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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 March 2015 (26.03.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/042464 A1

(51) International Patent Classification:
C07K 1/13 (2006.01)

(21) International Application Number:
PCT/US2014/056651

(22) International Filing Date:
19 September 2014 (19.09.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/880,012 19 September 2013 (19.09.2013) US

(71) Applicant: **UNIVERSITY OF WASHINGTON THROUGH ITS CENTER FOR COMMERCIALIZATION** [US/US]; 4311 11th Avenue NE, Suite 500, Seattle, WA 98105-4608 (US).

(72) Inventors: **BANEYX, Francois, C.**; 4311 11th Avenue NE, Suite 500, Seattle, WA 98105-4608 (US). **COYLE, Brandon, L.**; 4311 11th Avenue NE, Suite 500, Seattle, WA 98105-4608 (US).

(74) Agents: **MICHELETTI, Randy, R.** et al.; Perkins Coie LLP, P.O. Box 1247, Seattle, WA 98111-1247 (US).

(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, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: AFFINITY TAGS AND PROCESSES FOR PURIFYING AND IMMOBILIZING PROTEINS USING SAME

(57) Abstract: The present disclosure provides affinity tags, fusion proteins comprising one or more affinity tags, compositions comprising a fusion protein, methods of purifying a protein using an affinity tag, and devices for purifying a protein using an affinity tag.

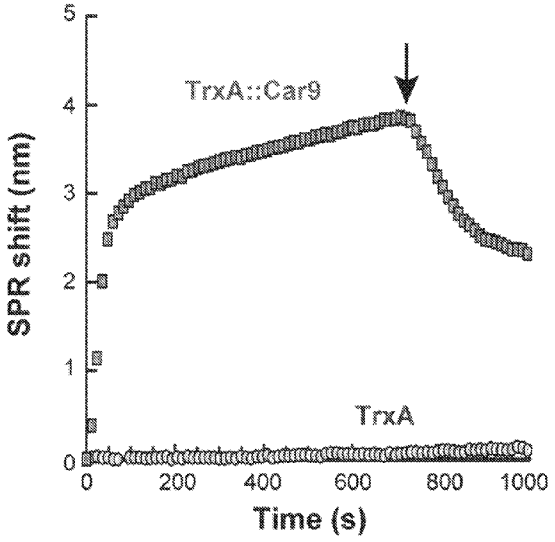


FIG. 2

WO 2015/042464 A1

Published:

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

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

AFFINITY TAGS AND PROCESSES FOR PURIFYING AND IMMOBILIZING PROTEINS USING SAME

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 61/880,012, filed September 19, 2013, the entire contents of which is incorporated herein by reference.

STATEMENT REGARDING FEDERALLY FUNDED RESEARCH

[0002] This invention was made with government support under grant no. DMR-0520567, awarded by the National Science Foundation, and grant no. N00014-12-1-1013, awarded by the Office of Naval Research. The government has certain rights in the invention.

TECHNICAL FIELD

[0003] The present disclosure relates to affinity tags, fusion proteins comprising one or more affinity tags, compositions comprising a fusion protein, methods of purifying or immobilizing a protein using an affinity tag, and devices for purifying a protein using an affinity tag.

BACKGROUND

[0004] The advent of disposable plasmid DNA miniprep kits has revolutionized molecular biology by enabling rapid and inexpensive plasmid purification. While affinity tags for protein purification have been developed, no currently available affinity tag enables rapid, inexpensive protein purification without requiring expensive, specialized reagents (e.g., immobilized binding substrates, immobilized antibodies, nickel nitrilotriacetate resins, etc.).

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] Many aspects of the present technology can be better understood with reference to the following drawings. The relative dimensions in the drawings may be to scale with respect to some embodiments. With respect to other embodiments, the drawings may not be to scale. For ease of reference, throughout this disclosure identical reference numbers may be used to identify identical or at least generally similar or analogous components or features.

[0006] FIG. 1 is a representation of common surface chemistries of silica.

[0007] FIG. 2 is a plot showing a surface plasmon resonance ("SPR") analysis of adsorption of TrxA::Car9 onto a silica-coated SPR chip compared to TrxA over a 12-minute adsorption phase (1 μM reagent concentration at 50 μL/min flow rate), followed by a 10-minute buffer-only wash cycle (arrow).

[0008] FIGS. 3A and 3B show centrifuge tubes containing hydrated silica gel (60-220 μm) after one hour of incubation at room temperature with 150 μL (2.2 μM) of GFPmut2 (FIG. 3A) or GFPmut2-Car9 (FIG. 3B). FIG. 3C is a plot of free GFP protein in the supernatant fluids of FIGS. 3A-3B as a function of the amount of silica gel added.

[0009] FIG. 4 is a plot of Langmuir absorption isotherm constructed by plotting the SPR shift in nanometers induced by the binding of GFPmut2-Car9 (■) and GFPmut2 (●) on silica as a function of the protein concentration. Solid lines indicate theoretical fits using Langmuir adsorption kinetics.

[0010] FIGS. 5A and 5B show results of purification of GFPmut2-Car9 (FIG. 5A) and MBP-Car9 (FIG. 5B) by silica gel chromatography and 0.5M arginine elution. Lanes show whole cell lysates (WC), clarified lysates (L) flow through (FT) and eluted fractions (E). Arrows indicate purified protein.

[0011] FIG. 6 shows results of purification of GFPmut2-Car9, MBP-Car9 (labelled "MBP" in FIG. 6) and mCherry-Car9 (labelled "mCherry" in FIG. 6) by silica gel chromatography and 1M lysine elution. Lanes show clarified lysate (L) and eluted fractions (E). Arrows indicate the position of purified protein.

[0012] FIG. 7A illustrates a partially schematic view of a device for purifying a protein configured according to one embodiment of the present technology.

[0013] FIG. 7B shows results of purification of GFPmut2-Car9 with 1M L-lysine eluent using a device configured according to one embodiment of the present technology. Lanes indicate clarified lysate (L) and eluted fractions (E) after 50-fold concentration.

[0014] FIGS. 8A and 8B show results of purification of GFPmut2-Car9 from KS272 cells and MBP-Car9 from Top10 cells, respectively, using 3 g of silica gel and 0.5M arginine eluent after sonication and incubation of the lysates on ice for 30 minutes in the presence of OmpT. Lanes show whole cell (WC), clarified lysates after 30 minutes of incubation (L), flow through (FT), and eluted fractions (E). In each of FIGS. 8A and 8B, the top arrow indicates the

migration position of intact GFPmut2-Car9 and MBP-Car9, respectively, while the bottom arrows indicate the migration position of the degradation products.

[0015] FIGS. 9A and 9B show the results of incubating purified GFPmut2-Car9 (Control) with whole cells lacking (*ΔompT*) or overproducing (*ompT*$^{+}$) the outer membrane protease OmpT. Soluble proteins were subjected to SDS-PAGE analysis after 30 minutes (FIG. 9A) or 45 minutes (FIG. 9B) of incubation at room temperature and removal of the cells.

[0016] FIG. 10A shows a silica slide after incubation for one minute with 15 μL of a 10 μM solution of GFPmut2 (left region), GFPmut2-Car9 (middle region) and GFPmut2-Car15 (right region). FIG. 10B shows the slide of FIG. 10A after washing for 5 minutes with ddH_2O. FIG. 10C shows the slide of FIG. 10B after washing for 5 minutes with 0.5M arginine in buffer. In each of FIGS. 10A-10C, the slide was imaged under wet conditions using a Typhoon 9000 gel imaging scanner with excitation at 488 nm.

[0017] FIG. 11 shows results of a cellular adherence experiment indicating that Car8, Car9, Car12 and Car 36 all bind to silica and can be removed upon treatment with 1M L-lysine solution in 20 mM Tris buffer at pH 7.5.

[0018] FIG. 12A is an image of GFPmut2-Car9 immobilized onto and within a silica sol-gel after storage at 4°C for 2 months. FIG. 12B is an image of the sol-gel of FIG. 12A after incubation for 48 hours at 4°C with 0.5M arginine.

DETAILED DESCRIPTION

[0019] The present technology is generally directed to affinity tags for proteins, proteins comprising an affinity tag, devices configured to purify proteins using affinity tags, and methods of purifying or immobilizing a protein using an affinity tag. In some embodiments, the affinity tag is a silica binding peptide. In some embodiments, the silica binding peptide comprises, consists essentially of, or consists of a peptide of SEQ ID NOs.: 1 to 4. Systems and methods configured in accordance with embodiments of the present technology provide efficient, inexpensive methods of purifying or immobilizing a protein using an affinity tag as disclosed herein.

Selected Embodiments of Affinity Tags and Proteins Comprising Same

[0020] The present disclosure provides various affinity tags. Unlike affinity tags known in the art, affinity tags of the present disclosure bind to inexpensive substrates, such as silicon

oxide substrates, rather than to antibodies or immobilized metal ions such as nickel nitrilotriacetate. In some embodiments, the silicon oxide substrate comprises, consists essentially of, or consists of silica (e.g., silica gel). In some embodiments, the silicon oxide substrate comprises, consists essentially of, or consists of borosilicate, controlled pore glass, quartz, and/or oxidized silicon.

[0021] Silicon oxide substrates, as shown representatively in FIG. 1, include several different types of surface silanols and siloxanes. Affinity tags of the present disclosure are configured to releasably bind to the surface of silicon oxide substrates. In some embodiments, the affinity tag has an amino acid sequence of DSARGFKKPGKR (SEQ ID NO.: 1; referred to herein as "Car9"). In other embodiments, the affinity tag has an amino acid sequence of KKRSPILASKRR (SEQ ID NO.: 2; referred to herein as "Car8"). In still other embodiments, the affinity tag has an amino acid sequence of RDRGATYPKLGR (SEQ ID NO.: 3; referred to herein as "Car12"). In other embodiments, the affinity tag has an amino acid sequence of RNKRCSSKTRRG (SEQ ID NO.: 4; referred to herein as "Car36"). In other embodiments, the affinity tag has an amino acid sequence of RTYLPLPWMAAL (SEQ ID NO.: 5; referred to herein as "Car15").

[0022] In some embodiments, the affinity tag has an amino acid sequence generally homologous to any one of SEQ ID NOs.: 1 to 4, while still exhibiting an affinity for a silicon oxide substrate. In some embodiments, the affinity tag has an amino acid sequence having a homology of at least 70%, at least 80%, at least 90%, or at least 95% with the amino acid sequence of any one of SEQ ID NOs.: 1 to 4.

[0023] In one embodiment, the affinity tag has an amino acid sequence having a homology of at least 70% with the amino acid sequence of SEQ ID NO.: 1. In another embodiment, the affinity tag has an amino acid sequence having a homology of at least 80% with the amino acid sequence of SEQ ID NO.: 1; the affinity tag has an amino acid sequence having a homology of at least 90% with the amino acid sequence of SEQ ID NO.: 1. the affinity tag has an amino acid sequence having a homology of at least 95% with the amino acid sequence of SEQ ID NO.: 1. In some embodiments, the affinity tag has an amino acid sequence including a first arginine residue, a hydrophobic residue within two positions of the first arginine residue, two consecutive lysine residues within four positions of the first arginine residue, a third lysine residue within 7 positions of the first arginine residue, and a second arginine residue within 8 positions of the

first arginine residue (e.g., *XXX*R*X*H_yKK*XX*KR, wherein *X* is any amino acid residue and H_y is a hydrophobic amino acid residue).

[0024] In one embodiment, the affinity tag has an amino acid sequence having a homology of at least 70% with the amino acid sequence of SEQ ID NO.: 2. In another embodiment, the affinity tag has an amino acid sequence having a homology of at least 80% with the amino acid sequence of SEQ ID NO.: 2; the affinity tag has an amino acid sequence having a homology of at least 90% with the amino acid sequence of SEQ ID NO.: 2; the affinity tag has an amino acid sequence having a homology of at least 95% with the amino acid sequence of SEQ ID NO.: 2. In some embodiments, the affinity tag has an amino acid sequence including a first lysine residue, a second lysine residue within 2 positions of the first lysine residue, a first arginine residue within 3 positions of the first lysine residue, at least one hydrophobic residue within 7 positions of the first lysine residue, a third lysine residue within 9 positions of the first lysine residue, a second arginine residue within 10 positions of the first lysine residue, and a third arginine residue within 11 positions of the first lysine residue (e.g., KKR*XX*$H_y$$H_y$*XX*KRR, wherein *X* is any amino acid residue and H_y is a hydrophobic amino acid residue).

[0025] In one embodiment, the affinity tag has an amino acid sequence having a homology of at least 70% with the amino acid sequence of SEQ ID NO.: 3. In another embodiment, the affinity tag has an amino acid sequence having a homology of at least 80% with the amino acid sequence of SEQ ID NO.: 3; the affinity tag has an amino acid sequence having a homology of at least 90% with the amino acid sequence of SEQ ID NO.: 3; the affinity tag has an amino acid sequence having a homology of at least 95% with the amino acid sequence of SEQ ID NO.: 3. In some embodiments, the affinity tag has an amino acid sequence including a first arginine residue, a second arginine residue within 3 positions of the first arginine residue, a hydrophobic residue within 6 positions of the first arginine residue, a lysine residue within 8 positions of the first arginine residue, and a third arginine residue within 11 positions of the first arginine residue (e.g., R*X*R*XXX*H_y*X*K*XX*R, wherein *X* is any amino acid residue and H_y is a hydrophobic amino acid residue).

[0026] In one embodiment, the affinity tag has an amino acid sequence having a homology of at least 70% with the amino acid sequence of SEQ ID NO.: 4. In another embodiment, the affinity tag has an amino acid sequence having a homology of at least 80% with the amino acid sequence of SEQ ID NO.: 4; the affinity tag has an amino acid sequence having a homology of at least 90% with the amino acid sequence of SEQ ID NO.: 4; the affinity tag has an amino acid

sequence having a homology of at least 95% with the amino acid sequence of SEQ ID NO.: 4. In some embodiments, the affinity tag has an amino acid sequence including a first arginine residue, a first lysine residue within 3 positions of the first arginine residue, a second arginine residue within 4 positions of the first arginine residue, a hydrophobic residue within 5 positions of the first arginine residue, a second lysine residue within 7 positions of the first arginine residue, a third arginine residue within 9 positions of the first arginine residue, and a fourth arginine residue within 10 positions of the first arginine residue (e.g., R*X*KRH$_y$*XX*K*X*RR*X*, wherein *X* is any amino acid residue and H_y is a hydrophobic amino acid residue).

[0027] The present disclosure provides proteins comprising an affinity tag as disclosed herein (e.g., any one of SEQ ID NOs.: 1 to 4). In some embodiments, the protein is a fusion protein comprising an affinity tag as disclosed herein and a peptide, such as a therapeutic peptide, a diagnostic peptide, or a peptide of other function. The peptide and the affinity tag may be linked through covalent bonding including, but not limited to, covalent organic bonds (e.g., C-C, C-N, C-O, etc.), disulfide bonding, hydrogen bonding, electrostatic bonding, recombinant fusion, and/or conformational bonding. In some embodiments, the peptide and the affinity tag are linked by one or more linking compounds. In some embodiments, the fusion protein exhibits the same or similar function as the peptide. In some embodiments, the peptide is a polyclonal antibody, a monoclonal antibody, a complement determining region-grafted antibody preparation, a hybrid antibody, an altered antibody, a F(ab)′$_2$ fragment, a Fab molecule, a Fv fragment, a single domain antibody, a chimeric antibody, or a fragment of any of the foregoing. In some embodiments, the peptide is green fluorescent protein ("GFP") or a variant thereof (e.g., GFPmut2). In some embodiments, the peptide is maltose-binding protein ("MBP") or a variant thereof. In some embodiments, the peptide is mCherry fluorescent protein ("mCherry") or a variant thereof.

[0028] Proteins of the present disclosure may further comprise a cleavage site, for example a protease cleavage site. In some embodiments, the affinity tag and the peptide of a fusion protein are separated by the cleavage site. In some embodiments, the cleavage site is a portion of the affinity tag. In other embodiments, the cleavage site is a portion of the peptide. In yet other embodiments, the cleavage site comprises, consists essentially of, or consists of an amino acid sequence that is not a portion of the affinity tag or the peptide. In some such embodiments, treatment of the fusion protein with a suitable protease cleaves the affinity tag, or a portion thereof, from the fusion protein to release an amino acid sequence comprising or consisting

essentially of the protein. One of skill in the art will recognize that, when a protease is used to cleave the fusion protein, the resulting cleaved peptide will contain at least one terminal amino acid residue in addition to the peptide. For example, a fusion protein of formula $(X)_n$RR–*peptide*, wherein $(X)_n$RR is the affinity tag having two consecutive arginine residues, and *peptide* is the peptide, may be cleaved in one embodiment by a protease that targets consecutive basic residues, such as OmpT. In such an embodiment, the fusion protein cleavage products will be $(X)_n$R and R-*peptide*. In another example, a fusion protein of formula $(X)_n$KK–*peptide*, wherein $(X)_n$KK is the affinity tag having two consecutive lysine residues, and *peptide* is the peptide, may be cleaved in one embodiment by a protease that targets consecutive basic residues, such as OmpT. In such an embodiment, the fusion protein cleavage products will be $(X)_n$K and K-*peptide*.

[0029] Proteins of the present disclosure may further comprise a detectable moiety, such as an enzyme, a prosthetic group, a fluorescent material, a luminescent material, a bioluminescent material, a radioactive material, a positron emitting metal, and/or a nonradioactive paramagnetic metal ion. The specific detectable moiety used will depend on the method of detection used; non-limiting examples of detection methods for which a suitable detection moiety may be incorporated into a fusion protein of the present disclosure include: flow cytometric detection, scanning laser cytometric detection, fluorescent immunoassays, enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), bioassays (e.g., neutralization assays), and Western blotting applications. The detection moiety may be linked to the fusion protein through covalent bonding including, but not limited to, covalent organic bonds (e.g., C-C, C-N, C-O, etc.), disulfide bonding, hydrogen bonding, electrostatic bonding, recombinant fusion, and/or conformational bonding. In some embodiments, the detection moiety and the affinity tag are linked by one or more linking compounds.

[0030] In some embodiments, a protein of the present disclosure further comprises a marker sequence. The marker sequence, for example, may be linked to the fusion protein through covalent bonding including, but not limited to, covalent organic bonds (e.g., C-C, C-N, C-O, etc.), disulfide bonding, hydrogen bonding, electrostatic bonding, recombinant fusion, and/or conformational bonding. In some embodiments, the marker sequence and the affinity tag are linked by one or more linking compounds. In some embodiments, the marker sequence is selected from: a hexa-histidine tag, a myc tag, and a flag tag.

[0031] In some embodiments, a protein of the present disclosure comprises an affinity tag, a peptide, and a cleavage site. The affinity tag and the peptide may be separated by the cleavage site. In some embodiments, the affinity tag has an amino acid sequence at least 70% homologous to or identical to SEQ ID NO.: 1.

[0032] In some embodiments, a protein of the present disclosure comprises an affinity tag, a peptide, and a cleavage site. The affinity tag and the peptide may be separated by the cleavage site. In some embodiments, the affinity tag has an amino acid sequence at least 70% homologous to or identical to SEQ ID NO.: 2.

[0033] In some embodiments, a protein of the present disclosure comprises an affinity tag, a peptide, and a cleavage site. The affinity tag and the peptide may be separated by the cleavage site. In some embodiments, the affinity tag has an amino acid sequence at least 70% homologous to or identical to SEQ ID NO.: 3.

[0034] In some embodiments, a protein of the present disclosure comprises an affinity tag, a peptide, and a cleavage site. The affinity tag and the peptide may be separated by the cleavage site. In some embodiments, the affinity tag has an amino acid sequence at least 70% homologous to or identical to SEQ ID NO.: 4.

Selected Embodiments of Nucleic Acids Encoding Affinity Tags and Proteins Comprising Same

[0035] The present disclosure provides isolated nucleic acids encoding a protein consistent with the present technology (e.g., a protein comprising an affinity tag, a peptide and a cleavage site). In some embodiments, the isolated nucleic acid comprises RNA or DNA. As used herein, the phrase "isolated nucleic acid" refers to a protein that has been removed from its normal surrounding nucleic acid sequences in the genome or in cDNA sequences, and from which introns (if any) have been removed. Isolated nucleic acids consistent with the present disclosure may comprise additional sequences useful for promoting expression and/or purification of the encoded protein, including but not limited to polyA sequences, modified Kozak sequences, and sequences encoding epitope tags, export signals, and secretory signals, nuclear localization signals, and plasma membrane localization signals. It will be apparent to those of skill in the art, based on the teachings herein, what nucleic acid sequences will encode the proteins of the present disclosure.

[0036] In some embodiments, the present disclosure provides a recombinant expression vector comprising an isolated nucleic acid disclosed herein operatively linked to a suitable

control sequence. As used herein, the phrase "recombinant expression vector" includes vectors that operatively link a nucleic acid coding region or gene to any control sequences capable of effecting expression of the gene product. The term "control sequences" refers to nucleic acid sequences capable of affecting expression of the nucleic acid molecules. As will be recognized by one of skill in the art, a control sequence need not be contiguous with the associated nucleic acid sequence, so long as it functions to direct the expression of the nucleic acid sequence. Thus, for example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the nucleic acid sequences and the promoter sequence can still be considered "operably linked" to the coding sequence. Other such control sequences include, but are not limited to, polyadenylation signals, termination signals, and ribosome binding sites. Such expression vectors can be of any type known in the art including, but not limited to, plasmid and viral-based expression vectors. The control sequence used to drive expression of the disclosed nucleic acid sequences in a mammalian system may be constitutive (driven by any of a variety of promoters including, but not limited to, CMV, SV40, RSV, actin, EF) or inducible (driven by any of a number of inducible promoters including, but not limited to, tetracycline, ecdysone, steroid- responsive). The construction of expression vectors for use in transfecting prokaryotic cells is also well known in the art, and thus can be accomplished via standard techniques, such as those described in Sambrook, Fritsch, and Maniatis, in: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989; Gene Transfer and Expression Protocols, pp. 109-128, ed. E.J. Murray, The Humana Press Inc., Clifton, N.J.), and the Ambion 1998 Catalog (Ambion, Austin, TX). In general the expression vector must be replicable in a host organism either as an episome or by integration into host chromosomal DNA. In a preferred embodiment, the expression vector comprises a plasmid. However, the scope of the present disclosure includes other expression vectors that serve equivalent functions, such as viral vectors.

[0037] In some embodiments, the present disclosure provides a host cell that has been transfected with a recombinant expression vector disclosed herein. The host cells may be prokaryotic (such as bacteria) or eukaryotic. The host cell can be transiently or stably transfected according to any technique known in the art, such as standard bacterial transformations, calcium phosphate co-precipitation, electroporation, or liposome mediated-, DEAE dextran mediated-, polycationic mediated-, or viral-mediated transfection.

[0038] The present disclosure provides a method of producing a protein, the method comprising: (a) culturing a host according to this aspect of the invention under conditions conducive to the expression of the protein, and (b) optionally, recovering the expressed protein. The expressed protein can be recovered, for example, by heating host bacterial cells expressing the protein to approximately up to 80 °C. Therefore, little to no other physical or chemical step is necessary to release the polypeptides from the cells. Centrifugation eliminates cell debris and aggregates. Purification of the expressed protein from the supernatant may be accomplished by any purification method disclosed herein.

Selected Methods of Purifying Proteins Comprising and Affinity Tag

[0039] The present disclosure provides methods for purifying a protein. In some embodiments, the protein is a fusion protein comprising an affinity tag, a peptide, and optionally a cleavage site. Generally, the method of purifying a protein comprises contacting a silicon oxide substrate with a fluid comprising the protein (e.g., fusion protein), contacting the silicon oxide substrate with a washing fluid, and contacting the silicon oxide substrate with a releasing agent to provide the purified fusion protein. The method optionally additionally includes contacting the purified fusion protein with a protease to cleave the affinity tag or a portion thereof from the fusion protein to provide a purified peptide.

[0040] The present disclosure also provides methods for purifying a peptide, the method comprising contacting a silicon oxide substrate with a fluid comprising a fusion protein, the fusion protein comprising an affinity tag, a peptide, and optionally a cleavage site; contacting the silicon oxide substrate with a washing fluid; and contacting the silicon oxide substrate with a releasing agent comprising a protease to provide the purified peptide. In some embodiments, the purified peptide additionally includes at least one terminal amino acid residue left over from the cleaved cleavage site and/or from the affinity tag.

[0041] The protein to be purified may be any protein that is capable of releasably binding to a silicon oxide substrate as disclosed herein. In some embodiments, the protein is a fusion protein described herein, such as a fusion protein comprising an affinity tag, a peptide, and optionally, a cleavage site.

[0042] The step of contacting the silicon oxide substrate may include, for example, contacting a suitable silicon oxide substrate (e.g., silica) with a suspension or solution of a fusion protein including an affinity tag capable of binding to silica. In some embodiments, the silicon oxide substrate (e.g., silica) is housed in a suitable vessel, such as a column, which is

formed of a material other than a silicon oxide-based material. In some embodiments, the vessel is a plastic or inert metal vessel.

[0043] The step of contacting the silicon oxide substrate with the fluid comprising the protein (e.g., fusion protein) may be performed at a temperature sufficient to enable efficient pouring of the fusion protein suspension or solution while avoiding denaturing of the fusion protein or any portion thereof. Thus, a person of skill in the art will be able to select a suitable temperature based on the peptide(s) of the fusion protein and their sensitivities to elevated temperatures.

[0044] The step of contacting the silicon oxide substrate with the fluid comprising the protein (e.g., fusion protein) may be performed for a period of time sufficient to enable complete adsorption of the fusion protein to the silicon oxide substrate. Generally, when the fusion protein includes an affinity tag as disclosed herein, adsorption occurs quickly at room temperature, usually within several minutes. Thus, the step of contacting the silicon oxide may be performed for a period of time of at least about 1 minute, 2 minutes, at least about 3 minutes, at least about 4 minutes, at least about 5 minutes, at least about 6 minutes, at least about 7 minutes, at least about 8 minutes, at least about 9 minutes, at least about 10 minutes, at least about 11 minutes, at least about 12 minutes, at least about 13 minutes, at least about 14 minutes, at least about 15 minutes, at least about 16 minutes, at least about 17 minutes, at least about 18 minutes, at least about 19 minutes, at least about 20 minutes, at least about 25 minutes, at least about 30 minutes, at least about 35 minutes, at least about 40 minutes, at least about 45 minutes, at least about 50 minutes, at least about 55 minutes, at least about 60 minutes, or more than 60 minutes.

[0045] The step of contacting the silicon oxide substrate with a washing fluid may be performed using any suitable washing fluid capable of flushing any unbound fusion proteins, contaminating proteins, or other undesired components of the fusion protein suspension or solution. When the fusion protein comprises an affinity tag disclosed herein, the washing fluid does not include arginine or lysine, or includes arginine or lysine (individually or combined) at a level below about 0.1 M, for example no greater than 0.1 M, no greater than about 0.09 M, no greater than about 0.08 M, no greater than about 0.07 M, no greater than about 0.06 M, no greater than about 0.05 M, no greater than about 0.04 M, no greater than about 0.03 M, no greater than about 0.02 M, no greater than about 0.01 M, no greater than about 9 mM, no greater than about 8 mM, no greater than about 7 mM, no greater than about 6 mM, no greater than

about 5 mM, no greater than about 4 mM, no greater than about 3 mM, no greater than about 2 mM, or no greater than about 1 mM.

[0046] The step of contacting the silicon oxide substrate with a washing fluid may be performed at a temperature sufficient to enable efficient pouring and flow of the washing fluid suspension or solution through the silicon oxide substrate, while avoiding denaturing of the fusion protein or any portion thereof. Thus, a person of skill in the art will be able to select a suitable temperature for the washing step based on the peptide(s) of the fusion protein and their sensitivities to elevated temperatures.

[0047] The step of contacting the silicon oxide substrate with a washing fluid may be performed for a period of time sufficient to enable complete flushing of non-adhered contaminants from the silicon oxide substrate. As will be recognized by one of skill in the art, the amount of time required will depend on a number of factors including, for example, the solubility of contaminants in the washing fluid, the viscosity of the washing fluid, the amount of silicon oxide substrate used (e.g., the length of a silica column), the amount of washing fluid to be used, and the pressure applied to the washing fluid (if any) and/or the reduced pressure applied downstream of the silicon oxide substrate (if any). Thus, the step of contacting the silicon oxide with the washing fluid may be performed for a period of time of at least about 1 minute, 2 minutes, at least about 3 minutes, at least about 4 minutes, at least about 5 minutes, at least about 6 minutes, at least about 7 minutes, at least about 8 minutes, at least about 9 minutes, at least about 10 minutes, at least about 11 minutes, at least about 12 minutes, at least about 13 minutes, at least about 14 minutes, at least about 15 minutes, at least about 16 minutes, at least about 17 minutes, at least about 18 minutes, at least about 19 minutes, at least about 20 minutes, at least about 25 minutes, at least about 30 minutes, at least about 35 minutes, at least about 40 minutes, at least about 45 minutes, at least about 50 minutes, at least about 55 minutes, at least about 60 minutes, or more than 60 minutes.

[0048] The releasing agent used to contact the silicon oxide substrate to release the bound protein (e.g., adhered fusion protein) may comprise a solution or suspension of the releasing agent in a solvent. When the fusion protein includes an affinity tag as disclosed herein, the releasing agent may comprise, consist essentially of, or consist of arginine (e.g., L-arginine) and/or lysine (e.g., L-lysine). The releasing agent may be at any suitable concentration to induce the fusion protein to separate from the silicon oxide substrate. In some embodiments, the releasing agent comprises arginine (e.g., L-arginine) in a concentration of at least about 0.1 M,

for example at least about 0.1 M, at least about 0.2 M, at least about 0.3 M, at least about 0.4 M, at least about 0.5 M, at least about 0.6 M, at least about 0.7 M, at least about 0.8 M, at least about 0.9 M, at least about 1 M, or greater than about 1 M. In one embodiment, the releasing agent includes 0.5 M arginine (e.g., L-arginine). In another embodiment, the releasing agent includes 1 M arginine (e.g., L-arginine).

[0049] In some embodiments, the releasing agent comprises lysine (e.g., L-lysine) in a concentration of at least about 0.1 M, for example at least about 0.1 M, at least about 0.2 M, at least about 0.3 M, at least about 0.4 M, at least about 0.5 M, at least about 0.6 M, at least about 0.7 M, at least about 0.8 M, at least about 0.9 M, at least about 1 M, or greater than about 1 M. In a preferred embodiment, the releasing agent includes 0.5 M lysine (e.g., L-lysine). In another preferred embodiment, the releasing agent includes 1 M lysine (e.g., L-lysine).

[0050] Alternatively, the releasing agent may comprise, consist essentially of, or consist of a protease for cleaving the fusion protein, for example at a cleavage site. In some embodiments, the protease is selected to cleave the fusion protein at a cleavage site between the affinity tag and the peptide. In other embodiments, the protease is selected to cleave the fusion protein at a cleavage site within the affinity tag. In still other embodiments, the protease is selected to cleave the fusion protein at a cleavage site within the peptide. In some embodiments, the releasing agent comprises, consists essentially of, or consists of a protease selective for a cleavage site at or near the C-terminal or N-terminal of the affinity tag (e.g., whichever terminus is covalently bound to the peptide of a fusion protein). In some embodiments, the releasing agent comprises, consists essentially of, or consists of a protease selected from the group consisting of: OmpT, TEV protease, Thrombin, Factor Xa and Enterokinase.

[0051] The step of contacting the silicon oxide substrate with a releasing agent to release the bound protein (e.g., adhered fusion protein) or to cleave the peptide from the fusion protein (e.g, on-gel cleavage) may be performed at a temperature sufficient to enable efficient pouring of the releasing agent suspension or solution while avoiding denaturing of the fusion protein or any portion thereof. Thus, a person of skill in the art will be able to select a suitable temperature based on the peptide(s) of the fusion protein and their sensitivities to elevated temperatures.

[0052] The step of contacting the silicon oxide substrate with the releasing agent may be performed for a period of time sufficient to enable complete desorption of the fusion protein from the silicon oxide substrate or cleavage of the peptide from the fusion protein. Generally, when the fusion protein includes an affinity tag as disclosed herein, desorption occurs quickly at

room temperature, usually within several minutes, depending on the concentration and flow rate of the releasing agent. Thus, the step of contacting the silicon oxide may be performed for a period of time of at least about 1 minute, 2 minutes, at least about 3 minutes, at least about 4 minutes, at least about 5 minutes, at least about 6 minutes, at least about 7 minutes, at least about 8 minutes, at least about 9 minutes, at least about 10 minutes, at least about 11 minutes, at least about 12 minutes, at least about 13 minutes, at least about 14 minutes, at least about 15 minutes, at least about 16 minutes, at least about 17 minutes, at least about 18 minutes, at least about 19 minutes, at least about 20 minutes, at least about 25 minutes, at least about 30 minutes, at least about 35 minutes, at least about 40 minutes, at least about 45 minutes, at least about 50 minutes, at least about 55 minutes, at least about 60 minutes, or more than 60 minutes.

EXAMPLES

Example 1: Assay for Determining Binding Affinity to Silica.

[0053] A surface plasmon resonance ("SPR") experiment was conducted using TrxA::Car9 (SEQ ID NO.: 6), a derivative of *E. coli* thioredoxin (TrxA) including a disulfide-constrained Car9 sequence flanked by Cys-Gly-Pro and Gly-Cys-Pro tripeptides in place of thioredoxin's native Cys-Gly-Pro-Cys active site. This fusion protein was expressed along with wild type TrxA (SEQ ID NO.: 7).

[0054] Silica-coated SPR chips were prepared by chemical vapor deposition (CVD) to deposit a thin (~3nm) silica film onto gold-coated SPR chips. The silica-coated SPR chips were washed with ethanol and adsorbed organic matter was removed by exposure to UV-ozone for 20 minutes. Solutions of TrxA and TrxA::Car9 (1 μM each) were simultaneously flowed over the chip using a multichannel apparatus. The SPR chips were washed with buffer for 10 minutes to remove loosely bound protein.

[0055] As shown in FIG. 2, TrxA::Car9 rapidly adsorbed to the silica-coated SPR chips; more than 50% of the bound protein remained on the silica surface after washing. In contrast, TrxA did not bind. These data indicate that disulfide-constrained Car9 confers silica binding affinity to TrxA protein.

Example 2: Affinity of Car9 for Silica.

[0056] To determine if presentation in a disulfide-bonded loop is required for Car9 to bind silica and to facilitate the construction of fusion proteins containing a C-terminal Car9

extension, a cassette specifying a *Hin*dIII restriction site, a Gly-Gly-Gly-Ser linker and the Car9 dodecamer was inserted into plasmid pBLN200, a pET-24a(+) (Novagen) derivative in which the T7 promoter was replaced by a DNA segment encoding the *araC* gene and the arabinose-inducible P_{BAD} promoter. The resulting plasmid was named pBLN200-Car9 (SEQ ID NO.: 8).

[0057] A *Nde*I-*Hin*dIII cassette encoding the green fluorescent protein (GFP) variant GFPmut2 was next cloned into the same sites of pBLN200-Car9. Both authentic GFPmut2 (SEQ ID NO.: 9) and the derivative containing the Car9 extension at its C-terminus (SEQ ID NO.: 10 hereinafter referred to as GFPmut2-Car9) were expressed at high level and in a soluble form in *E. coli* BL21(DE3), purified by anion exchange chromatography, and dialyzed against 20 mM Tris-HCl pH 7.5.

[0058] GFPmut2-Car9 and GFPmut2 (as a control) were incubated separately with increasing amounts of silica gel for 1 hour, and the amount of unbound protein was quantified by assaying the supernatant by fluorescence spectroscopy. FIG. 3A shows that while GFPmut2 exhibited the characteristic behavior of a protein that adsorbs non-specifically to a solid, FIG. 3B demonstrates that GFPmut2-Car9 could be depleted from solution upon addition of small amounts of silica. In fact, while about 35 mg of silica gel was sufficient to capture all soluble GFPmut2-Car9 under these experimental conditions, data extrapolation indicates that more than 10 times that amount would be needed to quantitatively adsorb all GFPmut2 (FIG. 3C).

[0059] To more accurately quantify the affinity of the Car9 tag for silica, a series of SPR experiments were conducted on silica-coated chips prepared according to Example 1 at increasing protein concentrations. Consistent with the results discussed above, GFPmut2 had little affinity for the silica-coated SPR chip surface (FIG. 4, –●–). By contrast, GFPmut2-Car9 adhered to the silica film in a concentration-dependent manner (FIG. 4, –■–) and with a K_d of 1 μM. This equilibrium dissociation constant is identical to that exhibited by hexahistidine-tagged proteins for Ni-NTA. These data demonstrate that Car9 retains its ability to bind silica when in a linear conformation and that it does not induce GFP misfolding, as was expected from its hydrophilic and basic nature.

Example 3: Purification of Proteins Using Affinity Tags.

[0060] *E. coli* SF100 cells, which are deficient in the outer membrane associated protease OmpT, harboring plasmid pGFPmut2-Car9 were grown overnight at 37°C in 25 mL Luria Broth (LB) supplemented with 50 μg/mL kanamycin. Seed cultures were used to inoculate 500 mL of

LB medium, and cells were grown to $A_{600} \approx 0.5$ at 37°C. Cultures were transferred to a 25°C water bath for 10 minutes and protein synthesis was induced by supplementing the medium with 2% L-Arabinose. After 6 hours of cultivation at the same temperature, cells were harvested by centrifugation at 7,000g for 5 minutes, resuspended in 35 mL of 20 mM Tris-HCl pH 7.5 supplemented with 2 mM EDTA, and disrupted by 6 rounds of sonication for 3 minutes each at 30% duty cycle using a Branson sonifier. Lysates were clarified by centrifugation at 10,000g for 15 minutes.

[0061] For initial experiments, an aliquot (5 mL) of clarified lysate was mixed overnight at 8°C and with gentle agitation with 3 g of silica gel (63-200 μm spherical particles with 6 nm pore size; cat. no. 391484, Sigma-Aldrich, St. Louis, MO). The protein-loaded resin was loaded onto a 1 cm inner diameter chromatography column (Pharmacia) and washed at 1 mL/min with 20 mM Tris-HCl pH 7.5 until no protein was detected in the effluent. Addition of up to 5 M NaCl or $MgCl_2$ proved ineffective at releasing the bound protein. However, supplementing the buffer with 0.5 M arginine resulted in quantitative elution of highly pure GFPmut2-Car9 (FIG. 5A). A repeat of the experiment is shown in Fig. 6 (middle two lanes labelled "MBP"). These data seem to suggest that the guanidium group of arginine is an effective competitor of the electrostatic and H-bonding interactions that stabilize Car9-silica interactions.

[0062] A *Nde*I-*Hin*dIII cassette encoding maltose binding protein (MBP) was cloned into pBLN200-Car9. The resulting protein, MBP-Car9 (SEQ ID NO.: 11), was expressed at high level and in a partially soluble form in *E. coli* SF100 as described in section. An aliquot (5 mL) of clarified lysate was mixed overnight at 8°C and with gentle agitation with 3 g of silica gel (63-200 μm spherical particles with 6 nm pore size; cat. no. 391484, Sigma-Aldrich, St. Louis, MO). The protein-loaded resin was loaded onto a 1 cm inner diameter chromatography column (Pharmacia) and washed at 1 mL/min with 20 mM Tris-HCl pH 7.5 until no protein was detected in the effluent. Supplementing the buffer with 0.5 M arginine resulted in quantitative elution of highly pure MBP-Car9 (FIG. 5B).

[0063] A *Nde*I-*Hin*dIII cassette encoding the fluorescent protein mCherry with a N-terminal hexa-histidine tag was cloned into pBLN200-Car9. The resulting protein, mCherry-Car9 (SEQ ID NO.: 12), was expressed at high level and in a partially soluble form in *E. coli* SF100 as described above. An aliquot (5 mL) of clarified lysate was mixed overnight at 8°C and with gentle agitation with 3 g of silica gel (63-200 μm spherical particles with 6 nm pore size; cat. no. 391484, Sigma-Aldrich, St. Louis, MO). The protein-loaded resin was loaded onto

a 1 cm inner diameter chromatography column (Pharmacia) and washed at 1 mL/min with 20 mM Tris-HCl pH 7.5 until no protein was detected in the effluent. Supplementing the buffer with 0.5 M arginine resulted in quantitative elution of highly pure mCherry-Car9 (FIG. 6, right two lanes labelled "mCherry").

[0064] Clarified extracts from SF100 cells expressing GFPmut2-Car9 ("GFPmut2" in FIG. 6), MBP-Car9 ("MBP" in FIG. 6) or mCherry-Car9 ("mCherry" in FIG. 6) were prepared as described above. Aliquots (5 mL) of clarified lysates were mixed overnight at 8°C and with gentle agitation with 3 g of silica gel (63-200 μm spherical particles with 6 nm pore size; Sigma). Each preparation of protein-loaded resin was loaded onto a 1 cm inner diameter chromatography column (Pharmacia) and washed at 1 mL/min with 20 mM Tris-HCl pH 7.5 until no protein was detected in the effluent. Supplementing the buffer with 1.0 M L-lysine resulted in quantitative elution of highly pure GFP-mut2, MBP-Car9 and mCherry-Car9 (FIG. 6).

Example 4: Rapid Purification from Crude Cellular Extracts Using Affinity Tags.

[0065] A device 100 consistent with that shown in FIG. 7A was constructed using two 30-mL syringes 110, 120 tethered to each other by a valve 130. The valve 130 was closed and a glass wool plug 112 was inserted into the first syringe 110. About 3 g of silica gel was loaded into the first syringe 110, on top of which a perforated plastic screen 116 was placed.

[0066] 90 mL of 20 mM Tris-HCl at pH 7.5 was loaded into the first syringe 110 to wash the matrix 112-116. The buffer was removed by pulling on the plunger 122 of the second syringe 120. 0.5 mL of clarified extract 140 from induced SF100(pGFPmut2-Car9) cultures was loaded into the first syringe 110 and flowed through the matrix 112-116 by pulling the plunger 122. 90 mL of buffer supplemented with 1M L-lysine was then added to the first syringe 110 and drawn through the matrix 112-116 by pulling on the plunger 122 to yield purified protein 150.

[0067] As shown in FIG. 7B, the purified protein 150 (lane "E") essentially included only the GFPmut2-Car9 fusion protein. The lysate (lane "L") included a small amount of the GFPmut2-Car9 fusion protein along with other impurities. The entire process of purifying the lysate was accomplished in less than 15 minutes.

Example 5: Cleavage of Affinity Tags in Cell Lysates.

[0068] OmpT is an aspartate protease that exhibits narrow specificity for paired basic residues. It is located in the outer membrane of *E. coli* K12 strains but absent in *E. coli* B strains (e.g., BLD1); cell integrity must be disrupted in order for OmpT to have access to intracellular substrates. The Car9 sequence contains two sets of paired basic residues: an internal Lys-Lys and a C-terminal Lys-Arg sequence. In addition, fusion of the Car9 to the C-terminus of GFPmut2 via a Gly-Gly-Gly-Ser linker introduces a Lys-Lys dipeptide at the fusion joint. To determine if these sites would be accessible to OmpT, GFPmut2-Car9 and MBP-Car9 were expressed as described above, except that the *ompT*$^+$ strains KS272 (F′ Δ*lacX*74 *galE galK thi rps*L(*strA*) Δ*phoA*) and Top10 (F$^-$ *endA*1 *recA*1 *hsdR*17 (r^-_k,m^+_k) λ$^-$ *supE*44 *thi*1 *gyrA*96 *relA*1 φ80Δ*lac*Δ*M15*Δ(*lacZYA-argF*)U169 *deo*R) were used. Expressed fusion protein was assayed according to the protocol of Example 3, except that the lysates were held for 30 minutes on ice to allow membrane-associated OmpT to access its potential substrates.

[0069] As shown in FIGS. 8A and 8B, whole cell lysates (lanes "WC") indicated the presence of intact GFPmut2-Car9 and MPR-Car9, indicated by the asterisked arrows, ←*. A lower molecular weight product, indicated by the daggered arrows, ←(†), accumulated in lysed fractions (lanes "L") did not show affinity for silica; these products flowed through the silica gel (lanes "FT") and were not released from silica by the 1M lysine eluent (lanes "E").

[0070] These data indicate that the Car9 affinity tag is solvent/protease accessible, and that the terminal Lys-Pro-Gly-Lys-Arg sequence contributes to silica adhesion. In addition, these data suggest that OmpT could be used to remove almost half of a C-terminal Car9 affinity tag from a fusion protein such as MBP-Car9 or all of the affinity tag from a fusion protein such as GFPmut2-Car9 instead of more expensive site-specific proteases (e.g., thrombin, factor Xa, TEV protease, etc.). Similarly, OmpT digestion may be used to remove nearly all of an N-terminal Car9 affinity tag (i.e., only the terminal arginine residue would remain). Thus, purification of a fusion protein comprising a Car9 affinity tag by silica gel chromatography, followed by OmpT digestion, can rapidly and inexpensively provide purified proteins.

Example 6: Cleavage of Affinity Tags Within Intact Cells

[0071] To determine if the OmpT protease would be useful within the context of intact cells to inexpensively remove the Car9 affinity tag, SF100 cells transformed or not with pML19 (SEQ ID NO.: 13), a multicopy plasmid encoding OmpT, were grown at 37°C in 25 mL of LB medium supplemented with 100 μg/mL carbenicillin. After 5h of growth at 37°C, cells were

harvested from culture volumes corresponding to 3 absorbance units at 600 nm (A_{600}) and washed 3 times with 2 mL of 50 mM sodium phosphate pH 7.5 with intervening 5000*g* centrifugation steps. After the final wash, cells were resuspended in 100 µL of 50 mM phosphate buffer pH 7.5 containing purified GFPmut2-Car9 at 10 µM final concentration. After 30 or 45 min incubation at room temperature, cells were sedimented at 5000*g*, and aliquots of the supernatants were fractionated by SDS-PAGE.

[0072] FIG. 9A shows that, while the fusion protein remained intact when placed in contact with Δ*ompT* cells (gray, asterisked arrow), it was converted into its expected degradation products (black, daggered arrows) after 30 min of incubation with OmpT-overproducing cells. In addition, tag-free GFPmut2 was the dominant product after 15 additional minutes of incubation (FIG. 9B, daggered arrow).

Example 7: Cost of Use.

[0073] Because the present technology utilizes inexpensive, unmodified silica gel instead of a specialty solid phase, the cost to purify a protein is significantly lower than presently available alternatives. Purification of about 15 mg of purified protein using methods disclosed herein may require about 3 g of silica. Thus, purification reagent costs are estimated to be on the order of US$1.50 per 10 mg of purified protein. In contrast, purifying the same amount of His-tagged protein using Ni-NTA technology, the least expensive alternative widely available at this time, would cost more than 10 times that amount.

Example 8: Functionalization of Glass Surfaces Using Fusion Proteins.

[0074] Decorating glass, borosilicate and silicon with proteins is important in sensor and microarray development, as well as for electronic, bioimaging and medical applications. Typically, silane coupling agents terminated with an amino (e.g., aminopropylsilane) or a thiol (e.g., mercaptopropylsilane) group that is suitable for subsequent protein conjugation. The affinity tags disclosed herein (e.g., Car9) may be used instead to functionalize silica surfaces by simply contacting the silica surface with a solution of a fusion proteins as disclosed herein. Unlike functionalization using a traditional silane coupling agent, the use of a fusion protein additionally enables oriented immobilization (through the tagged end of the fusion protein) and selective release (by addition of arginine or lysine).

[0075] In one exemplary illustration of the above, aliquots of 10 µM solutions of GFPmut2, GFPmut2-Car9 and GFPmut2-Car15 (a derivative a GFPmut2 fitted with the Arg-

Thr-Tyr-Leu-Pro-Leu-Pro-Trp-Met-Ala-Ala-Leu carbon binding sequence Car15 (SEQ ID NO.: 5) in place of Car9; SEQ ID NO.: 14) were deposited using a toothpick onto a glass microscope slide that had been cleaned with ddH_2O and dried. Imaging with a GE Typhoon FLA 9000 scanner revealed that equivalent amount of fluorescent material had been deposited onto the surface (FIG. 10A). After a 5 minute wash in ddH_2O with gentle shaking, all GFPmut2 and most of GFPmut2-Car15 had been removed, while a monolayer of GFPmut2-Car9 remained adhered to the glass slide (Fig. 10B). Immersion of the slide in 20 mM Tris-HCl at pH 7.5 supplemented with 0.5 M L-arginine led to almost complete release of the adsorbed protein (FIG. 10C), confirming glass adhesion of the fusion protein through the Car9 affinity tag. These data indicate that the Car9 affinity tag represents a straightforward one-step alternative to silane chemistry for functionalizing silica surfaces. In addition, and unlike silane chemical techniques, functionalization of glass surfaces with a fusion protein as disclosed herein is chemically reversible.

Example 9: Identification of Additional Releasable Affinity Tags

[0076] *E. coli* GI826 cells (F^- *lacI*^{q} *ampC*::P_{trp}*cI* Δ*fliC* Δ*motB eda*::Tn*10*) lacking a plasmid or transformed with derivatives of plasmid pFliTrx (Invitrogen) specifying Car8, Car9, Car12 or Car36 flanked by a N-terminal Cys-Gly-Pro tripeptide and a C-terminal Gly-Pro-Cys tripeptide, and inserted within the FliTrx protein (SEQ ID NO.: 15; a synthetic, cell surface exposed fusion protein; Invitrogen) were grown in 5 mL of IMC medium (6 g of Na_2HPO_4, 3 g of KH_2PO_4, 0.5 g of NaCl, and1 g of NH_4Cl in 1L of deionized water with 0.2% casamino acids, 0.5% glucose and 1 mM $MgCl_2$) supplemented with 100 µg/mL carbenicillin for 24h at 25°C. Culture aliquots (1 mL) were used to inoculate 125 mL shake flasks containing 25 mL of IMC medium supplemented with 100 µg/mL carbenicillin. Cells were grown at 25°C to $A_{600} \approx 0.5$ and the synthesis of FliTrx::Car8 (SEQ ID NO.: 16), FliTrx::Car9 (SEQ ID NO.: 17), FliTrx::Car12 (SEQ ID NO.: 18) and FliTrx::Car36 (SEQ ID NO.: 19) was induced by addition of 0.1 mg/mL L-tryptophan. Cultivation was continued for 5 hours or more and until A_{600} exceeded 1. The turbidity of all cultures was then adjusted to 1.0 using IMC medium.

[0077] Aliquots of these solutions (2 mL) were transferred in duplicate to the wells of 24-wells cell culture plates containing a ~1x1 cm piece of glass microscope slide (VWR). After 15 minutes of incubation at room temperature with gentle agitation, the liquid was aspired and 2 mL of Buffer A (20 mM Tris-HCl, pH 7.5) was added as wash. After 1 minute, the liquid was aspired. This wash cycle was repeated 2 additional times.

[0078] Duplicate samples were then treated with either 2 mL of Buffer A or 2 mL of Buffer A supplemented with 1 M L-lysine. After 10 minutes incubation at room temperature with gentle agitation, the glass slides were removed, transferred to the stage of an optical microscope, and cells present in 3 different fields were counted at 90X magnification. This number was averaged to represent one measurement. The experiment was replicated three times in its entirety.

[0079] FIG. 11 shows that, in addition to the FliTrx::Car9 positive control, the FliTrx::Car8, FliTrx::Car12 and FliTrx::Car36 fusion proteins enable efficient cell attachment to microscope glass slides. By contrast, plasmid-free cells that do not express any FliTrx fusion protein on their surface ("GI826" in FIG. 11) do not appreciably adhere to the glass. FIG. 11 also shows that incubation with 1M L-lysine partially (FliTrx::Car8; FliTrx::Car9; FliTrx::Car12) or efficiently (FliTrx::Car36) precludes cell attachment to glass, demonstrating that L-lysine and/or arginine triggers release of the Car8, Car12 and Car36 affinity tags from silica-containing surfaces.

Example 10: Immobilization and Controlled Release of Proteins From Silica Sol-Gels

[0080] Sols are dispersions of colloidal particles in a liquid phase that can be polymerized into a rigid mass known as a gel through a variety of chemistries. The resulting porous networks can be used as matrices to entrap enzymes, proteins or other payloads for applications ranging from biosensing to controlled drug delivery. Some of the problems of these technologies include loss of bioactivity over time and uncontrolled leaking of adsorbed proteins into the surrounding medium.

[0081] A silica sol-gel was prepared according to standard methods. Briefly, 5.4 mL of concentrated NH_4OH was added to 1 L of water, and 5 mL of the solution was mixed with 10 mL of methanol in a beaker to produce a catalyst solution. Separately, 10 mL of tetramethoxysilane (TMOS) was mixed with an equal volume of methanol to create the alkoxide solution. The contents of the two beakers were combined and the solution was poured into a Petri dish and allowed to sit at room temperature for 15 min to allow for sol-gel formation. Methanol was added to cover the gel and replaced every 24 hours for a total period of 4 days. The gel was then washed with 20 mM Tris-HCl pH 7.5 for 72 hours as above, except that the solution was changed every 12 hours.

[0082] For protein immobilization, the gel was broken into pieces and a ~1 cm^3 piece was immersed into 1 mL of clarified lysate from induced SF100(pGFPmut2-Car9) cultures and incubated overnight at 4°C. The protein-loaded gel was then placed into 10 mL of 20 mM Tris-HCl at pH 7.5, allowed to sit for 10 minutes, and washed 5 times until all fluorescence had disappeared from the liquid (FIG. 12A). The hydrated gel could be stored for over two months at 4°C without any detectable decrease of fluorescence or leaching of GFP into the buffer. Furthermore, the immobilized protein could be released when by immersing the gel in 1 mL of 20 mM Tris-HCl at pH 7.5 buffer supplemented with 0.5M arginine (Fig. 12B).

[0083] These data indicate that the Car9 affinity tag enables efficient immobilization of proteins with silica networks prepared by sol-gel processes, and that the stably entrapped protein can be chemically released in an active form by chemical treatment.

Further Examples

1. A protein comprising SEQ ID NO.: 1.

2. A protein comprising SEQ ID NO.: 2.

3. A protein comprising SEQ ID NO.: 3.

4. A protein comprising SEQ ID NO.: 4.

5. A protein comprising an amino acid sequence having a homology of at least 70%, at least 80%, at least 90%, or at least 95% with the amino acid sequence of any one of SEQ ID NOs.: 1 to 4.

6. The protein of any one of examples 1 to 5, wherein the protein is a fusion protein.

7. The protein of example 6, wherein the fusion protein comprises a polyclonal antibody, a monoclonal antibody, a complement determining region-grafted antibody preparation, a hybrid antibody, an altered antibody, a $F(ab)'_2$ fragment, a Fab molecule, a Fv fragment, a single domain antibody, a chimeric antibody, or a fragment thereof.

8. A substrate comprising a bound protein of any one of examples 1 to 4.

9. The substrate of example 8, wherein the substrate includes a surface comprising a silicon oxide.

10. The substrate of example 8 or example 9, wherein the substrate comprises silica.

11. A method of isolating a protein, the method comprising:

contacting a silicon oxide substrate with a fluid comprising the protein for a time sufficient to allow the protein to bind to the silicon oxide substrate;

contacting the silicon oxide substrate with a washing fluid; and

contacting the silicon oxide substrate with a releasing agent to release the bound protein.

12. The method of example 11, wherein the protein is a protein of any one of examples 1 to 7.

13. The method of example 11 or example 12, wherein the releasing agent comprises lysine and/or arginine.

14. The method of example 11 or example 12, wherein the releasing agent comprises a protease.

15. The method of any one of examples 11 to 14, wherein the silicon oxide substrate comprises silica, borosilicate, controlled pore glass, quartz, and/or oxidized silicon.

16. The method of any one of examples 12 to 13, further comprising contacting the released protein with a protease to cleave the SEQ ID NOs.: 1 to 4, or homologous portion thereof, from the protein.

17. A kit for purifying the protein of any one of examples 1 to 7, the kit comprising;

a silicon oxide substrate;

a vessel for retainably holding the silicon oxide substrate;

a receptacle for receiving eluent passing through the vessel; and

a releasing agent.

18. The kit of example 17 further comprising instructions for using the kit to purify the protein.

19. The kit of example 17 or example 18, wherein the releasing agent comprises arginine and/or lysine.

20. The kit of example 17 or example 18, wherein the releasing agent comprises a protease

21. A nucleic acid encoding the protein of any one of examples 1 to 7.

22. A recombinant expression vector comprising the nucleic acid of example 21.

23. A host cell transfected with the recombinant expression vector of example 22.

24. A silica sol-gel comprising a protein comprising any one of SEQ ID NOs.: 1 to 4.

25. The silica sol-gel of example 24, wherein the protein is a fusion protein.

26. The silica sol-gel of example 23 or example 24, wherein the protein elutes from the silica sol-gel in the presence of a solution comprising arginine and/or lysine.

27. The silica sol-gel of example 26, wherein the solution comprises arginine and/or lysine in a concentration of at least about 0.1M.

Conclusion

[0084] The above detailed descriptions of embodiments of the technology are not intended to be exhaustive or to limit the technology to the precise form disclosed above. Although specific embodiments of, and examples for, the technology are described above for illustrative purposes, various equivalent modifications are possible within the scope of the technology, as those skilled in the relevant art will recognize. For example, while steps are presented in a given

order, alternative embodiments may perform steps in a different order. The various embodiments described herein may also be combined to provide further embodiments.

[0085] All references cited are herein incorporated by reference in their entirety. Within this application, unless otherwise stated, the techniques utilized may be found in any of several well-known references such as: Molecular Cloning: A Laboratory Manual (Sambrook, et al., 1989, Cold Spring Harbor Laboratory Press), Gene Expression Technology (Methods in Enzymology, Vol. 185, edited by D. Goeddel, 1991. Academic Press, San Diego, CA), "Guide to Protein Purification" in Methods in Enzymology (M.P. Deutshcer, ed., (1990) Academic Press, Inc.); PCR Protocols: A Guide to Methods and Applications (Innis, et al. 1990. Academic Press, San Diego, CA), Culture of Animal Cells: A Manual of Basic Technique, 2nd Ed. (R.I. Freshney. 1987. Liss, Inc. New York, NY), Gene Transfer and Expression Protocols, pp. 109-128, ed. E.J. Murray, The Humana Press Inc., Clifton, N.J.), and the Ambion 1998 Catalog (Ambion, Austin, TX). Aspects of the disclosure can be modified, if necessary, to employ the systems, functions, and concepts of the above references and application to provide yet further embodiments of the disclosure. These and other changes can be made to the disclosure in light of the detailed description.

[0086] As used herein, the amino acid residues are abbreviated as follows: alanine (Ala; A), asparagine (Asn; N), aspartic acid (Asp; D), arginine (Arg; R), cysteine (Cys; C), glutamic acid (Glu; E), glutamine (Gln; Q), glycine (Gly; G), histidine (His; H), isoleucine (Ile; I), leucine (Leu; L), lysine (Lys; K), methionine (Met; M), phenylalanine (Phe; F), proline (Pro; P), serine (Ser; S), threonine (Thr; T), tryptophan (Trp; W), tyrosine (Tyr; Y), and valine (Val; V).

[0087] As used throughout the present application, the term "polypeptide" is used in its broadest sense to refer to a sequence of subunit amino acids. The polypeptides of the invention may comprise L-amino acids, D-amino acids (which are resistant to L-amino acid-specific proteases in vivo), or a combination of D- and L-amino acids. The polypeptides described herein may be chemically synthesized or recombinantly expressed.

[0088] From the foregoing, it will be appreciated that specific embodiments of the invention have been described herein for purposes of illustration, but well-known structures and functions have not been shown or described in detail to avoid unnecessarily obscuring the description of the embodiments of the technology. Where the context permits, singular or plural terms may also include the plural or singular term, respectively. For example, as used herein, the singular forms "a", "an" and "the" include plural referents unless the context clearly dictates

otherwise. "And" as used herein is interchangeably used with "or" unless expressly stated otherwise. Moreover, unless the word “or” is expressly limited to mean only a single item exclusive from the other items in reference to a list of two or more items, then the use of “or” in such a list is to be interpreted as including (a) any single item in the list, (b) all of the items in the list, or (c) any combination of the items in the list. Additionally, the words “herein,” “above,” and “below” and words of similar import, when used in this application, shall refer to this application as a whole and not to any particular portions of the application. The term “comprising” is used throughout to mean including at least the recited feature(s) such that any greater number of the same feature and/or additional types of other features are not precluded.

[0089] It will also be appreciated that specific embodiments have been described herein for purposes of illustration, but that various modifications may be made without deviating from the technology. Further, while advantages associated with certain embodiments of the technology have been described in the context of those embodiments, other embodiments may also exhibit such advantages, and not all embodiments need necessarily exhibit such advantages to fall within the scope of the technology. Accordingly, the disclosure and associated technology can encompass other embodiments not expressly shown or described herein.

I/We claim:

1. A protein comprising SEQ ID NO.: 1.

2. A protein comprising SEQ ID NO.: 2

3. A protein comprising SEQ ID NO.: 3

4. A protein comprising SEQ ID NO.: 4

5. A protein comprising an amino acid sequence having a homology of at least 70%, at least 80%, at least 90%, or at least 95% with the amino acid sequence of any one of SEQ ID NOs.: 1 to 4.

6. The protein of any one of claims 1 to 5, wherein the protein is a fusion protein.

7. The protein of claim 6, wherein the fusion protein comprises a polyclonal antibody, a monoclonal antibody, a complement determining region-grafted antibody preparation, a hybrid antibody, an altered antibody, a $F(ab)'_2$ fragment, a Fab molecule, a Fv fragment, a single domain antibody, a chimeric antibody, or a fragment thereof.

8. A substrate comprising a bound protein of any one of claims 1 to 4.

9. The substrate of claim 8, wherein the substrate includes a surface comprising a silicon oxide.

10. The substrate of claim 8 or claim 9, wherein the substrate comprises silica.

11. A method of isolating a protein, the method comprising:

contacting a silicon oxide substrate with a fluid comprising the protein for a time sufficient to allow the protein to bind to the silicon oxide substrate;

contacting the silicon oxide substrate with a washing fluid; and

contacting the silicon oxide substrate with a releasing agent to release the bound protein.

12. The method of claim 11, wherein the protein is a protein of any one of claims 1 to 7.

13. The method of claim 11 or claim 12, wherein the releasing agent comprises lysine and/or arginine.

14. The method of claim 11 or claim 12, wherein the releasing agent comprises a protease.

15 The method of any one of claims 11 to 14, wherein the silicon oxide substrate comprises silica, borosilicate, controlled pore glass, quartz, and/or oxidized silicon.

16. The method of claim 12 or claim 13, further comprising contacting the released protein with a protease to cleave the SEQ ID NOs.: 1 to 4, or homologous portion thereof, from the protein.

17. A kit for purifying the protein of any one of claims 1 to 7, the kit comprising;
a silicon oxide substrate;
a vessel for retainably holding the silicon oxide substrate;
a receptacle for receiving eluent passing through the vessel; and
a releasing agent.

18. The kit of claim 17 further comprising instructions for using the kit to purify the protein.

19. The kit of claim 17 or claim 18, wherein the releasing agent comprises arginine and/or lysine.

20. The kit of claim 17 or claim 18, wherein the releasing agent comprises a protease.

21. A nucleic acid encoding the protein of any one of claims 1 to 7.

22. A recombinant expression vector comprising the nucleic acid of claim 21.

23. A host cell transfected with the recombinant expression vector of claim 22.

24. A silica sol-gel comprising a protein comprising any one of SEQ ID NOs.: 1 to 4.

25. The silica sol-gel of claim 24, wherein the protein is a fusion protein.

26. The silica sol-gel of claim 23 or claim 24, wherein the protein elutes from the silica sol-gel in the presence of a solution comprising arginine and/or lysine.

27. The silica sol-gel of claim 26, wherein the solution comprises arginine and/or lysine in a concentration of at least about 0.1M.

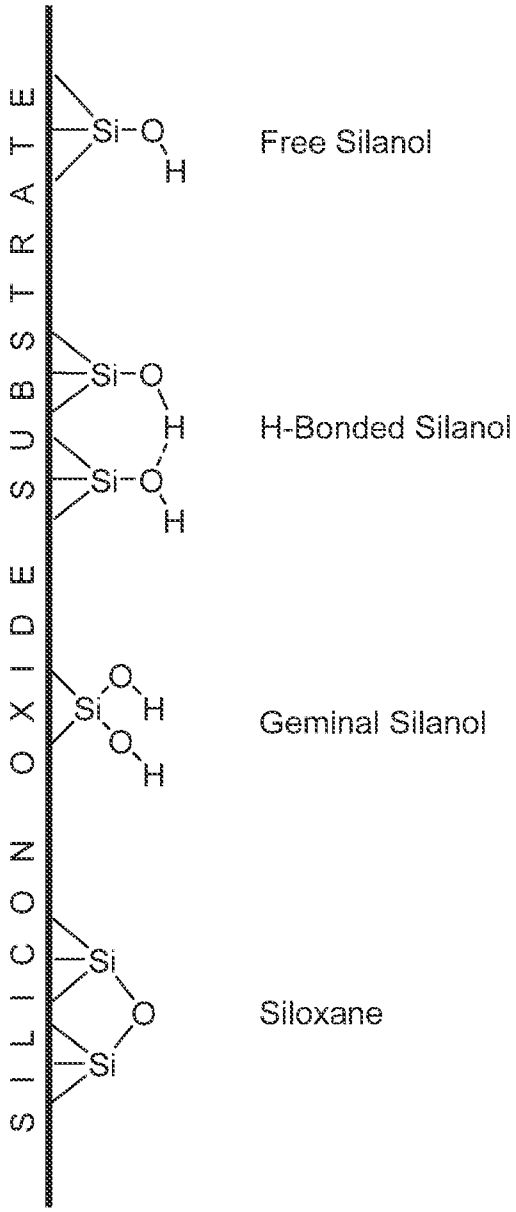

SILICON OXIDE SUBSTRATE
Free Silanol
H-Bonded Silanol
Geminal Silanol
Siloxane

FIG. 1

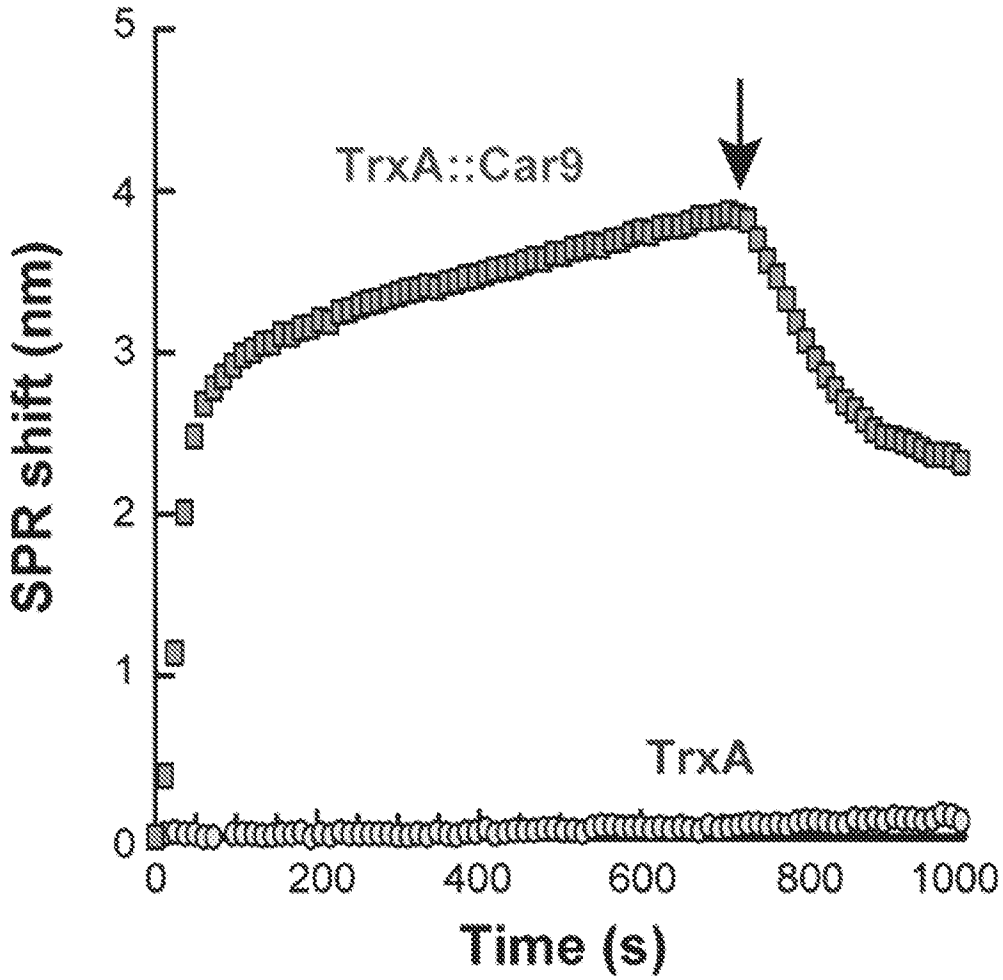

SPR shift (nm)
5
4
3
2
1
0
TrxA::Car9
TrxA
0
200
400
600
800
1000
Time (s)

FIG. 2

FIG. 3A

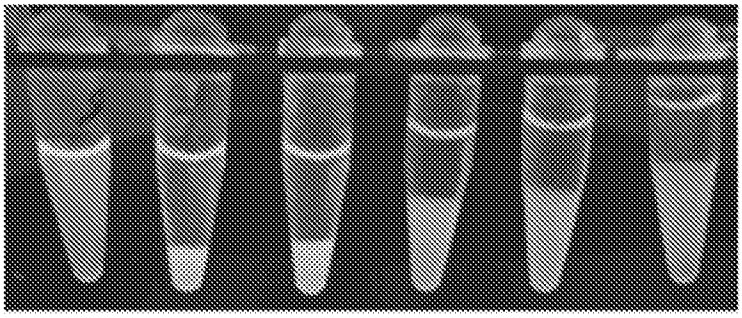
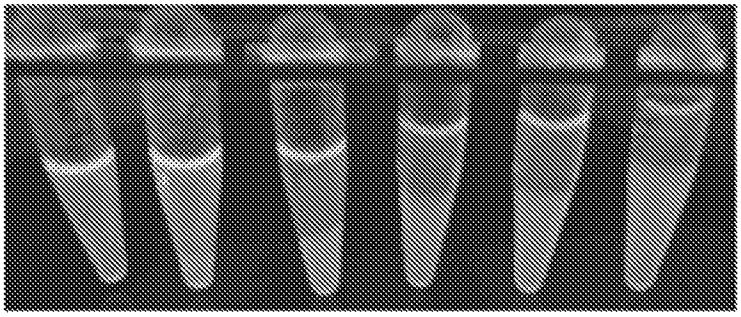


FIG. 3B



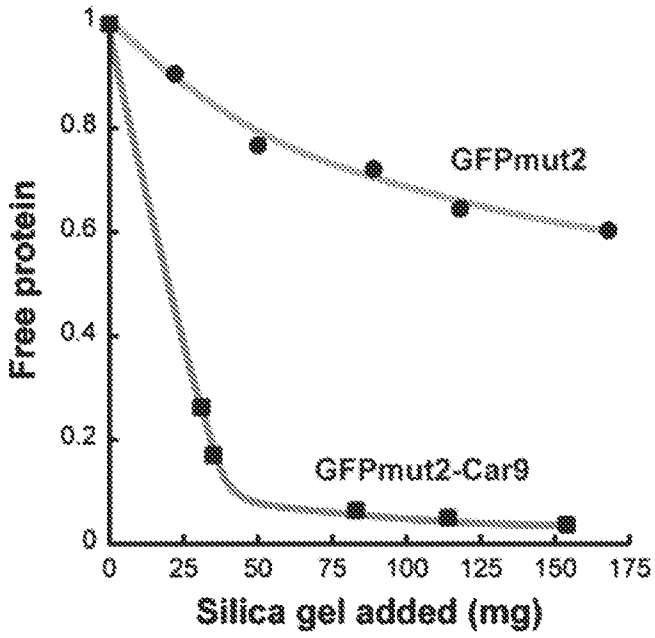


FIG. 3C

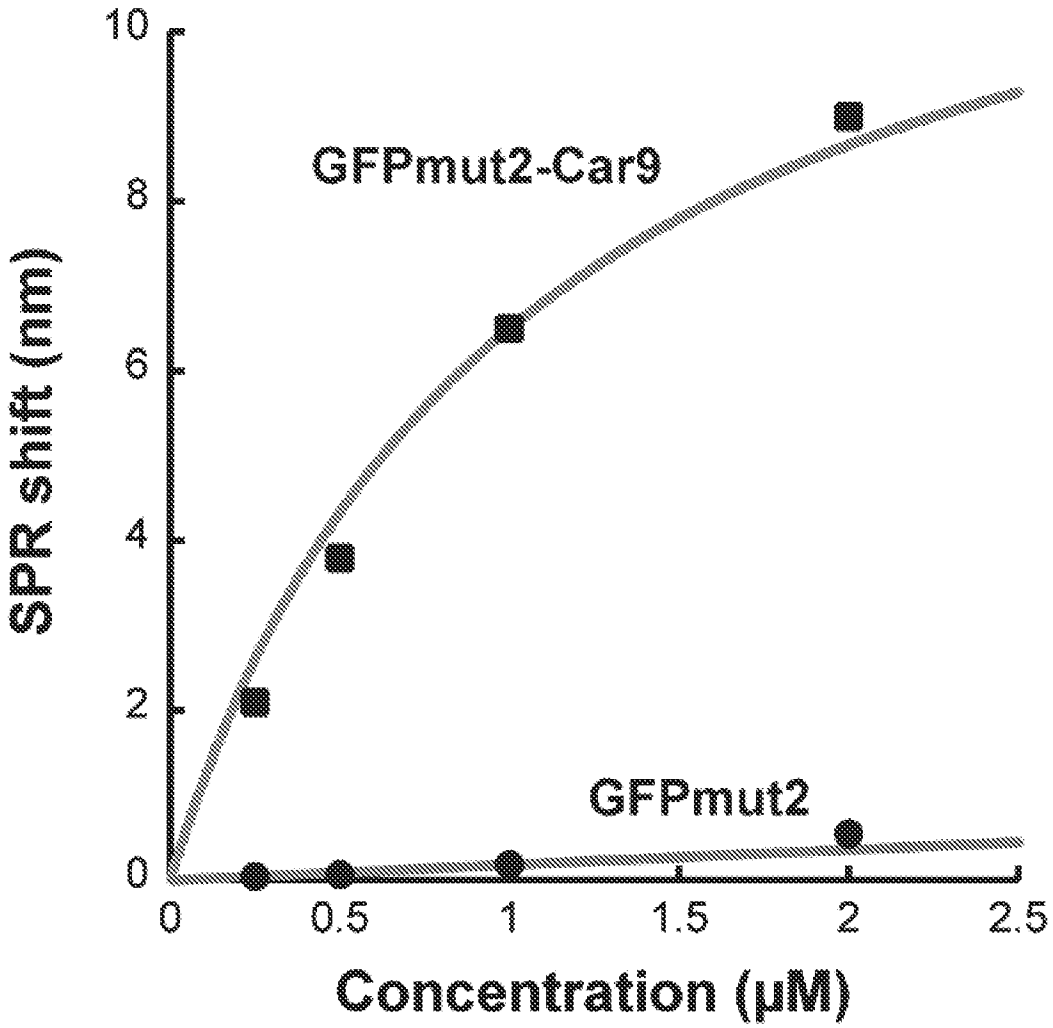

10
8
6
4
2
0
SPR shift (nm)
GFPmut2-Car9
GFPmut2
0
0.5
1
1.5
2
2.5
Concentration (μM)

FIG. 4

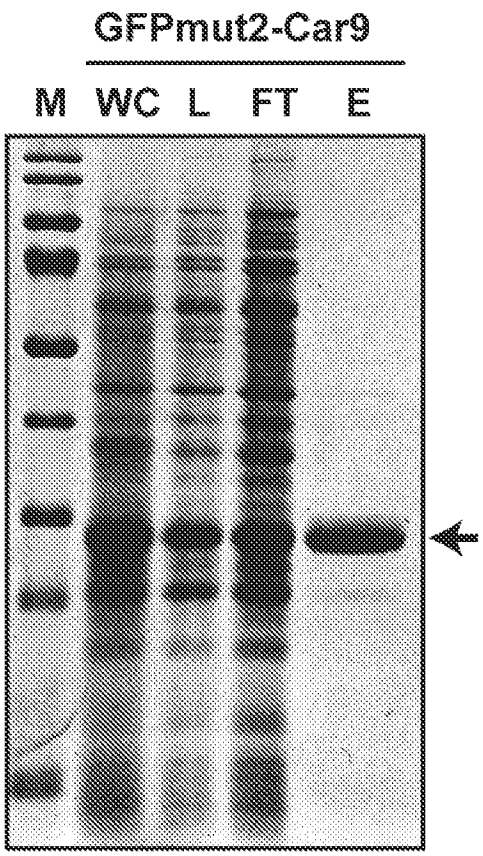


FIG. 5A

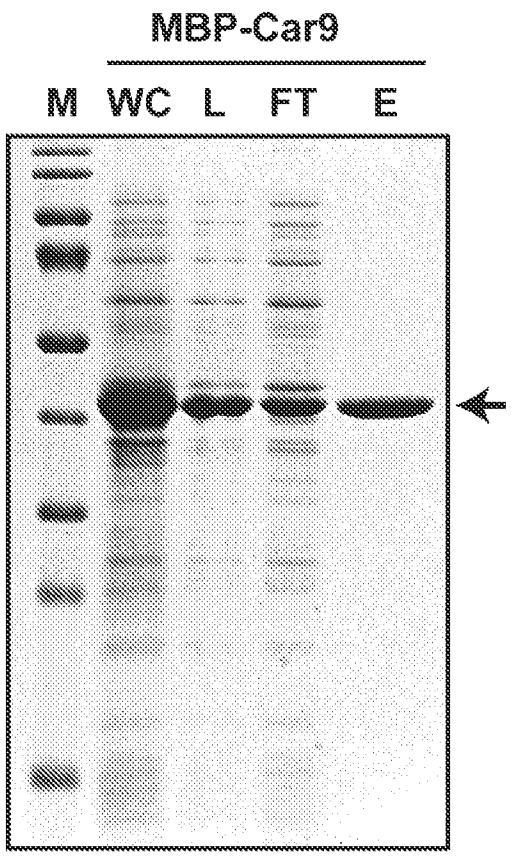


FIG. 5B

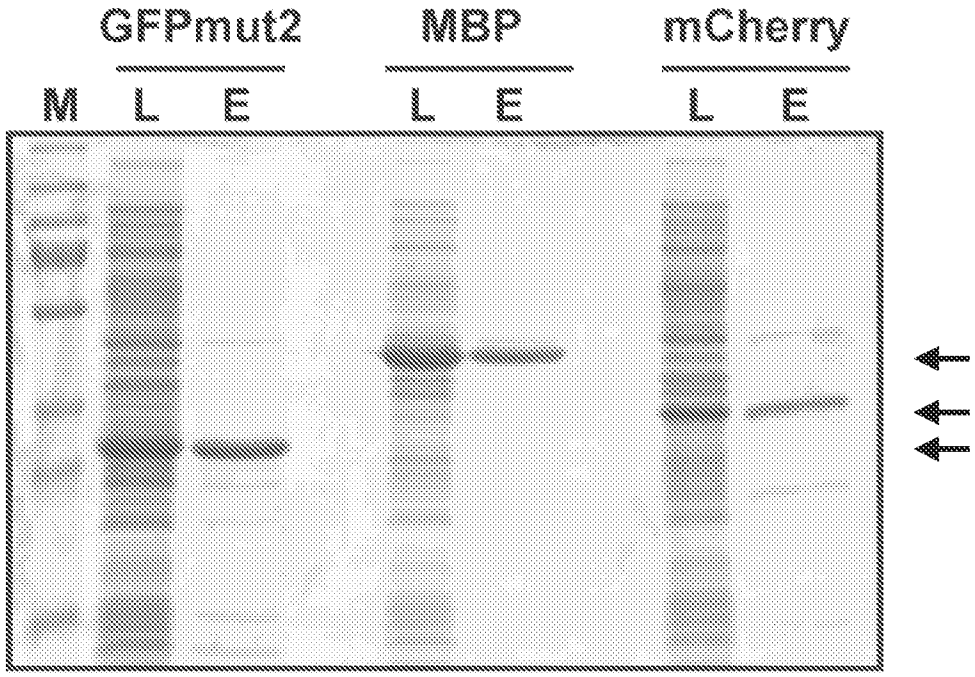


FIG. 6

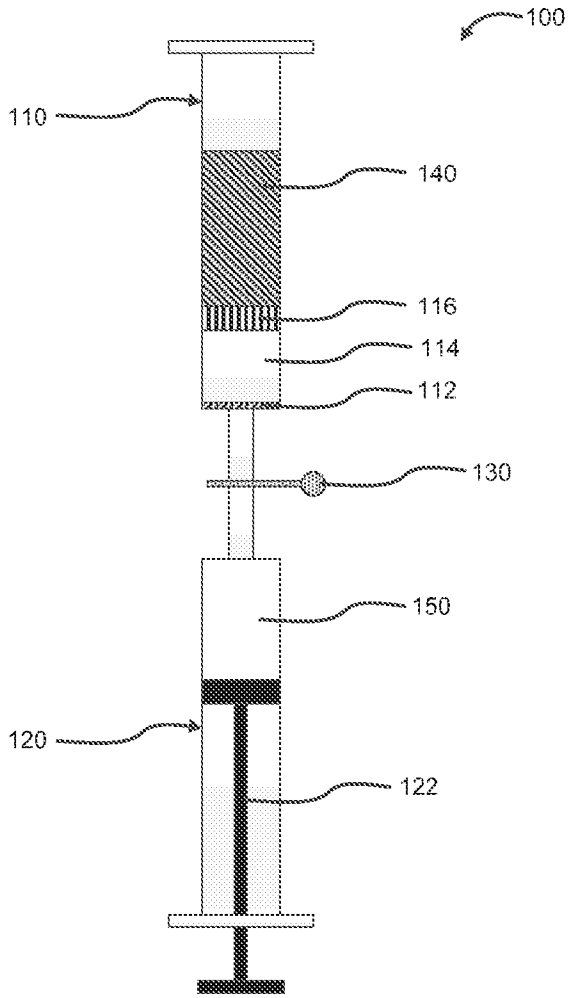


FIG. 7A

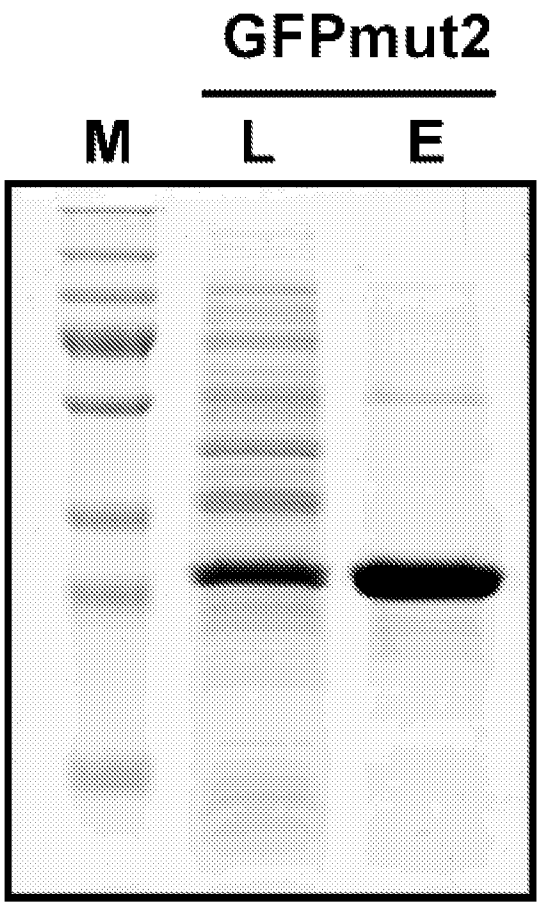


FIG. 7B

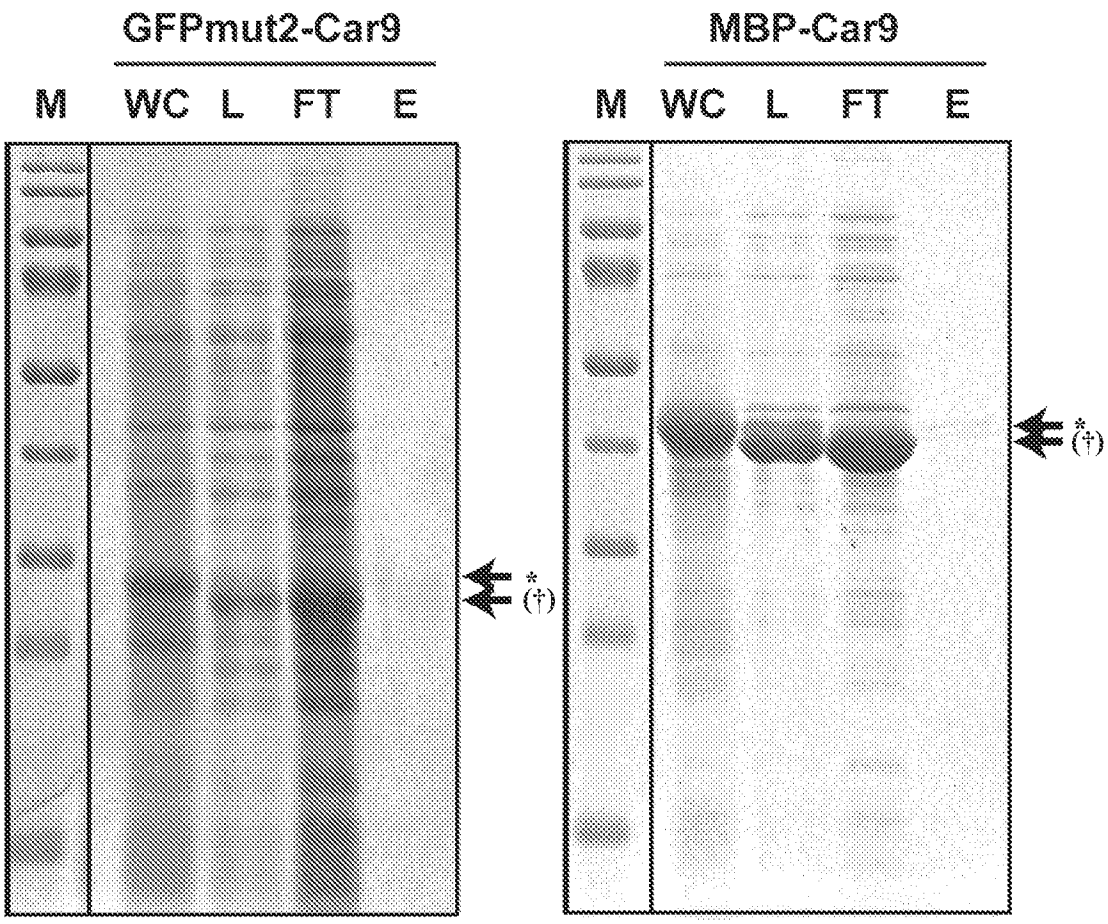


FIG. 8A FIG. 8B

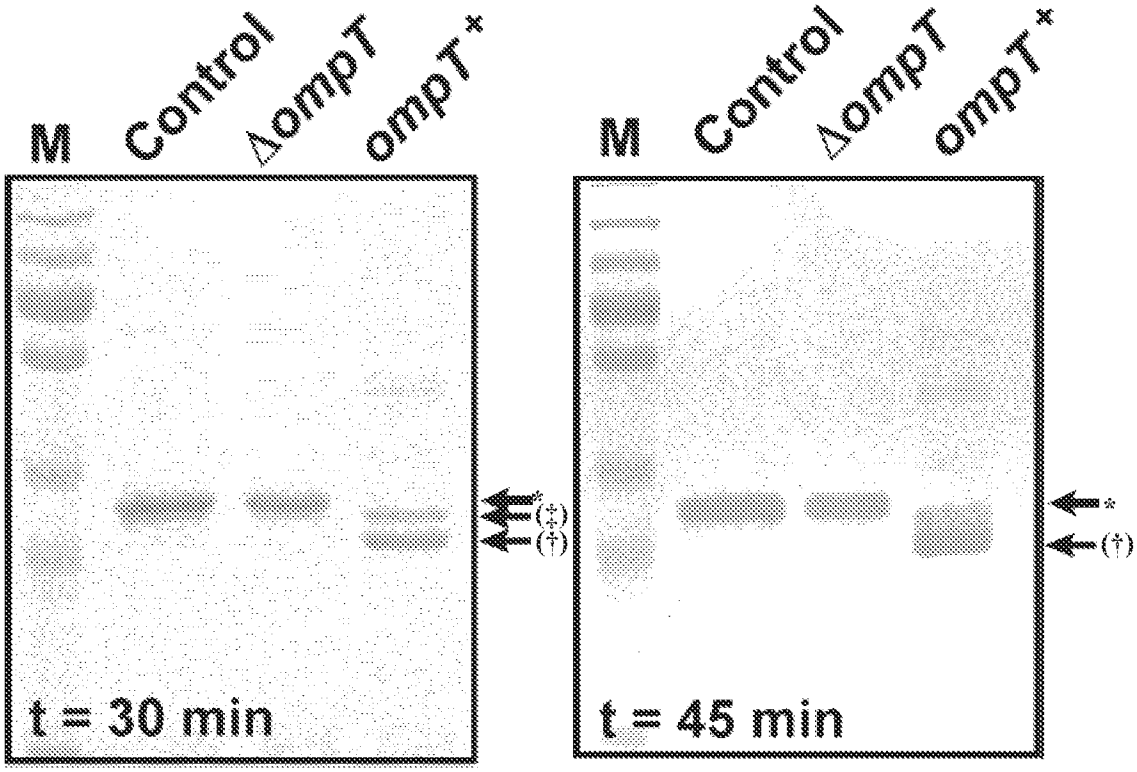


FIG. 9A FIG. 9B

FIG. 10A

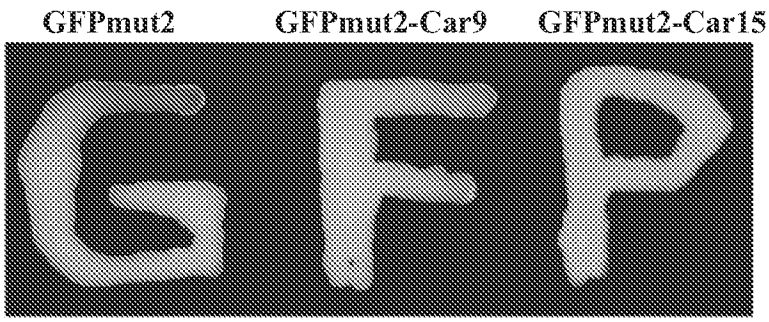


FIG. 10B

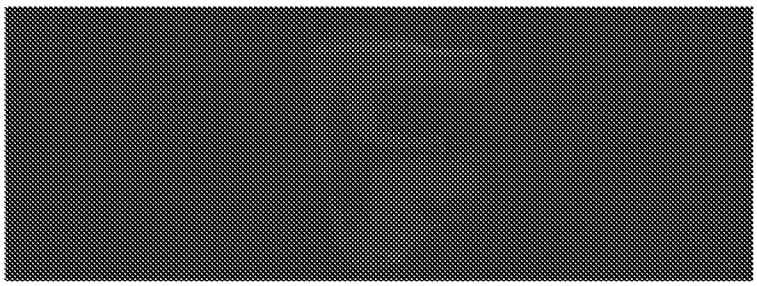
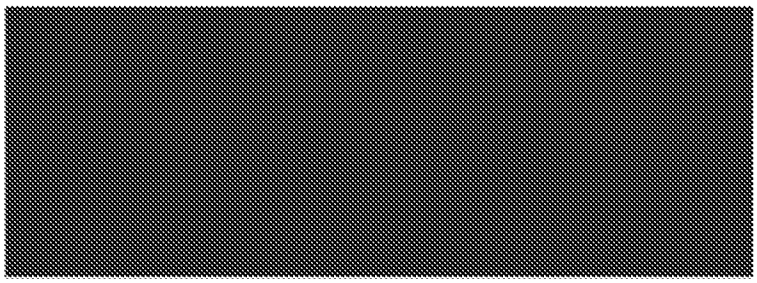


FIG. 10C



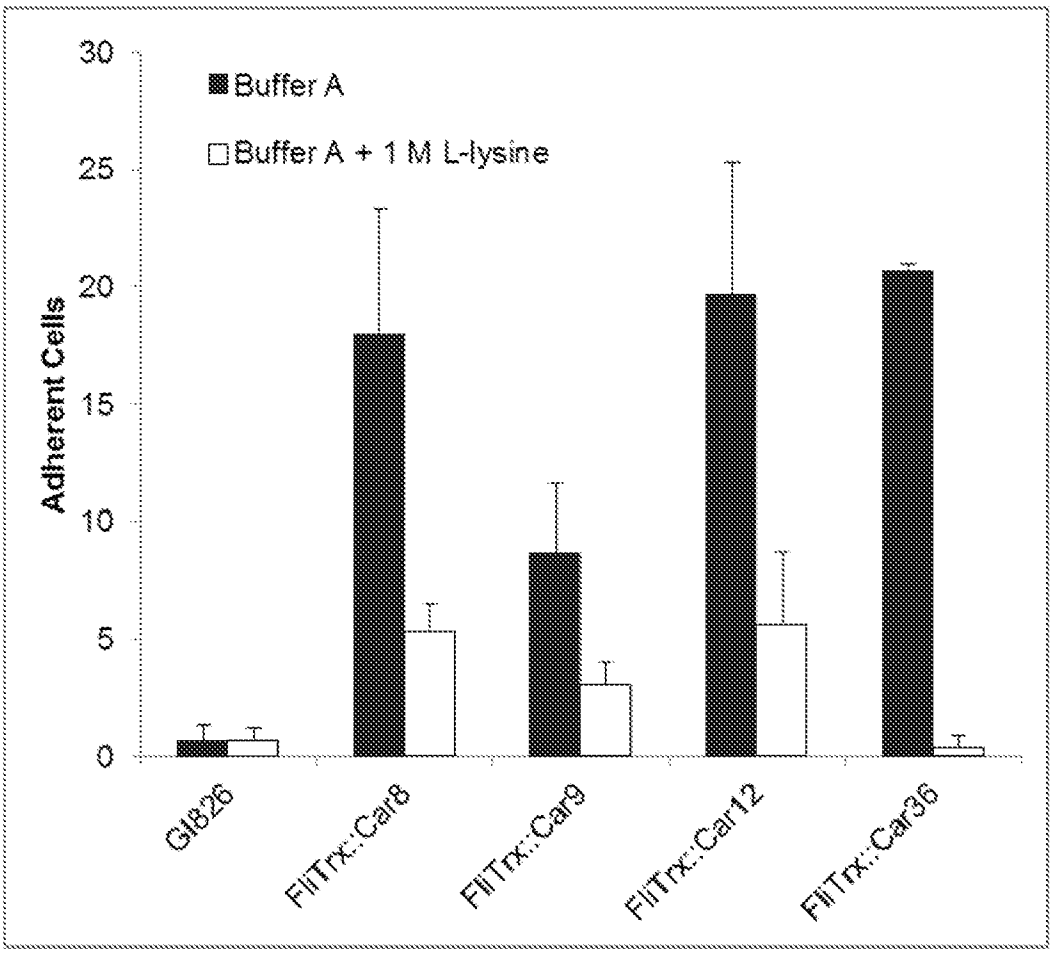

30
25
20
15
10
5
0
Adherent Cells
Buffer A
Buffer A + 1 M L-lysine
GI826
FliTrx::Car8
FliTrx::Car9
FliTrx::Car12
FliTrx::Car36

FIG. 11

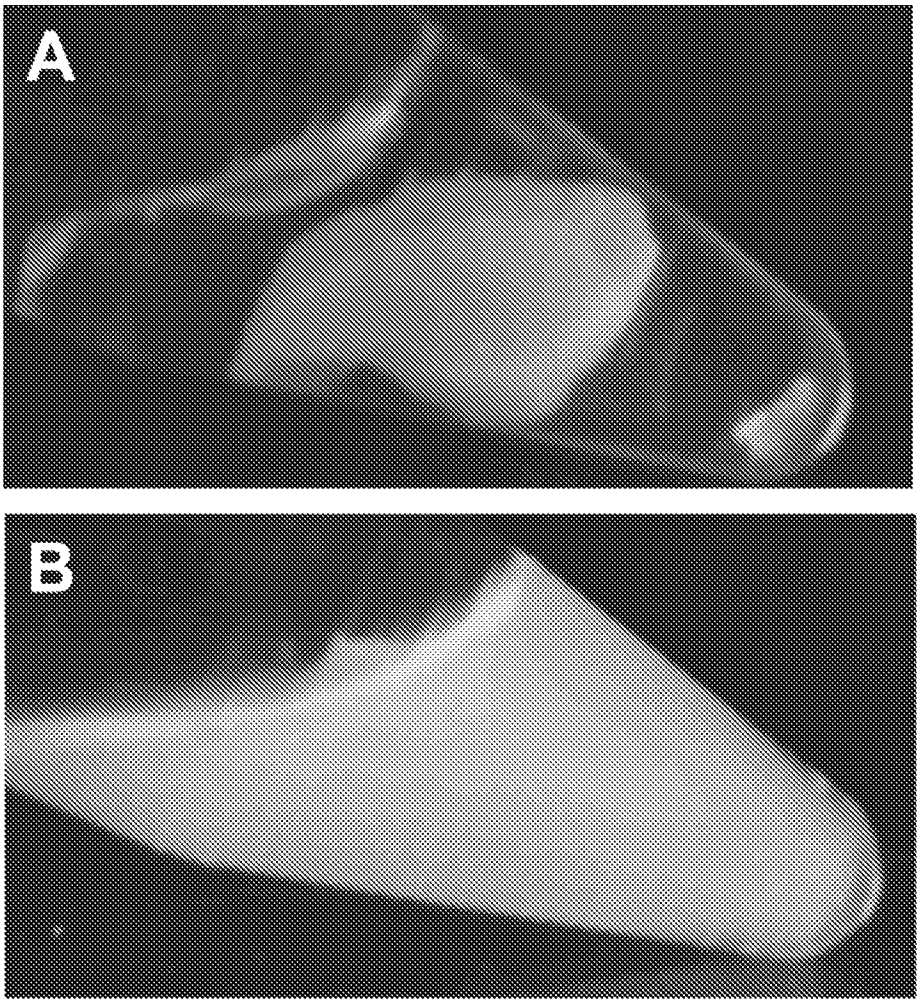


FIG. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/056651

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C07K 1/13 (2014.01)
CPC - C07K 1/13 (2014.12)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - C07K 1/13, 1/22; C08G 77/00, 77/02 (2014.01)
USPC - 436/86; 506/41; 530/300, 413

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
CPC - B01D 15/22, 15/206; C07F 5/022; C07K 1/13, 2319/40; C08G 77/02; G01N 30/463, 30/6052 (2014.12) (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatBase, Google Patents, PubMed

Search terms used:silicon silica borosilicate quartz silica sol gel fusion protein antibody protein purification isolation

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	COYLE et al. "Carbon-binding designer proteins that discriminate between sp2- and sp3-hybridized carbon surfaces," Langmuir. 02 April 2013 (02.04.2013), Vol. 29, Pgs. 4839-46. entire document	1, 5, 6, 8, 11 ---------- 7, 9, 24-27
X	IKEDA et al. "Single-step affinity purification of recombinant proteins using the silica-binding Si-tag as a fusion partner," Protein Expr Purif. 23 December 2009 (23.12.2009), Vol. 71, Pgs. 91-95. entire document	11
Y	MASSARI et al. "Dynamics of proteins encapsulated in silica sol-gel glasses studied with IR vibrational echo spectroscopy," J Am Chem Soc. 29 March 2006 (29.03.2009), Vol. 128, Pgs. 3990-3997. entire document	24-27
Y	US 2013/0115635 A1 (PARDON et al) 09 May 2013 (09.05.2013) entire document	7
A	WO 2007/115860 A1 (BIO-KER SRL) 18 October 2007 (18.10.2007) entire document	8, 9, 24-27
A	GHOSE et al. "Preparative protein purification on underivatized silica," Biotechnol Bioeng. 12 July 2004 (12.07.2004), Vol. 87, Pgs. 413-423. entire document	11
A	US 2005/0233307 A1 (KOREN ANDERSON MOLECULAR PROBES, INC) 20 October 2005 (20.10.2005) entire document	6-7
P, Y	COYLE et al. "Carbon-binding designer proteins that discriminate between sp2- and sp3-hybridized carbon surfaces," Langmuir, 02 April 2013 (02.04.2013), Vol. 29, Pgs. 4839-46.	8, 11, 24

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 December 2014	30 DEC 2014
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/056651

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:

   a. (means)

   ☐ on paper

   ☒ in electronic form

   b. (time)

   ☒ in the international application as filed

   ☐ together with the international application in electronic form

   ☐ subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs:1-4 were searched.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/056651

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 10, 12-23
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.