAGE and transferred to nitrocellulose membranes. The blot was blocked with 2% milk and washed 5 times with 20ml PBS/Tween. NbSyn2-AP conjugate (monovalent) was diluted 1 in 500 in 0,2 % milk in PBS and shaken for 1h. The blot was washed for 5 more times and developed with NBT/BCIP substrate (Roche) as described by the manufacturer.

To purify the expressed EPEA tagged CsgG, 110g of frozen cell pellet was resuspended in 330ml 50mM Tris, 150mM NaCl, 1mM DTT, 1mM EDTA pH8.0. 20µg DNase/1g pellet and 0.1mg lysozyme/ml buffer was added. Cells were lysed by the French press and the cell debris was removed by a short centrifugation step (7,500g for 12 min at 4°C). The supernatant was then centrifuged at high speed to recover the membranes (100000g for 50min at 4°C). Membranes were homogenized in 100ml 25mM Tris, 150mM NaCl, 1mM DTT, 1mM EDTA pH8.0 with 0.5% sarcosine. The suspension was centrifuged

at high speed .(100000g for 50min at 4°C) and the pellet (=outer membranes) was collected. The pellet was resuspended in 100ml 25mM Tris, 150mM NaCl, 1mM DTT, 1mM EDTA pH8.0 with 2mM LDAO using a homogeniser. After centrifugation, the pellet (membranes) was washed by resuspending it 50ml 25mM Tris, 150mM NaCl, 1mM DTT, 1mM EDTA pH8.0 with 0.002% DDM. Suspension was 5 centrifuged (100000g for 50min at 4°C) to collect the outer membranes. These were resuspended in 100ml 25mM Tris, 150mM NaCl, 1mM DTT and 1% DDM and left to shake overnight at 4°C to solubilize the membranes. One last centrifugation step (100000g for 50min at 4°C) was done to remove the insoluble parts: the supernatant was run over a small column containing the 1ml CNBr-NbSyn2 cross-linked to Sepharose-beads as described in example 4. The beads were washed with 5ml 25mM Tris 10 pH8, 1mM DTT, 0.02% DDM, 1% C8E4, 5mM LDAO. The elution of bound protein was done with 2 times 1ml 25mM Tris pH8, 150mM NaCl, 1mM DTT, 0.02% DDM containing 750µM of the small synthetic NH₂-SEPEA-COOH peptide.

REFERENCES

- Ausubel et al. (1992) Current Protocols in Molecular Biology, Greene/Wiley, New York, N.Y.
- Ausubel, F. M., et al. Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (2007). Current Protocols in Molecular Biology; John Wiley and Sons, Inc.
- 5 • Bermel, W., Bertini, I., Felli, I.C., Lee, Y.M., Luchinat, C. & Pierattelli, R. (2006). Protonless NMR experiments for sequence-specific assignment of backbone nuclei in unfolded proteins. J Am Chem Soc. 128, 3918-9.
- Brizzard, B.L., R.G. Chubet, and D.L. Vizard (1994). Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution.
- 10 BioTechniques 16:730-735.
- Brizzard, B. (2008). Epitope Tagging. BioTechniques 44:693-695 (25th Anniversary Issue, April 2008)
- Chiu J, March PE, Lee R, Tillett D. Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h Nucleic Acids Res. 2004, 32:e174.
- Chomczynski, P. and Sacchi, N., 1987. Single step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, p. 156
- 15 • Chong, S., F.B. Mersha, D.G. Comb, M.E. Scott, D. Landry, L.M. Vence, F.B. Perler, J. Benner, et al. (1997). Single-column purification of free recombinant proteins using self-cleavable affinity tag derived from a protein splicing element. Gene 192:271-281.
- Conrath, K. E., Lauwereys, M., Galleni, M., Matagne, A., Frere, J. M., Kinne, J., Wyns, L. &
- 20 Muyldermans, S. (2001). Beta-lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. *Antimicrob Agents Chemother.* **45**, 2807-12
- Conrath, K. E., Wernery U., Muyldermans S. and Nguyen V. K. (2003). Emergence and evolution of functional heavy-chain antibodies in Camelidae. Dev. Comp. Immunol. **27**, 87-103.
- Davis *et al.*, (1999) New fusion protein systems designed to give soluble expression in *Escherichia coli*, *Biotechnol. Bioeng.* **65**, 382–388.
- 25 • Dedmon, M.M., Christodoulou, J., Wilson, M.R. & Dobson, C.M. (2005a). Heat shock protein 70 inhibits alpha-synuclein fibril formation via preferential binding to prefibrillar species. J Biol Chem. 280, 14733
- Dedmon, M.M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M. & Dobson, C.M. (2005b).
- 30 Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. J Am Chem Soc. 127, 476-7
- Desmyter A, Transue TR, Ghahroudi MA, Thi MH, Poortmans F, Hamers R, Muyldermans S, Wyns L. (1996) Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. Nat Struct Biol. 3, 803-11.

- di Guan, C., P. Li, P.D. Riggs, and H. Inouye (1988). Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* 67:21-30.
- Evan, G.I., G.K. Lewis, G. Ramsay, and J.M. Bishop. 1985. Isolation of monoclonal antibodies specific
5 for human c-myc proto-oncogene product. *Mol. Cell. Biol.* 5:3610-3616.
- Field, J., J. Nikawa, D. Broek, B. MacDonald, L. Rodgers, I.A. Wilson, R.A. Lerner, and M. Wigler. 1988. Purification of RAS-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. *Mol. Cell. Biol.* 8:2159-2165.
- Ghahroudi M.A., Desmyter A., Wyns L., Hamers R., Muyldermans S. (1997) Selection and
10 identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.* 414, 521-526
- Giliams T, Degenst E & Dobson, C.M. (2010) Nanobodies to Investigate the Mechanisms Underlying the α -Synuclein Related Parkinson's Disease. First year report
- Gloeckner, C.J., K. Boldt, A. Schumacher, R. Roepman, and M. Ueffing (2007). A novel tandem
15 affinity purification strategy for the efficient isolation of native protein complexes. *Proteomics* 7:4228-4234.
- Hackbarth, J.S., S.H. Lee, X.W. Meng, B.T. Vroman, S.H. Kaufmann, and L.M. Karnitz (2004). S-peptide epitope tagging for protein purification, expression monitoring, and localization in mammalian cells. *BioTechniques* 37:835-839.
- Hamers-Casterman C., Atarhouch T., Muyldermans S., Robinson G, Hamers C, Songa E.B., Bendahman N. and Hamers R. (1993). Naturally occurring antibodies devoid of light chains. *Nature* 363, 446-8.
- Harlow, E. & Lane, D. P. (1988) *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Hochuli, E., W. Bannwarth, H. Dobeli, R. Gentz, and D. Stuber (1988) Genetic approach to facilitate
25 purification of recombinant proteins with a novel metal chelate. *BioTechnology* 6:1321-1325.
- Hopp et al. "A short polypeptide marker sequence useful for recombinant protein identification and purification", *Nature Biotechnology*, vol. 6, pp. 1204-1210 (1988).
- Jarvik, J.W. and C.A. Telmer (1998). Epitope tagging. *Annu. Rev. Genet.* 32:601-618.
- Kapust, R.B & Waugh, D.S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective
30 at promoting the solubility of polypeptides to which it is fused, *Protein Sci.* 1668-1674.
- Kaufmann, M., P. Linder, A. Honegger, K. Blank, M. Tschopp, G. Capitani, A. Pluckthun, and M.G. Grutter (2002). Crystal structure of the anti-His tag antibody 3D5 single-chain fragment complexed to its antigen. *J. Mol. Biol.* 318:135-147.

- Kim TD, Paik SR, Yang CH. (2002) Structural and functional implications of C-terminal regions of alpha-synuclein. *Biochemistry* 41(46):13782-90.
- Kimple ME, Sondek J (2004). Overview of affinity tags for protein purification. Curr. Protoc. Protein Sci. 2004 Sep;Chapter 9:Unit 9.9.
- 5 • Kohler and Milstein, Nature, 256, pp. 495-497 (1975).
- Korotkov KV, Pardon E, Steyaert J, Hol WG. Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. Structure. 2009, **17**, 255-265.
- Lauwereys, M., Arbabi Ghahroudi, M., Desmyter, A., Kinne, J., Hölzer, W., De Genst, E., Wyns, L. & 10 Muyldermans, S. (1998). Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *EMBO J.* **17**, 3512-3520
- Lefranc MP, Duprat E, Kaas Q, Tranne M, Thiriot A, Lefranc G. (2005) IMGT unique numbering for MHC groove G-DOMAIN and MHC superfamily (MhcSF) G-LIKE-DOMAIN. Dev Comp Immunol. 29, 917-38.
- 15 • Li Y, (2010). Commonly used tag combinations for tandem affinity purification. Biotechnol Appl Biochem 55, 73-83.
- Lichty J.J. et al., Protein Expression and Purification, 41, May 2005, Pages 98-105.
- Loris R, Marianovsky I, Lah J, Laeremans T, Engelberg-Kulka H, Glaser G, Muyldermans S, Wyns L. Crystal structure of the intrinsically flexible addiction antidote MazE. *J Biol Chem.* 2003 **278**, 28252-20 28257.
- Maniatis et al. (1982) Molecular Cloning: A Laboratory Manual; DNA cloning: A practical approach, volumes I and II (1985) (Glover, ed.).
- Metzger et al. (1988) Nature 334:31-36.
- Park SM, Jung HY, Kim TD, Park JH, Yang CH, Kim J. (2002) Distinct roles of the N-terminal-binding 25 domain and the C-terminal-solubilizing domain of alpha-synuclein, a molecular chaperone. *J Biol Chem.* 277(32), 28512-20.
- Perkins A., Cochrane A.W., Ruben S.M., Rosen C.A.(1989). Structural and functional characterization of the human immunodeficiency virus rev protein. J Acquir Immune Defic Syndr. **2**, 256-263.
- 30 • Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E, Bragado-Nilsson E, Wilm M, Séraphin B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods. 24, 218-29.

- Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann, and B. Seraphin (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17:1030-1032.
- Rivers, R.C., Kumita, J.R., Tartaglia, G.G., Dedmon, M.M., Pawar, A., Vendruscolo, M. & Dobson, C.M. Christodoulou, J. (2008). Molecular determinants of the aggregation behavior of alpha- and beta-synuclein. *Protein Sci.* 17, 887-98.
- Robinson LS, Ashman EM, Hultgren SJ, Chapman MR. Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. *Mol Microbiol.* 2006, 59:870-81.
- Saerens D, Ghassabeh GH, Muyldermans S (2008). Single-domain antibodies as building blocks for novel therapeutics. *Curr Opin Pharmacol.* 8, 600-8.
- Sambrook et al. *Molecular Cloning, a Laboratory Manual*, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 16.30-16. 55 (1989)
- Schmidt, T.G.M. and A. Skerra (2007). The Streptag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protocols* 2:1528-1535.
- Smith, M. C, Furman, T.C., Ingolia, T.D., Pidgeon, C. (1988) *J. Biol. Chem.* 263, 7211-7215
- Tucker, J. and R. Grisshammer. 1996. Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem. J.* 317:891-899.
- Uhlén, M., B. Nilsson, B. Guss, M. Lindberg, S. Gatenbeck, and L. Philipson. 1983. Gene fusion vectors based on the gene for staphylococcal protein A. *Gene* 23:369-378.
- Van Leene, J., *et al.* (2008) Boosting tandem affinity purification of plant protein complexes. *Trends in plant science* 13, 517-520
- Van Leene, J., *et al.* (2010) Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*. *Mol Syst Biol* 6, 397
- Vuchelen, A., O'Day, E., De Genst, E., Pardon, E., Wyns, L., Dumoulin, M., Dobson, C.M., Christodoulou J. & Hsu, S.T. (2009). (1)H, (13)C and (15)N assignments of a camelid nanobody directed against human alpha-synuclein. *Biomol NMR Assign.* 3, 231-3.
- Watson, J. D., et al., *Recombinant DNA*, 2nd Ed., New York: W. H. Freeman and Co. (1992)
- Waugh, D. (2005) Making the most of affinity tags. *Trends in Biotechnology* 23, 316-320
- Winnacker, E., *From Genes to Clones*, New York: VCH Publishers (1987)
- Wu (ed) *Molecular Cloning*, Cold Spring Harbor Laboratory, Plainview, N.Y. (1993)
- Y-B. Zhang et al. (2004) Protein aggregation during overexpression limited by peptide extensions with large net negative charge, *Protein Expr. Purif.* 36, 207–216.
- Zell, R & Fritz H. J. (1987). DNA mismatch-repair in *Escherichia coli* counteracting the hydrolytic deamination of 5-methyl-cytosine residues. *EMBO J.* 6, 1809–1815.

81584211

SEQUENCE LISTING IN ELECTRONIC FORM

In accordance with Section 111(1) of the Patent Rules, this description contains a sequence listing in electronic form in ASCII text format (file: 29775-121 Seq 04-11-12 v1.txt).

A copy of the sequence listing in electronic form is available from the Canadian Intellectual Property Office.

CLAIMS:

1. A fusion protein comprising a target protein linked to a heterologous C-terminal epitope tag, wherein said C-terminal epitope tag is defined by SEQ ID NO: 1, and wherein said SEQ ID NO: 1 represents the last four C-terminal amino acid residues of
5 said fusion protein, provided the target protein does not comprise fragments of alphasynuclein.
2. The fusion protein according to claim 1 wherein the C-terminal epitope tag is directly coupled to the C-terminal end of the target protein or via a linker.
3. An isolated nucleic acid encoding the fusion protein according to claim 1 or 2.
- 10 4. A vector comprising the nucleic acid according to claim 3.
5. A host cell comprising the vector according to claim 4 or the nucleic acid according to claim 3.
6. A complex of
 - (i) an antibody fragment or single-domain antibody specifically binding to a C-terminal
15 epitope tag defined by SEQ ID NO: 1; and
 - (ii) the fusion protein according to claim 1 or 2; and
 - (iii) optionally a ligand bound to the target protein comprised in said fusion protein.
7. The complex according to claim 6 wherein said complex is crystalline.
8. A solid surface comprising an antibody fragment or single-domain antibody
20 specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1.
9. Use of an antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1 and/or the complex according to claim 6 or 7 for affinity chromatography, affinity purification, immunoprecipitation, in-vivo imaging, tandem affinity purification, protein detection, immunochemistry,
25 surface-display, FRET-type applications and/or co-crystallisation.

10. Use according to claim 9 wherein said protein detection is performed in an ELISA detection assay or Western blot analysis.
11. Use according to claim 9 wherein said affinity purification is performed on a column or on beads.
- 5 12. A method for purifying a fusion protein comprising the steps of
- a) applying a solution containing the fusion protein according to claim 1 or 2 to a solid support possessing an immobilized antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1, and
- b) forming a complex between said immobilized antibody fragment or single-domain
10 antibody and said fusion protein,
- c) removing weakly bound or unbound molecules, and
- d) eluting the bound molecules.
13. A method for immobilizing and/or capturing a fusion protein on a solid support comprising the steps of
- 15 a) applying a solution containing the fusion protein according to claim 1 or 2 to a solid support possessing an immobilized antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1,
- b) forming a complex between said immobilized antibody fragment or single-domain antibody and said fusion protein, and
- 20 c) removing weakly bound or unbound molecules.
14. A method for detecting the fusion protein according to claim 1 or 2 in a sample or composition comprising using an antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1.

15. A kit comprising the fusion protein according to claim 1 or 2, and/or the nucleic acid according to claim 3, and/or the vector according to claim 4, and/or an antibody fragment or single-domain antibody specifically binding to a C-terminal epitope tag defined by SEQ ID NO: 1, said antibody fragment or single-domain antibody being
5 pre-immobilized to a non-soluble matrix or solid support, and/or the complex according to claim 6 or 7, and instructions for using said kit.
16. Use of the kit according to claim 15 for affinity chromatography, affinity purification, immunoprecipitation, in-vivo imaging, tandem affinity purification, protein detection, immunochemistry, surface-display, FRET-type applications and/or
10 co-crystallisation.
17. The fusion protein according to claim 1 or 2 or the complex according to claim 6 or 7, for use in affinity chromatography, affinity purification, immunoprecipitation, in-vivo imaging, tandem affinity purification, protein detection, immunochemistry, surface-display, FRET-type applications and/or co-crystallisation.

Figure 1

		--CDR1--	--CDR2---
Nb_85	qvqlqesgggsvqaggsrlrlscaas	gltysnyc mgwfrqapgkarevvar	istr-gikt yyadsvngrf
Nb_88	qvqlqesggglvqpggslrlscaas	gytysric mgwfrqapgkarevvar	istr-gikt yyadsvngrf
Nb_Syn2	qgqlvesgggsvqaggsrlrlscaas	gidsssysc mgwfrqrpqkeregvar	inglggvkt ayadsvkdrf
Nb_Syn2 PstI	qvqlqesgggsvqaggsrlrlscaas	gidsssysc mgwfrqrpqkeregvar	inglggvkt ayadsvkdrf
Nb_Syn1a	evqlvesgggsvqaggsrlslacaps	glnassyc mgwfrqspgkeregvar	ingnagikt ayadsvkdrf
Nb_Syn1b	dvqlvesgggsvqaggsrlrlscaps	glnassyc mgwfrqspgkeregvar	ingnagikt ayadsvkdrf
Nb_Syn1c	qvqlvesgggsvqaggsrlrlscapl	glnassyc mgwfrqspgkeregvar	ingnagikt ayadsvkdrf
Nb_86	qvqlqesgggsvqaggsrlrlscals	gytfrgnr mawfrqapgkeregvar	intg-gvnt yvadsvkgrf
Nb_87	qvqlqesgggsvqtggslrlscvas	gys--gy mawfrqapgkeregiaa	iyrg-dkit yyahsvqgrf
		-----CDR3-----	
Nb_85	tisrdnaknmvylqmnsllkpedtaiiyc	aaviypgygdscpwttsvny	wgqgtqvtvss
Nb_88	tisrdnaknmvylqmnsllkpedtaiiyc	aaviypgygdscpwttsvny	wgqgtqvtvss
Nb_Syn2	tisrdnaentvylqmnsllkpedtaiiyc	aakfspgyc--ggswsnfgy	wgqgtqvtvss
Nb_Syn2 qvq	tisrdnaentvylqmnsllkpedtaiiyc	aakfspgyc--ggswsnfgy	wgqgtqvtvss
Nb_Syn1a	tlstrdnakntvylqmnsllkaedsaiyfc	aaksspgyc--ggnwdnfgy	wgqgtqvtvss
Nb_Syn1b	tlstrdnakntvylqmnsllkaedsamiyc	aaksspgyc--ggnwdnfgy	wgqgtqvtvss
Nb_Syn1c	tlstrdnakntvylqmnsllkaedsamiyc	aaksspgyc--ggnwdnfgy	wgqgtqvtvss
Nb_86	tisqdnakntvylqmnnlqpedtamiyfc	aadltgwrp---vgfsgyny	wgqgtqvtvss
Nb_87	tisqanakntvylmnsllkpedtaiiyc	aarrvvadp--sllsktyay	wgqgtqvtvss

Figure 2

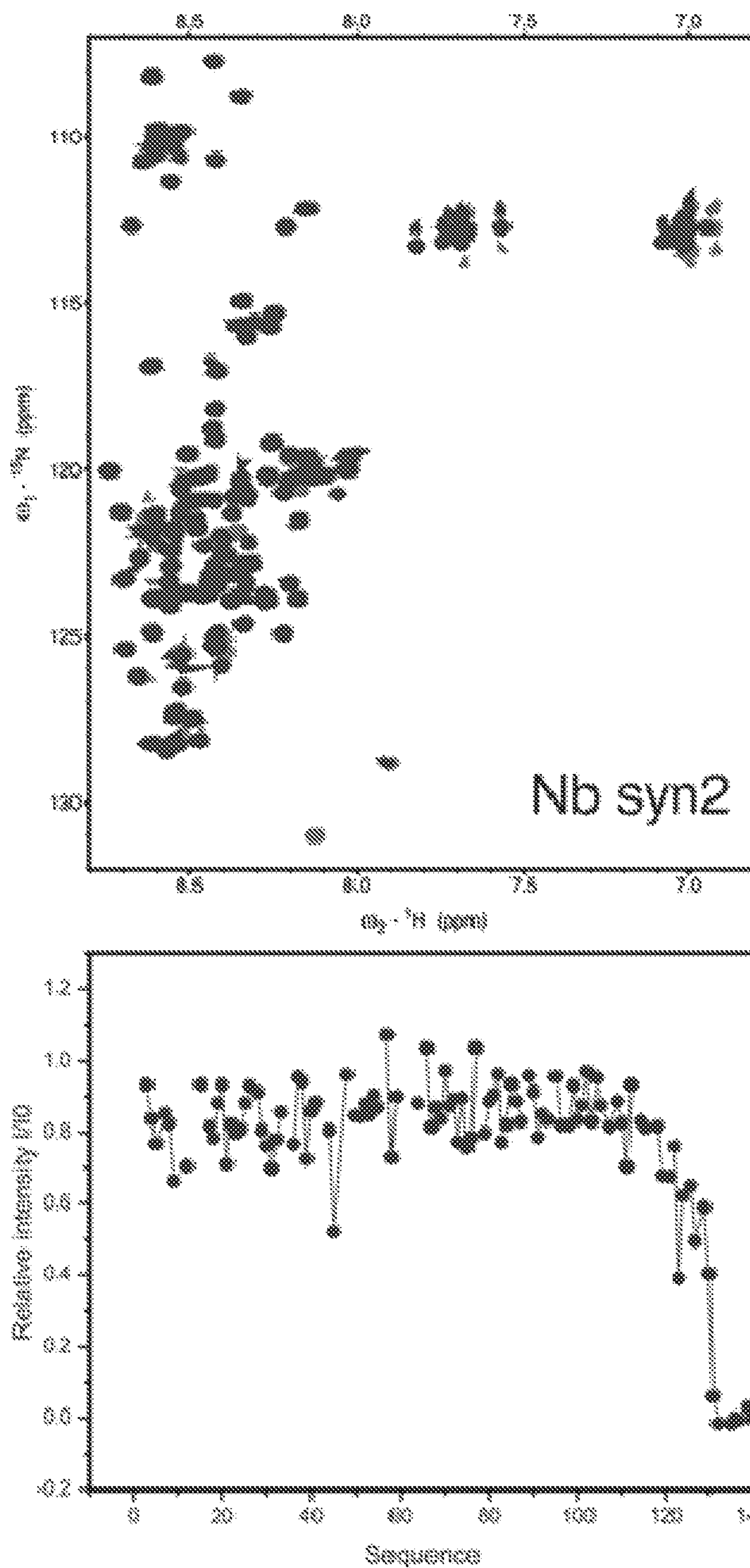


Figure 2 continued

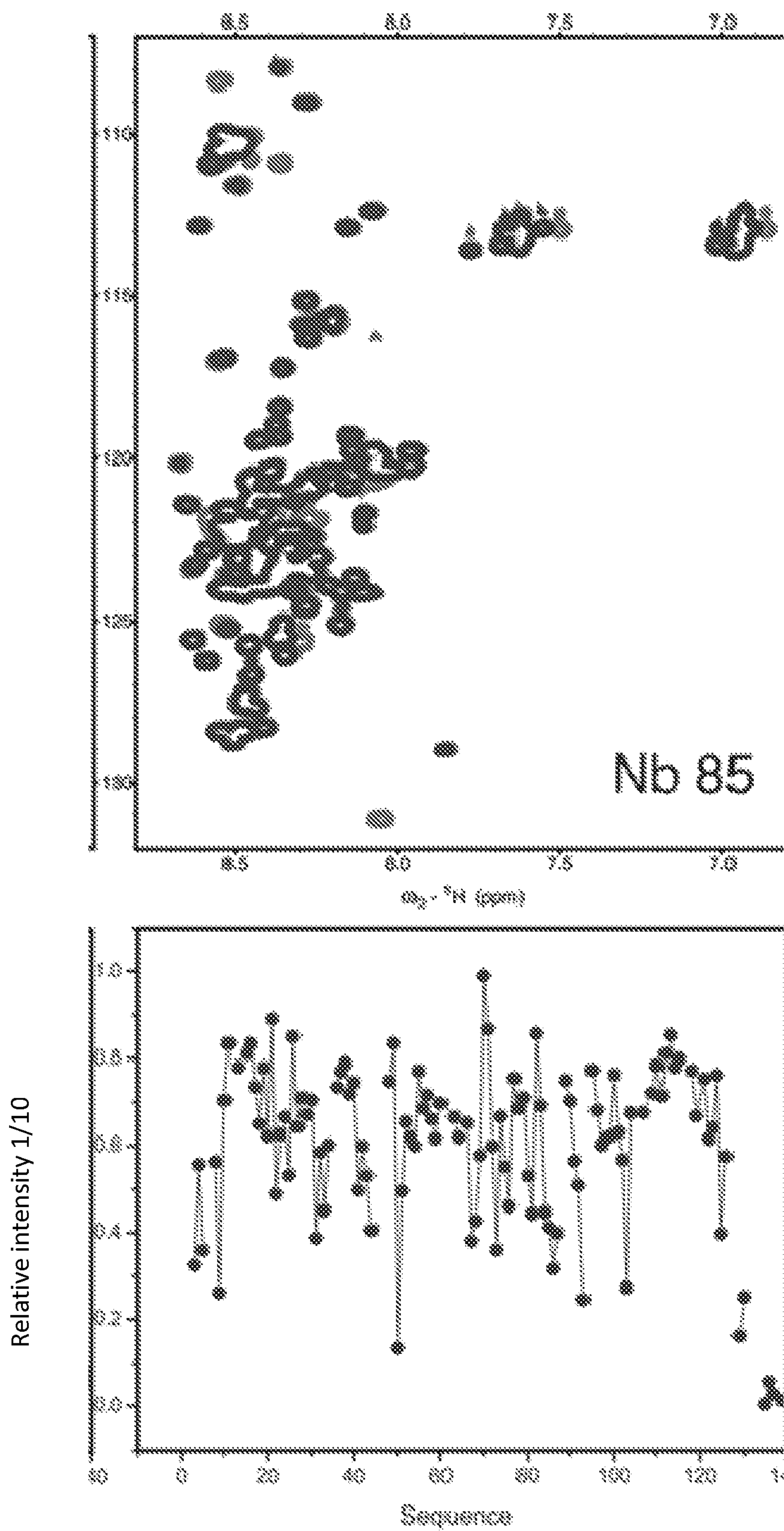


Figure 2 continued

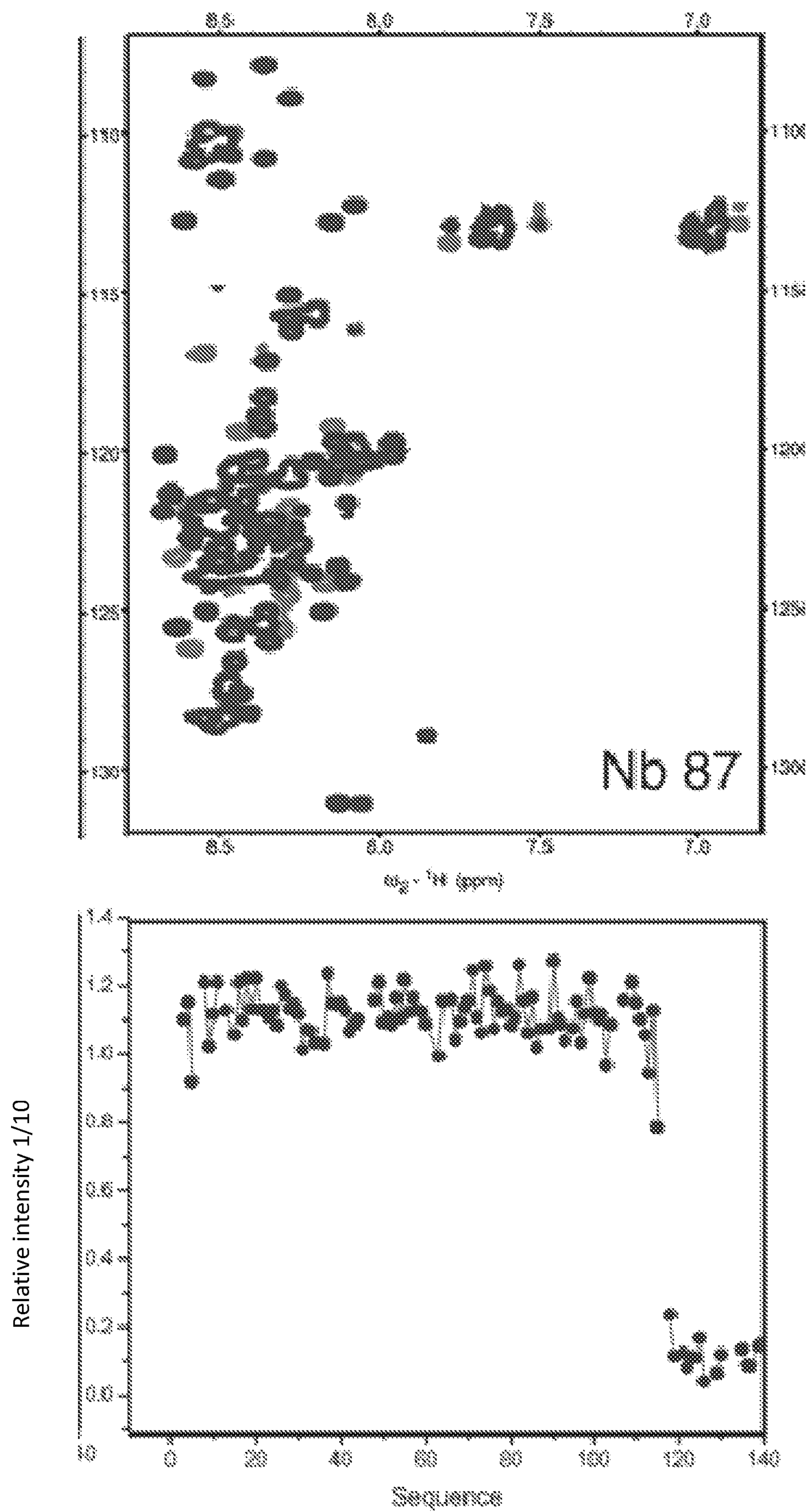


Figure 3

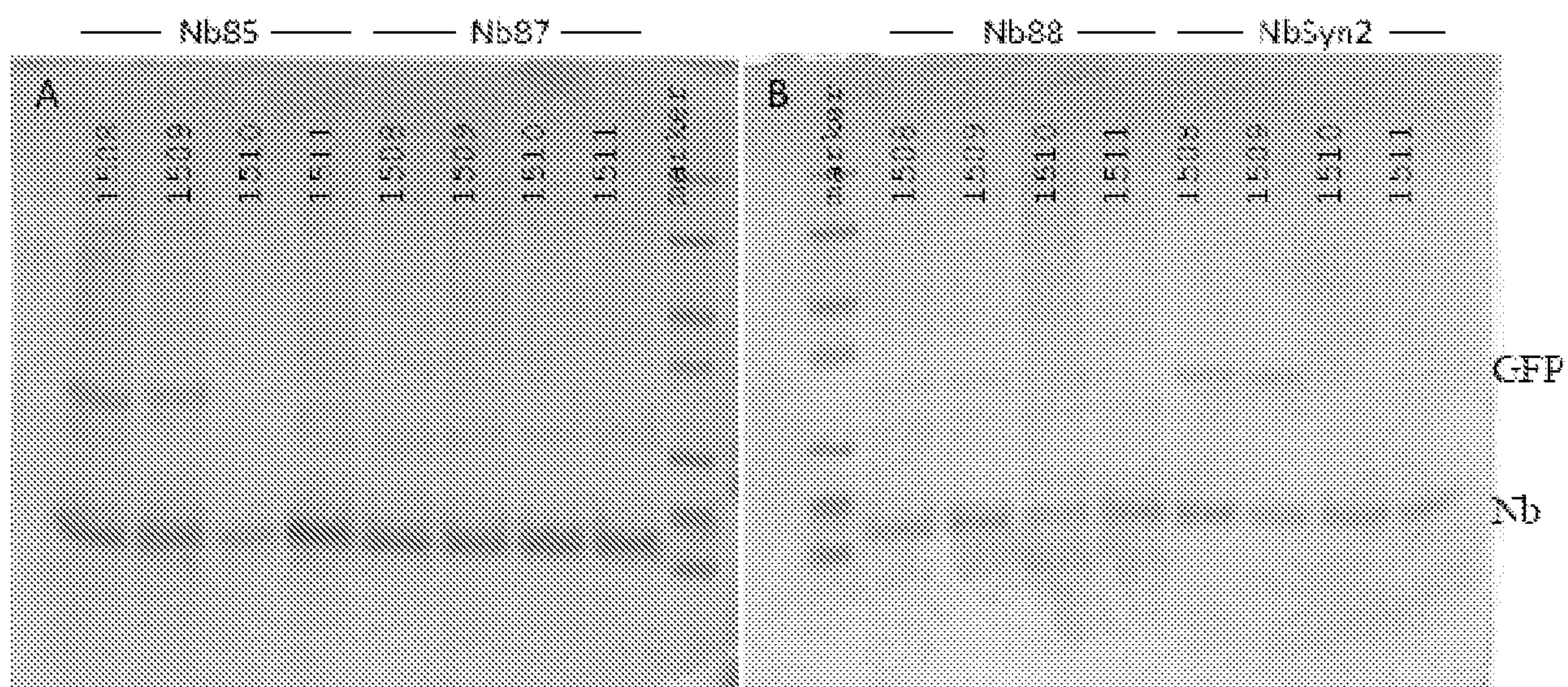


Figure 4

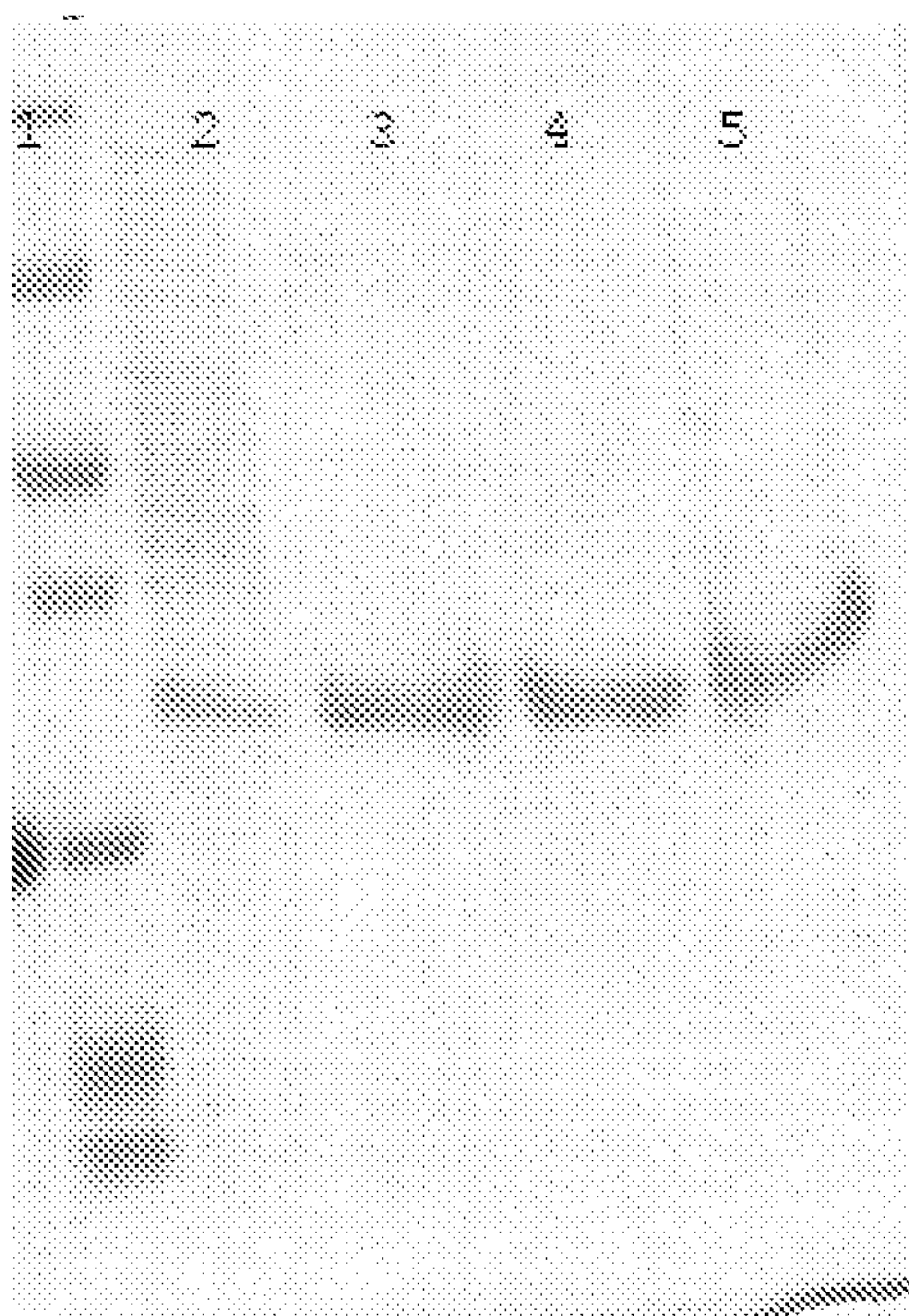


Figure 5

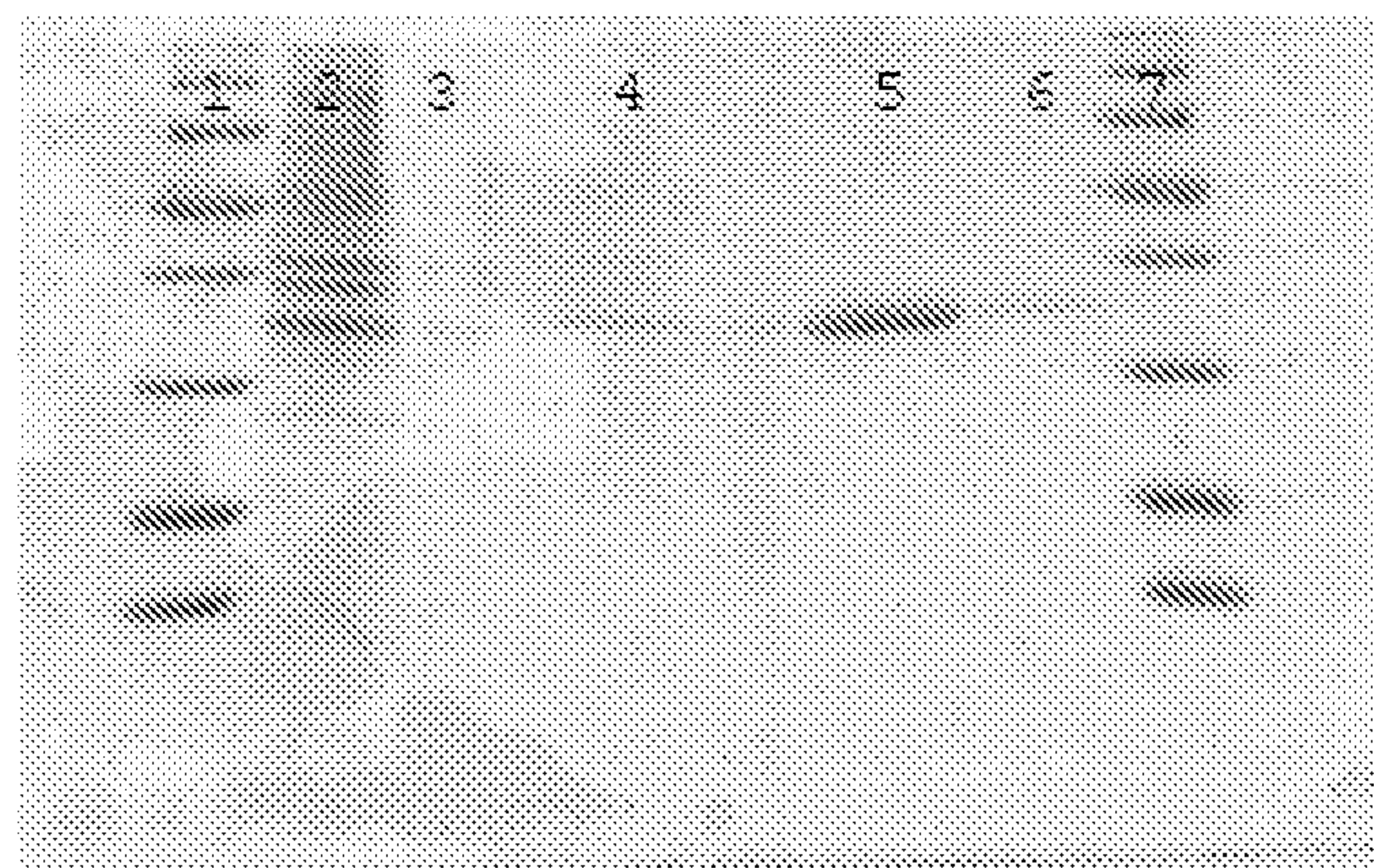


Figure 6

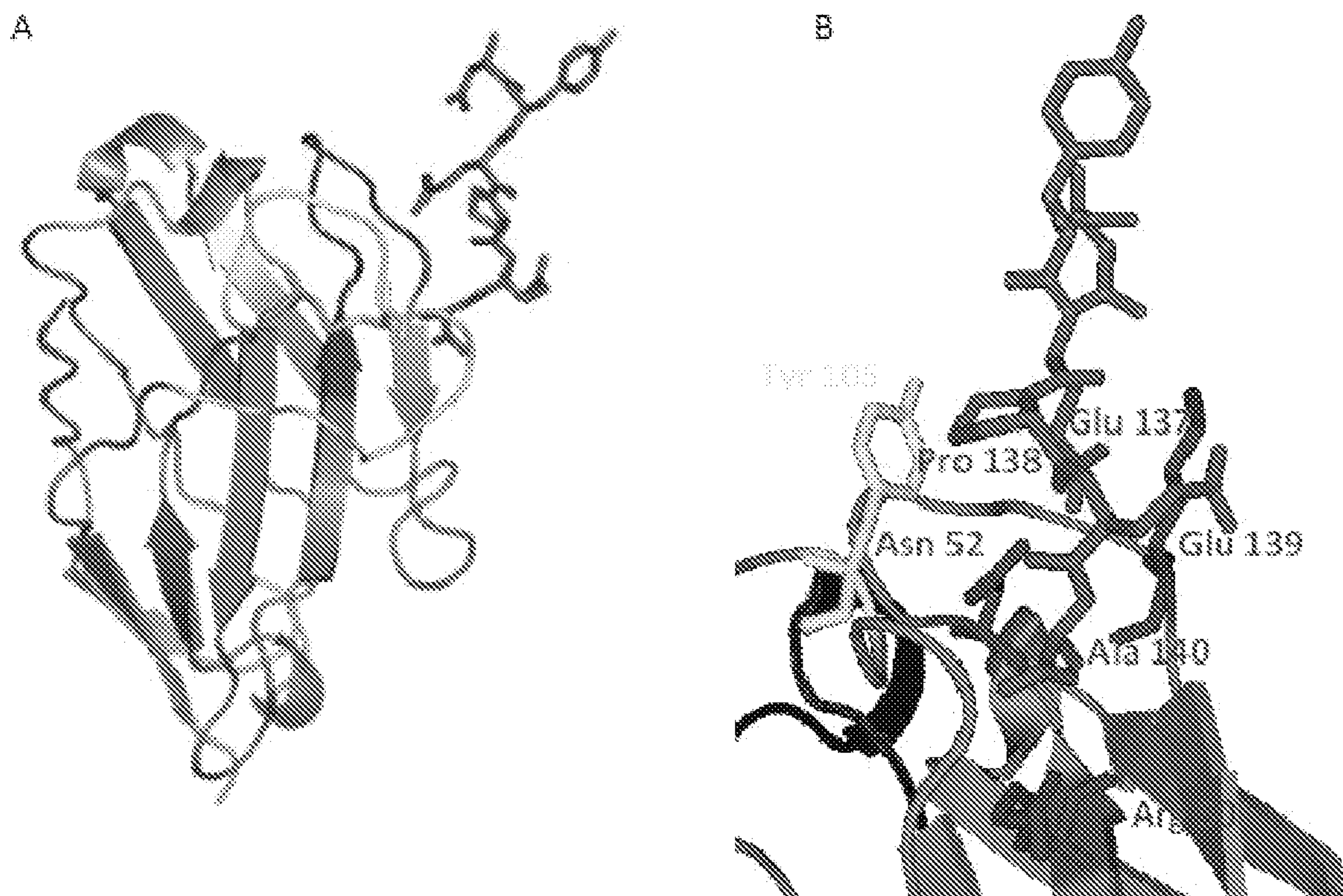


Figure 7

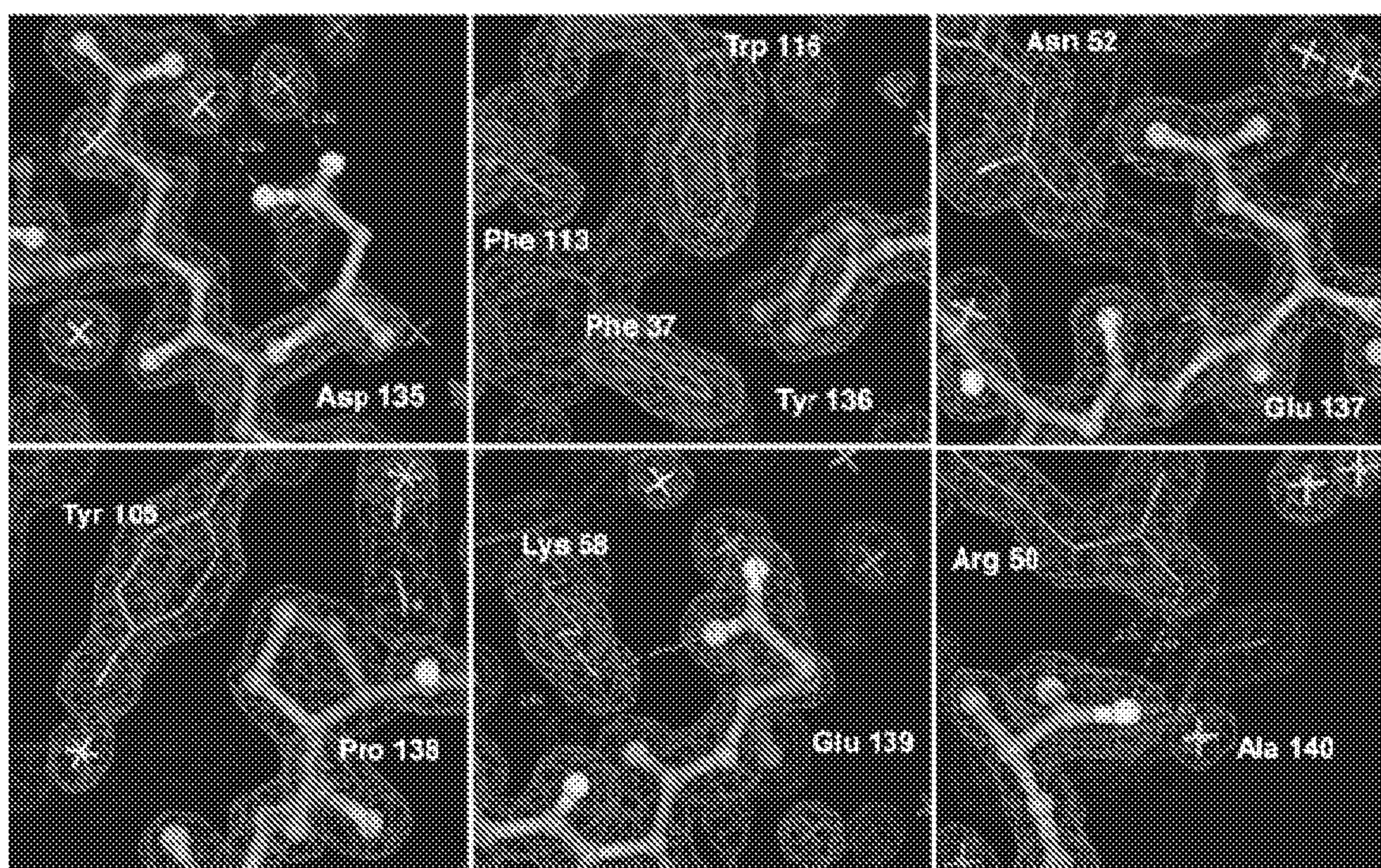


Figure 8



Figure 9

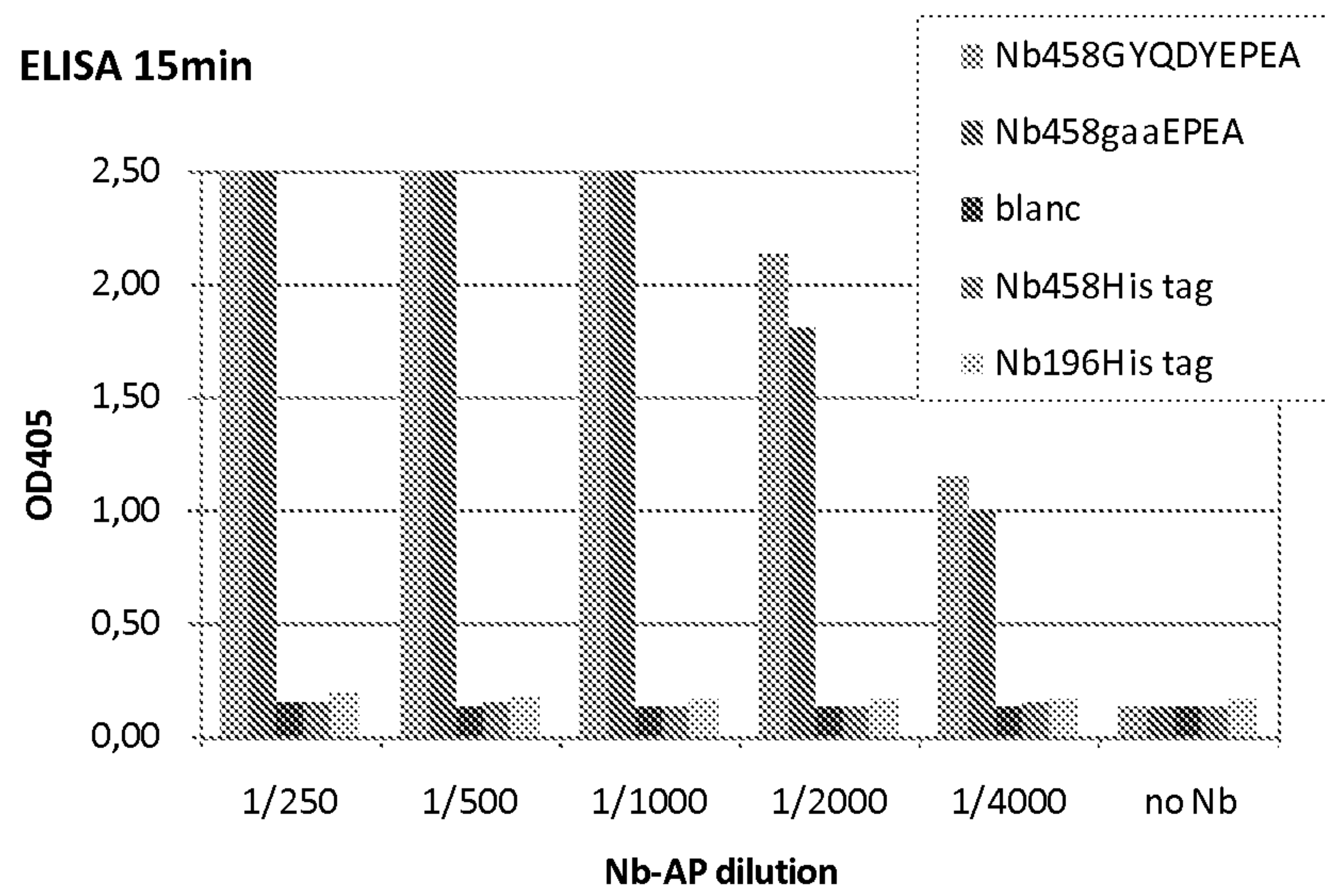


Figure 10

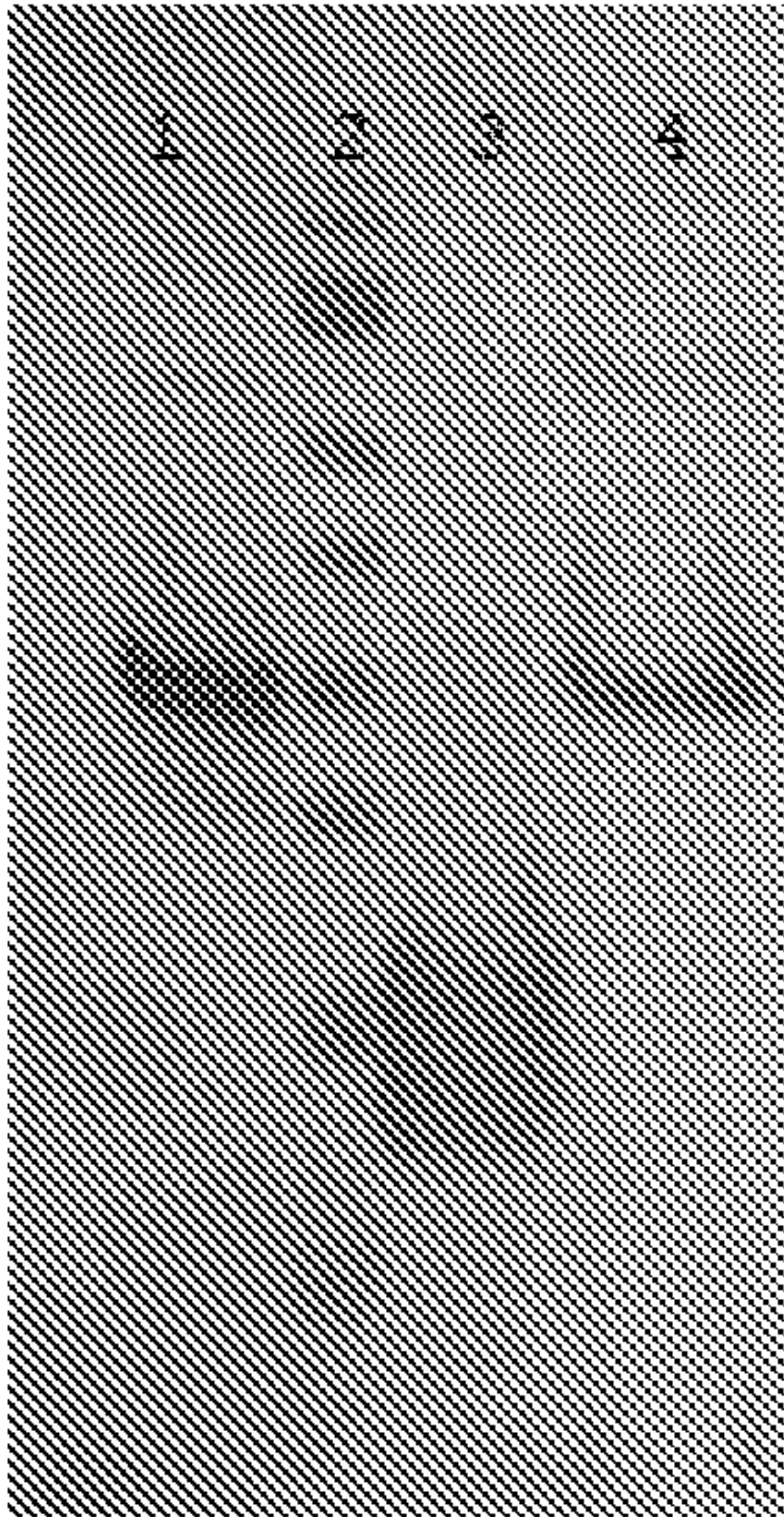


Figure 11

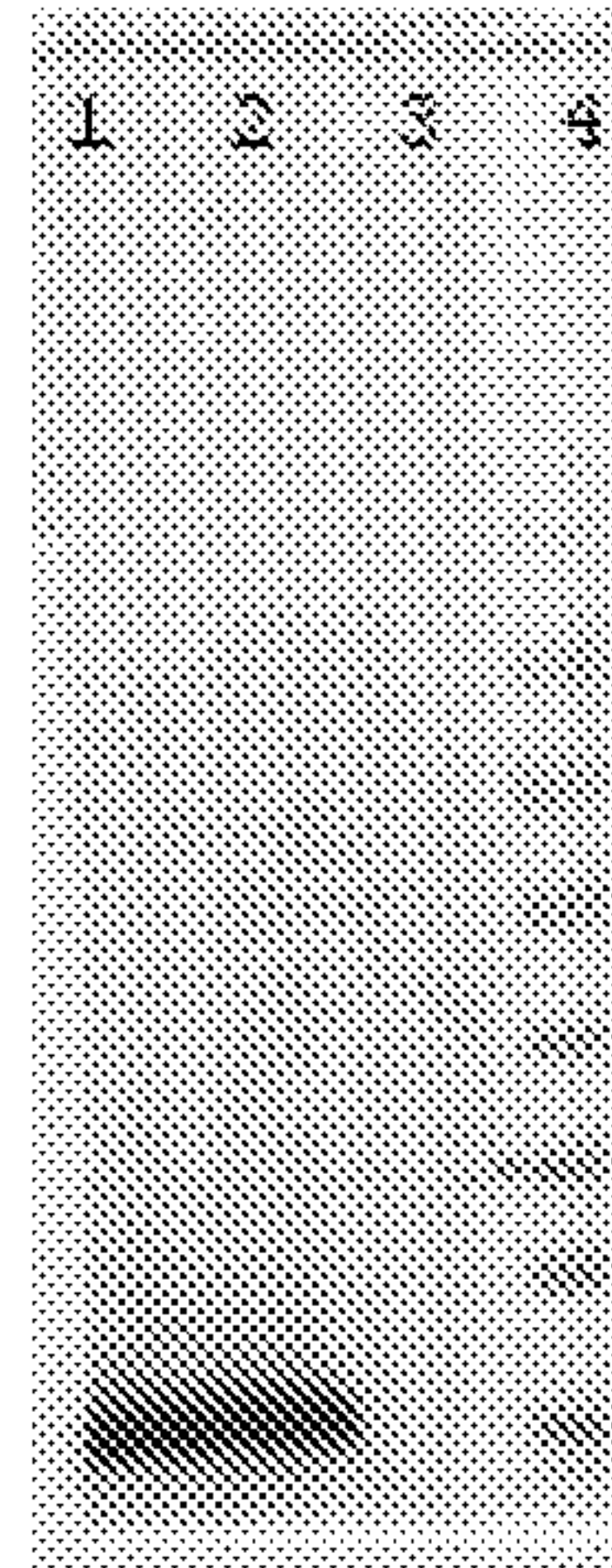


Figure 12

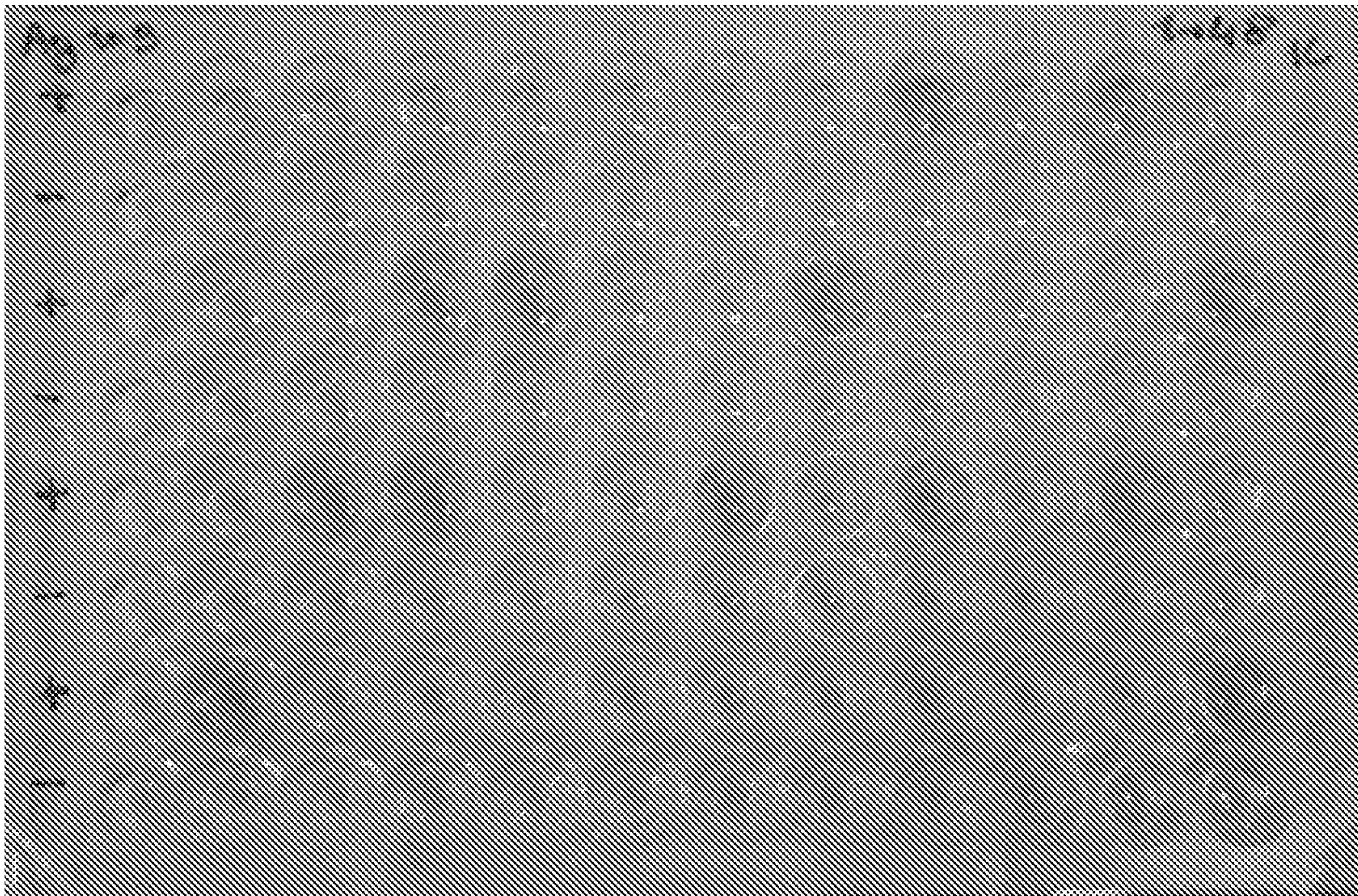


Figure 13

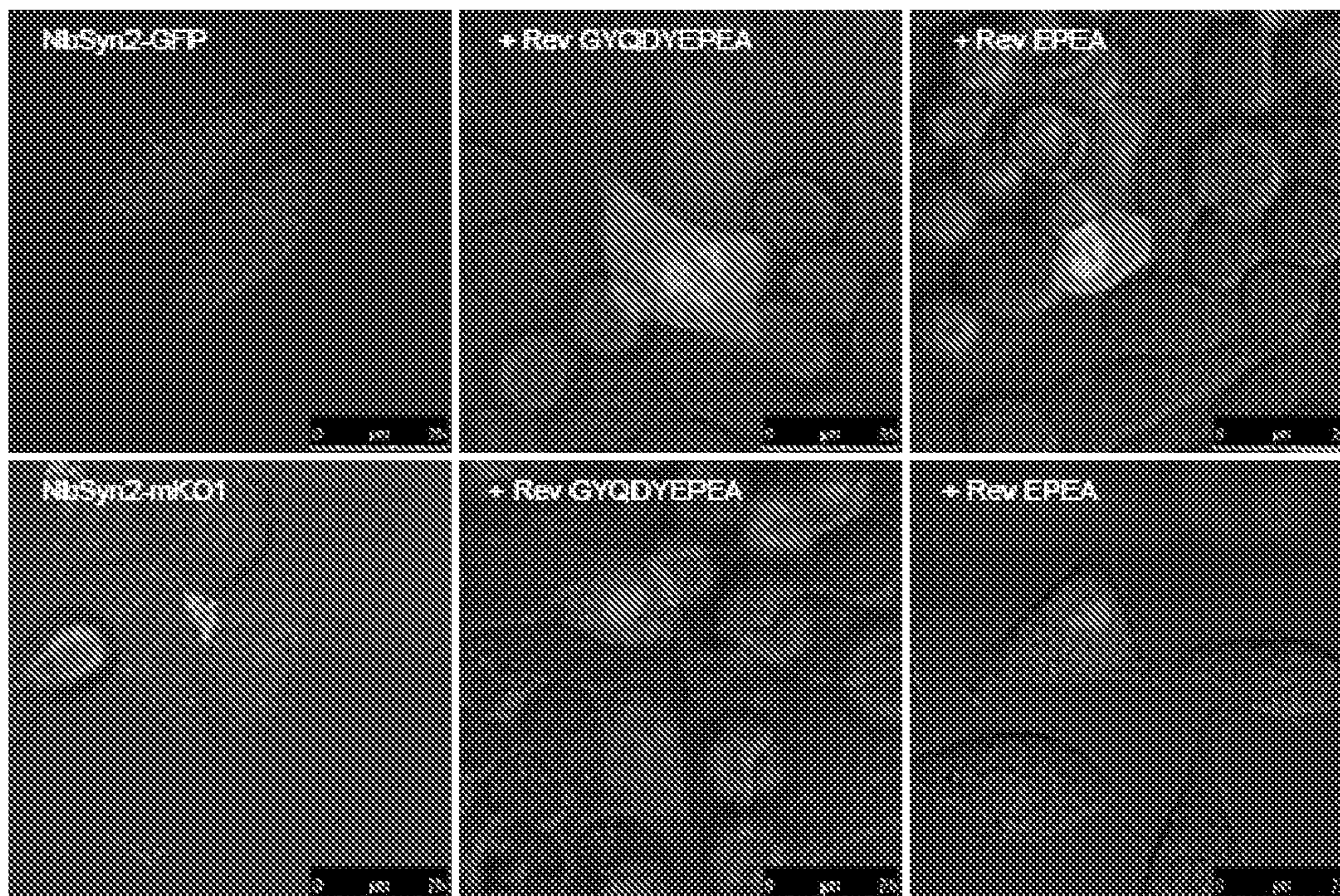


Figure 14

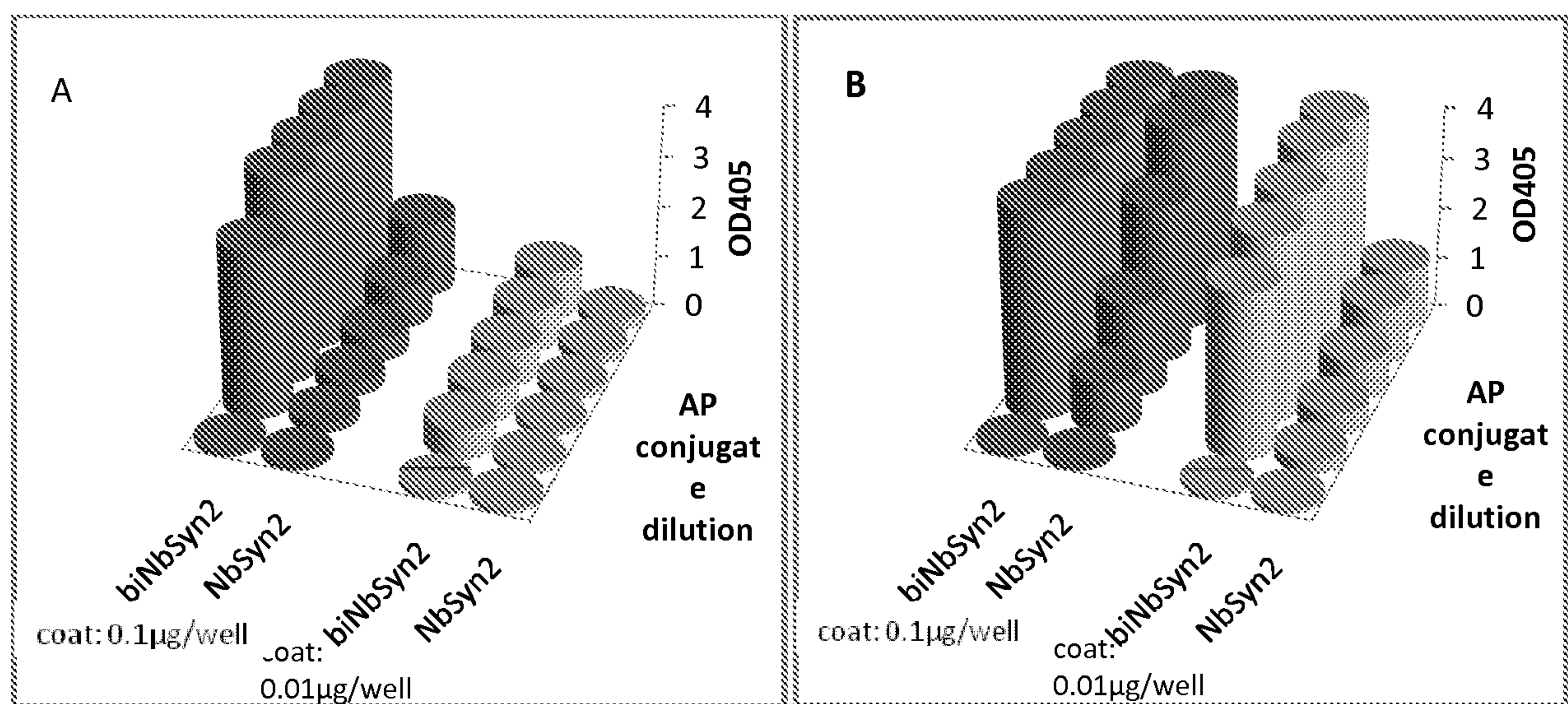


Figure 15

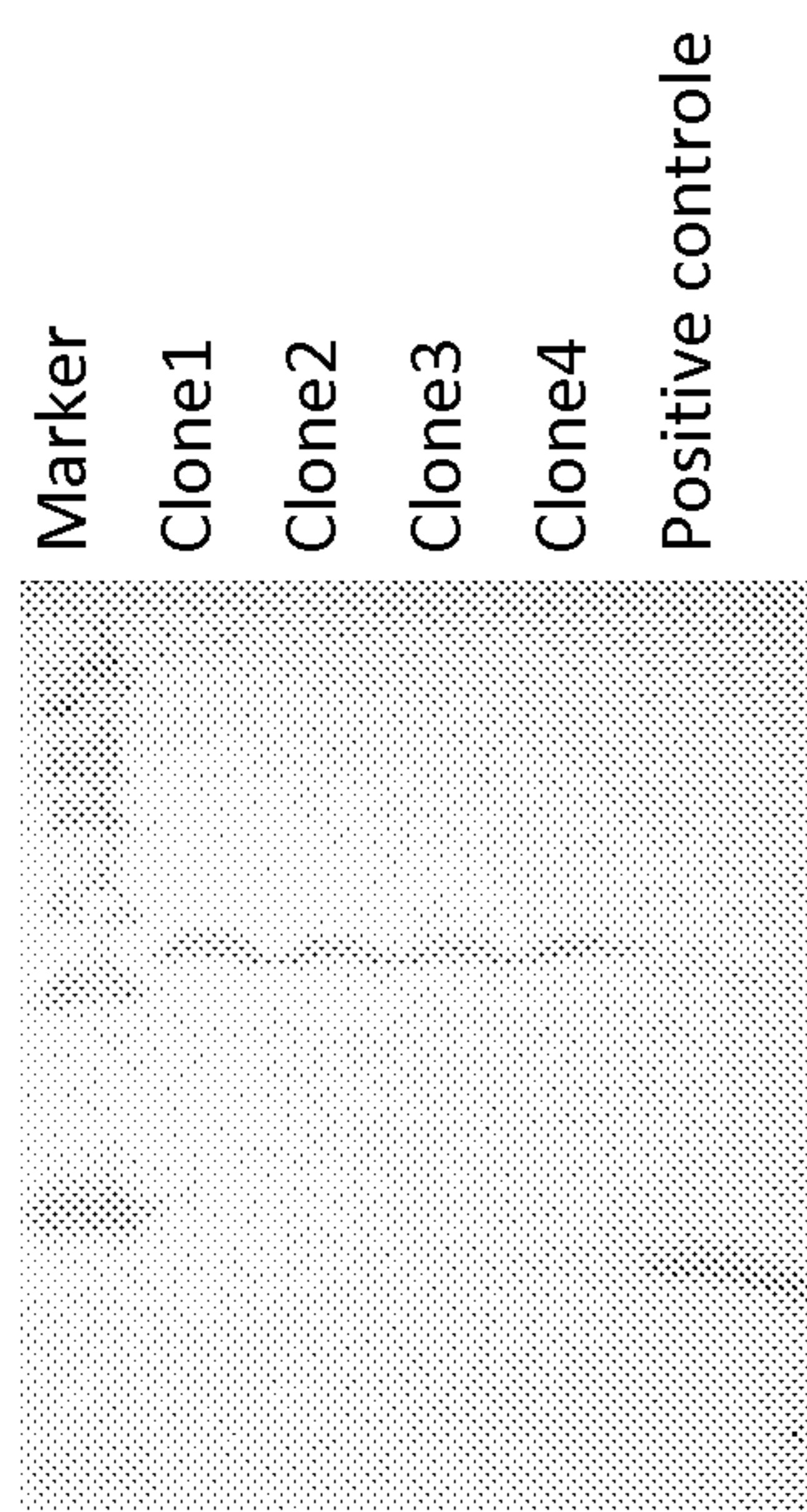


Figure 16

