



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2017/02/10
(87) Date publication PCT/PCT Publication Date: 2017/08/17
(85) Entrée phase nationale/National Entry: 2018/08/10
(86) N° demande PCT/PCT Application No.: EP 2017/053034
(87) N° publication PCT/PCT Publication No.: 2017/137579
(30) Priorité/Priority: 2016/02/12 (US62/294,470)

(51) Cl.Int./Int.Cl. *C07K 16/28* (2006.01),
C12N 15/09 (2006.01), *C12N 5/10* (2006.01),
C12P 21/02 (2006.01)
(71) Demandeur/Applicant:
ABLYNX NV, BE
(72) Inventeurs/Inventors:
SCHOTTE, PETER, BE;
DE GROEVE, MANU, BE
(74) Agent: BORDEN LADNER GERVAIS LLP

(54) Titre : PROCEDE DE PRODUCTION DE DOMAINES VARIABLES UNIQUES D'IMMUNOGLOBULINE
(54) Title: METHOD FOR THE PRODUCTION OF IMMUNOGLOBULIN SINGLE VARIABLE DOMAINS

(57) **Abrégé/Abstract:**

Methods are provided for the manufacture of polypeptides comprising at least one immunoglobulin variable domain that result in an increased yield. The methods are based on simultaneous enhancement of one or more auxiliary proteins in the host.

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

(19) World Intellectual Property Organization

International Bureau

(43) International Publication Date
17 August 2017 (17.08.2017)





(10) International Publication Number
WO 2017/137579 A1

- (51) International Patent Classification:

C07K 16/28 (2006.01) *C12N 5/10* (2006.01)

C12P 21/02 (2006.01) *C12N 15/09* (2006.01)
- (21) International Application Number:

PCT/EP2017/053034
- (22) International Filing Date:

10 February 2017 (10.02.2017)
- (25) Filing Language:

English
- (26) Publication Language:

English
- (30) Priority Data:

62/294,470 12 February 2016 (12.02.2016) US
- (71) Applicant: **ABLYNX NV** [BE/BE]; Technologiepark 21, 9052 Ghent-Zwijnaarde (BE).
- (72) Inventors: **SCHOTTE, Peter**; Hugo Verriestlaan 34, 9840 De Pinte (BE). **DE GROEVE, Manu**; Kwalestraat 85-C, 9320 Nieuwerkerken (BE).
- (74) Agents: **WIEDEMANN, Peter** et al.; Hoffmann Eitle Patent- Und Rechtsanwälte Partmbb, Arabellastraße 30, 81925 Munich (DE).
- (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, DJ, 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, KH, KN, KP, KR, KW, 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).
- Published:
- *with international search report (Art. 21(3))*
 - *with sequence listing part of description (Rule 5.2(a))*

(54) Title: METHOD FOR THE PRODUCTION OF IMMUNOGLOBULIN SINGLE VARIABLE DOMAINS

(57) Abstract: Methods are provided for the manufacture of polypeptides comprising at least one immunoglobulin variable domain that result in an increased yield. The methods are based on simultaneous enhancement of one or more auxiliary proteins in the host.

METHOD FOR THE PRODUCTION OF IMMUNOGLOBULIN SINGLE VARIABLE DOMAINS

FIELD OF THE INVENTION

The present invention relates to a method for the manufacture of immunoglobulin single variable domains. More specifically, the present invention provides an improved method of producing immunoglobulin single variable domains wherein increased yields are obtained. The invention further provides nucleic acids, genetic constructs and host cells for use in the method of the invention as well as immunoglobulin single variable domains obtainable by the method of the invention.

BACKGROUND ART

For therapeutic applications, antibodies or antibody fragments must be of very high product quality. This puts high demands on the production processes for biological therapeutics. The production costs of these therapeutic compounds are strongly influenced by difficulties encountered during the production process. Low yields or lack of homogeneity will impact the economics of the production process, and hence, the costs for the therapeutic, overall.

The limitation of obtaining adequate yields of functional product has been reported for conventional immunoglobulins and their fragments across a broad range of expression systems, including in vitro translation, *E. coli*, yeasts such as e.g. *Saccharomyces cerevisiae* and *Pichia pastoris*, mammalian cells such as e.g. Chinese hamster ovary cells and baculovirus systems in insect cells. Amongst others, a bottleneck for antibody expression appears to be insufficient supply of light chains, inappropriate processing and folding in the endoplasmic reticulum (ER) and intracellular accumulation of heavy chain fragments. (Lange et al. 2001, J. Immunol. Methods 255: 103; Gasser et al. 2006, Biotechnol Bioeng. 94: 353; Gach et al. 2007, J. Biotechnol. 128: 735; Jenkins et al. 2009, Biotechnol. Appl. Biochem. 53: 73).

Intervention in the protein folding and secretory pathways has been described as one of the different strategies for improving the expression and quality of recombinant proteins, such as monoclonal antibodies (MAbs). However, overexpression of one or more components of the ER secretion machinery has yielded mixed results as regards improving

productivity. Many challenges still exist to achieve consistently high yields in biopharmaceutical production. (Jenkins et al. 2009, Biotechnol. Appl. Biochem. 53: 73).

Pichia pastoris has been developed as a host for heterologous protein production. Although this host is known as a highly efficient expression system, especially the production of complex proteins has turned out to have a rather low success rate. In *P. pastoris*, co-overexpression of Immunoglobulin binding protein (BiP) resulted in increased secretion levels of a scFv (A33scFv) by approximately threefold. In contrast, co-overexpression of protein disulfide isomerase (PDI1) had no apparent effect on secretion of A33scFv. Co-overexpressing BiP and PDI1 in *P. pastoris* did not increase A33scFv secretion and protein levels remained the same as the control strain (Damasceno et al. 2007, Appl. Microbiol. Biotechnol. 74: 381). Compared to that of a control strain, 2F5 Fab fragment productivity in *P. pastoris* could be improved ranging from 1.2 fold in the case of co-overexpression with BFR2 to 2.3 fold when SSE1 or KIN2 was overexpressed (Gasser et al. 2007, Appl. Environ. Microbiol. 73: 6499). Overexpression of basic leucine zipper (bZIP) transcription factor HAC1 had a slight effect (1.3 fold) on Mab 2F5 Fab fragment secretion in *P. pastoris*, while overexpression of PDI1 enabled an increase of the Fab level by 1.9 fold (Gasser et al. 2006, Biotechnol Bioeng. 94: 353). The authors conclude that sufficient supply of light chain and the formation of interchain disulfide bonds can be seen as a major rate limiting factors to Fab assembly and subsequent secretion.

In contrast to these difficulties observed with conventional four-chain antibodies or their fragments, including Fabs and scFvs, immunoglobulin single variable domain are known to be readily expressed and secreted from hosts like *E. coli* or *P. pastoris* at a sufficient rate and level. Immunoglobulin single variable domains do not possess interchain disulfide bridges and are characterized by formation of the antigen binding site by a single immunoglobulin variable domain, which does not require interaction with a further domain (e.g. in the form of VH/VL interaction) for antigen recognition. For example, production of Nanobodies, as one specific example of an immunoglobulin single variable domain, in prokaryotic hosts such as *E. coli* has been extensively described (see e.g. Ghahroudi et al. 1997, FEBS Letters 414: 521-526; Muyldermans 2001, J. Biotechnol. 74: 277-302; Vranken et al. 2002, Biochemistry 41: 8570-8579).

Production of Nanobodies in lower eukaryotic hosts such as *P. pastoris* has been described by Frenken et al. 2000 (J. Biotechnol. 78: 11-21), WO 94/25591, WO 2010/125187, WO 2012/056000 and WO 2012/152823.

Without any optimization of conditions, recombinant camelid single variable domains are routinely obtained at levels of 5-10 mg/l when expressed in *E. coli* grown in shaking culture flasks (Ghahroudi et al. 1997). With other expression systems it is even possible to obtain higher yields of VHH expression. Production levels of 9.3 mg/l/OD660 or ~250 mg secreted protein per litre of *Saccharomyces* yeast culture in shake flasks have been described by Frenken et al. 2000. More recently, Nanobody yields of more than 1 g per litre have been described (WO 2010/139808, WO 2012/152823) upon expression in *P. pastoris*.

WO 2010/125187 describes methods for producing a single variable domain in yeast (such as *P. pastoris*). The methods of WO 2010/125187 apply conditions that promote the formation of disulfide bridges in the single variable domain. One of the conditions proposed is enhancing the expression of a thiol isomerase (such as e.g. PDI1).

The fact that fully functional immunoglobulin single variable domains are readily produced in e.g. *E. coli* or yeast at a sufficient rate and level represents an important advantage of this immunoglobulin-format over conventional immunoglobulins.

SUMMARY OF THE INVENTION

Although VHH's and Nanobodies can be expressed using the expression systems described in the art for the expression of the same (WO 1994/25591; Ghahroudi et al. 1997, FEBS Letters 414: 521-526; Frenken et al. 2000, J. Biotechnol. 78: 11-21; Muyldermans 2001, J. Biotechnol. 74: 277-302; Vranken et al. 2002, Biochemistry 41: 8570-8579; WO 2010/125187; WO 2012/056000; WO 2012/152823 and other patent applications by Ablynx N.V.), the present inventors have found that in some cases (e.g. multivalent VHHs and Nanobodies and/or VHH's and Nanobodies with more than one disulfide bridge), the expression of VHH's and Nanobodies is more difficult leading to expression levels and/or yields much lower than expected. For example, the inventors have unexpectedly observed problems with the production of some therapeutic VHH1 type immunoglobulin single variable domains. Upon expression of these immunoglobulin single variable domains in *P. pastoris*, the present inventors obtained much lower yields of these immunoglobulin single variable domains compared to the yields normally obtained for VHH2 type and VHH3 type

immunoglobulin single variable domains. Contrary to what was established for immunoglobulin single variable domains, expression of certain immunoglobulin single variable domains, such as e.g. VHH1 type immunoglobulin single variable domains, in *P. pastoris* did not result in high amounts of functional product.

The inventors have provided a solution to this problem which is as set out further herein. In addition, they observed that the proposed solution to improve the yield of immunoglobulin single variable domains is generally applicable, i.e. not only to improve the yield of immunoglobulin single variable domains of the VHH1 type, but also to other immunoglobulin single variable domains, such as e.g. immunoglobulin single variable domains of the VHH2 and VHH3 type.

Hence, in one aspect the present invention relates to the observation of the low yields upon expression in *P. pastoris* of certain immunoglobulin single variable domains.

In a further aspect of the present invention, methods are provided for the production of immunoglobulin single variable domains, wherein the yield of the obtained product is increased. These methods are also referred to herein as “method(s) of the invention”. The present invention thus also provides methods of producing immunoglobulin single variable domains which overcome this unexpected problem. More in particular, the present inventors have found (by screening a library of auxiliary proteins) that enhancing the expression, in a *Pichia* host, of certain auxiliary proteins (PDI1 (SEQ ID NO: 5), Kar2p (SEQ ID NO: 4), RPP0 (SEQ ID NO: 6) and, in particular, HAC1spliced (SEQ ID NO: 14)) resulted in an increased yield of the immunoglobulin single variable domain when expressed in said *Pichia* host. Yield increases of more than 2 fold to more than 10 fold (or even more) were obtained.

In one aspect therefore, the present invention relates to a method for increasing the expression and/or production yield of immunoglobulin single variable domains in *Pichia* (such as *P. pastoris*). In the method of the invention, the immunoglobulin single variable domain is expressed while simultaneously the expression of HAC1spliced is enhanced.

The immunoglobulin single variable domains used in the method of the present invention may form part of a polypeptide (also referred to as “*polypeptide of the invention*”), which may comprise or essentially consist of one or more (i.e. at least one) immunoglobulin single variable domains and which may optionally further comprise one or more further amino acid sequences (all optionally linked via one or more suitable linkers).

Thus, the present invention provides methods for producing, in a *Pichia* host (such as *P. pastoris*), a polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain (referred to herein as “polypeptide of the invention”), said method comprising the step of expressing, in the *Pichia* host, the polypeptide of the invention and enhancing, in said *Pichia* host, the expression of HAC1spliced protein. The method of the invention may further comprise the step of isolating and/or purifying the polypeptide of the invention.

The method of the present invention is especially suited for immunoglobulin single variable domains and polypeptides of the invention that are not easily expressible in *Pichia*, such as *P. pastoris*, or that are expressed with very low yields, when expressed under standard conditions (as further defined herein); and for which thus, upon expression of these immunoglobulin single variable domains or polypeptides of the invention in *Pichia*, such as *P. pastoris*, sufficient amounts of immunoglobulin single variable domain and/or polypeptide cannot be obtained. Accordingly, in one aspect, the immunoglobulin single variable domain and/or polypeptide of the invention is selected from the immunoglobulin single variable domains and/or polypeptides for which, upon expression in a *Pichia* host (such as *P. pastoris*) under standard *Pichia* expression conditions (as further defined herein), a yield is obtained of 0.5 g/l or lower, such as 0.4 g/L, 0.3 g/L, 0.2 g/L, 0.1 g/L, 0.05 g/L, 0.01 g/L, or even less. In another aspect, the immunoglobulin single variable domain and/or polypeptide of the invention is selected from the immunoglobulin single variable domains and/or polypeptides for which, upon expression in a *Pichia* host (such as *P. pastoris*) under *Pichia* expression conditions (as further defined herein), the yield obtained shows an inverse correlation (as further defined herein) with the copy number of the nucleic acid encoding said immunoglobulin single variable domain and/or polypeptide.

In a specific aspect of the invention, the method of the present invention is especially suited for immunoglobulin single variable domains and polypeptides of the invention that are easily expressible in *Pichia*, such as *P. pastoris*. Accordingly, in one aspect, the immunoglobulin single variable domain and/or polypeptide of the invention is selected from the immunoglobulin single variable domains and/or polypeptides for which, upon expression in a *Pichia* host (such as *P. pastoris*) under standard *Pichia* expression conditions (as further defined herein), a yield is obtained of 0.5 g/l or higher, such as 1 g/L, 2 g/L, 3 g/L, 4 g/L, 5 g/L, 6 g/L, or even higher.

In the above method, the polypeptide of the invention may comprise or essentially consist of two or more immunoglobulin single variable domains. Such polypeptides are also referred to as multivalent polypeptides. Accordingly, in a specific aspect, the polypeptide of the invention to be expressed by the method of the invention, is a multivalent polypeptide.

Alternatively, the polypeptide expressed by the method of the invention may comprise or essentially consist of one immunoglobulin single variable domain. Such polypeptide will also be referred to herein as a monovalent polypeptide. Accordingly, in another specific aspect, the polypeptide of the invention to be expressed by the method of the invention, is a monovalent polypeptide.

In the above method, the immunoglobulin single variable domain (potentially present in the polypeptide of the invention) can be (without being limited) an immunoglobulin single variable domain that is a light chain variable domain or a heavy chain variable domain, more specifically an immunoglobulin single variable domain which is a heavy chain variable domain that is derived from a conventional four-chain antibody or a heavy chain variable domain that is derived from a heavy chain antibody, in particular a domain antibody (or an amino acid sequence that is suitable for use as a domain antibody), a single domain antibody (or an amino acid sequence that is suitable for use as a single domain antibody), a “dAb” (or an amino acid sequence that is suitable for use as a dAb) or a Nanobody (including but not limited to a VHH sequence), preferably a Nanobody.

In a preferred aspect, the Nanobody expressed in the method of the invention is a VHH sequence (VHH), a (partially) humanized VHH sequence (humanized VHH), a camelized heavy chain variable domain (camelized VH) or a Nanobody, that has been obtained by techniques such as affinity maturation.

In another specific aspect of the invention, the immunoglobulin single variable domain (potentially present in the polypeptide of the invention), expressed in the method of the invention, comprises at least two disulfide bridges. In another specific aspect, the immunoglobulin single variable domain belongs to the group of VHH1 type immunoglobulin single variable domains.

In another aspect, the immunoglobulin single variable domain (potentially present in the polypeptide of the invention), does not belong to the group of VHH1 type immunoglobulin single variable domains, but belongs to another group of immunoglobulin

single variable domains, such as the VHH2, VHH3 or any other type immunoglobulin single variable domains.

In yet another specific aspect, the immunoglobulin single variable domain (potentially present in the polypeptide of the invention), expressed in the method of the invention, specifically binds c-Met, such as e.g. the immunoglobulin single variable domains described in WO 2012/042026 and WO 2013/045707. Thus, polypeptides of the invention that comprise at least one immunoglobulin single variable domain that specifically binds c-Met, wherein said immunoglobulin single variable domain comprises two disulfide bridges, form a specific but non-limiting aspect of the invention. A particular Nanobody for use in the method of the invention is SEQ ID NO: 49.

In a preferred aspect, the one or more immunoglobulin single variable domain(s) (potentially present in the polypeptide of the invention), expressed in the method of the invention, specifically bind TNF, such as e.g. the immunoglobulin single variable domains described in US provisional application US 62/254,375 of Ablynx NV (see also PCT/EP2016/077595).

Thus, polypeptides of the invention that comprise one or more (at least one) immunoglobulin single variable domain that specifically bind(s) TNF form a specific and non-limiting aspect of the invention.

In a specific aspect, the polypeptide of the invention, expressed in the method of the invention, comprises or essentially consists of one immunoglobulin single variable domain. Such polypeptide will also be referred to herein as a monovalent polypeptide. Therefore, in a preferred aspect, in the method of the invention, the polypeptide is a monovalent polypeptide.

Such a monovalent polypeptide, comprising or essentially consisting of one immunoglobulin single variable domain, may comprise at least two disulfide bridges. Accordingly, the immunoglobulin single variable domain that specifically binds TNF belongs to the VHH1 type immunoglobulin single variable domains.

Alternatively, the immunoglobulin single variable domain comprises one disulfide bridge. Accordingly, in a preferred aspect, the immunoglobulin single variable domain that specifically binds TNF does not belong to the group of VHH1 type immunoglobulin single variable domains, but belongs to another group of immunoglobulin single variable domains, such as the VHH2, VHH3 or any other type immunoglobulin single variable domains.

In a preferred aspect, in the method of the present invention, the immunoglobulin single variable domain essentially consists of 4 framework regions (FR1 to FR4, respectively) and 3 complementarity determining regions (CDR1 to CDR3, respectively), in which CDR1 is SEQ ID NO: 58, CDR2 is SEQ ID NO: 60 and CDR3 is SEQ ID NO: 62.

Particular Nanobodies for use in the method of the invention are chosen from the group consisting of SEQ ID NO's: 55 and 56.

The polypeptide of the invention comprising or essentially consisting of one or more immunoglobulin single variable domains may further comprise one or more other residues or binding units, optionally linked via one or more peptidic linkers.

In one aspect, said one or more other residues may be effective in preventing or reducing binding of antibodies pre-existing in the serum (so-called "pre-existing antibodies") to the polypeptides of the invention (as is further described herein).

In another aspect, said one or more other residues or binding units may also be chosen from the group consisting of immunoglobulin single variable domains, domain antibodies, amino acid sequences that are suitable for use as a domain antibody, single domain antibodies, amino acid sequences that are suitable for use as a single domain antibody, "dAb"'s, amino acid sequences that are suitable for use as a dAb, or Nanobodies. Polypeptides comprising or essentially consisting of two or more binding units are also referred to as multivalent constructs.

In a specific aspect of the invention, the polypeptide comprising or essentially consisting of one or more immunoglobulin single variable domains expressed in the method of the invention is a multivalent construct.

In another specific aspect, the polypeptide of the invention is a bivalent, trivalent or tetravalent polypeptide.

In a particular aspect of the invention, said one or more other binding units may provide the polypeptide of the invention with increased half-life, compared to the polypeptide without said one or more binding units. Without being limiting, said one or more other binding units that provides the polypeptide with increased half-life may be chosen from the group consisting of binding units that can bind to serum albumin (such as human serum albumin) or a serum immunoglobulin (such as IgG).

In the method of the invention, the expression of HAC1spliced may be enhanced by introduction, into the *Pichia* host, of one or more nucleic acid(s) encoding HAC1spliced

protein. In another aspect, the expression of HAC1spliced protein may be enhanced by introduction, into the *Pichia* host, of one or more strong promoter(s) controlling the expression of a nucleic acid encoding HAC1spliced protein.

The polypeptide comprising or essentially consisting of the at least one immunoglobulin single variable domain and the HAC1spliced protein may be expressed from the same genetic construct. In this particular aspect of the invention, transcription of the nucleic acid encoding the polypeptide comprising or essentially consisting of the at least one immunoglobulin single variable domain and transcription of the nucleic acid encoding the HAC1spliced protein may be controlled by the same promoter or by a different promoter. The nucleic acid encoding the HAC1spliced protein may be located on the genetic construct downstream of the nucleic acid encoding the polypeptide comprising or essentially consisting of the at least one immunoglobulin single variable domain. In the alternative, the nucleic acid encoding the polypeptide comprising or essentially consisting of the at least one immunoglobulin single variable domain may be located on the genetic construct downstream of the nucleic acid encoding HAC1spliced protein.

In another aspect of the invention, the polypeptide comprising or essentially consisting of the at least one immunoglobulin single variable domain and the HAC1spliced protein may be expressed from different genetic constructs. In this particular aspect of the invention, transcription of the nucleic acid encoding the polypeptide comprising or essentially consisting of the at least one immunoglobulin single variable domain and transcription of the nucleic acid encoding the HAC1spliced protein from the different genetic constructs may be controlled by two separate promoters which may be the same or different.

The polypeptide comprising or essentially consisting of the at least one immunoglobulin single variable domain and/or the HAC1spliced protein can also be expressed from the chromosome. In this particular aspect, the expression of HAC1spliced protein may be enhanced by introduction of a strong promoter into the chromosome of the *Pichia* host. In the alternative, one or more nucleic acid(s) encoding HAC1spliced protein controlled by a strong promoter may be introduced into the chromosome of the *Pichia* host. Also one or more nucleic acids encoding the polypeptide comprising or essentially consisting of the at least one immunoglobulin single variable domain may be introduced into the chromosome of the *Pichia* host.

In one aspect of the invention, the number of the nucleic acid(s) encoding the polypeptide comprising or essentially consisting of an immunoglobulin single variable domain is one. In another aspect of the invention the number of the nucleic acid(s) encoding the polypeptide comprising or essentially consisting of an immunoglobulin single variable domain is two or more.

In another aspect, the invention also relates to a nucleic acid that encodes a polypeptide of the invention (or a suitable fragment thereof) and that encodes HAC1spliced protein. Such a nucleic acid will also be referred to herein as a “*nucleic acid of the invention*” and may for example be in the form of a genetic construct, as further described herein.

Accordingly, the present invention also relates to a nucleic acid of the invention that is in the form of a genetic construct. Such a genetic construct will also be referred to herein as a “*genetic construct of the invention*” and thus comprises a nucleic acid encoding a polypeptide of the invention and a nucleic acid encoding the HAC1spliced protein. In a specific aspect, in such a genetic construct of the invention, the nucleic acid encoding the HAC1spliced protein is located downstream of the nucleic acid encoding the polypeptide of the invention. In another specific aspect, in such a genetic construct of the invention, the nucleic acid encoding the polypeptide of the invention is located downstream of the nucleic acid encoding the HAC1spliced protein. In the genetic construct of the invention, expression of the nucleic acid encoding the polypeptide of the invention and expression of the nucleic acid encoding the HAC1spliced protein may be controlled by the same promoter or by a different promoter. The promoter may be a constitutive or an inducible promoter.

In one aspect of the invention, the copy number of the nucleic acid encoding the polypeptide of the invention is one. In another aspect of the invention the copy number of the nucleic acid encoding the polypeptide of the invention is two or more.

In one aspect of the invention, the number of auxiliary proteins is one, wherein the auxiliary protein is HAC1spliced. In another aspect of the invention, the expression of one or more additional auxiliary proteins is enhanced. In yet another aspect of the invention, the additional auxiliary protein(s) is(are) selected from PDI1, Kar2p and RPP0. In yet another aspect of the invention the number of auxiliary proteins is two. In yet another aspect of the invention, the number of auxiliary proteins is more than two, such as three or more. In yet another aspect, the two or more auxiliary proteins are selected from the following combination of auxiliary proteins:

- PDI1 and HAC1spliced;
- Kar2p and HAC1spliced;
- RPP0 and HAC1spliced;
- PDI1, Kar2p and HAC1spliced;
- PDI1, RPP0 and HAC1spliced;
- Kar2p, RPP0 and HAC1spliced; and
- PDI1, Kar2p, RPP0 and HAC1spliced.

In a preferred aspect, expression of following auxiliary proteins is enhanced:

- PDI1, Kar2p and Hac1spliced; or
- Kar2p, RPP0 and Hac1spliced.

In a preferred aspect of the invention, the auxiliary protein is HAC1spliced (SEQ ID NO: 14).

In another preferred aspect, an additional auxiliary protein is selected from PDI1 (SEQ ID NO: 5), Kar2p (SEQ ID NO: 4), RPP0 (SEQ ID NO: 6) and Hac1spliced (SEQ ID NO: 14).

In another preferred aspect, the two or more auxiliary proteins are selected from the following combination of auxiliary proteins:

- PDI1 (SEQ ID NO: 5) and Hac1spliced (SEQ ID NO: 14);
- Kar2p (SEQ ID NO: 4) and Hac1spliced (SEQ ID NO: 14);
- RPP0 (SEQ ID NO: 6) and Hac1spliced (SEQ ID NO: 14);
- PDI1 (SEQ ID NO: 5), Kar2p (SEQ ID NO: 4) and Hac1spliced (SEQ ID NO: 14);
- PDI1 (SEQ ID NO: 5), RPP0 (SEQ ID NO: 6) and Hac1spliced (SEQ ID NO: 14);
- Kar2p (SEQ ID NO: 4), RPP0 (SEQ ID NO: 6) and Hac1spliced (SEQ ID NO: 14); and
- PDI1 (SEQ ID NO: 5), Kar2p (SEQ ID NO: 4), RPP0 (SEQ ID NO: 6) and Hac1spliced (SEQ ID NO: 14).

In another aspect, the present invention relates to the introduction of the nucleic acid and genetic construct of the invention in a *Pichia* host, also referred to herein as “*Pichia* host of the invention”. In addition to plasmid or vector transformation of the *Pichia* host, chromosomal transformation of the *Pichia* host is also encompassed by the present invention. A strong (inducible) promoter (instead of the native promoter of a native auxiliary protein) may be introduced on the chromosome of the *Pichia* host; or another copy of an auxiliary protein gene sequence under control of another (strong) promoter may be introduced into the chromosome.

In another aspect, the invention relates to a *Pichia* host that expresses (or that under suitable circumstances is capable of expressing) a polypeptide of the invention and wherein the expression of HAC1spliced is enhanced; and/or that contains a nucleic acid of the invention and/or a genetic construct of the invention.

In a preferred aspect, the *Pichia* host is *Pichia pastoris*.

In another preferred aspect, the *Pichia pastoris* strain is selected from *Pichia pastoris* X33 and *Pichia pastoris* NRRL Y-11430.

The invention further relates to methods for the preparation of the nucleic acids of the invention, the genetic constructs of the invention and the *Pichia* hosts of the invention; and to the uses of said nucleic acids of the invention, genetic constructs of the invention and *Pichia* hosts of the invention for the production of immunoglobulin single variable domains and polypeptides comprising the same.

The invention further relates to a polypeptide and/or immunoglobulin single variable domain obtainable by any of the methods as set forth herein, pharmaceutical compositions and other compositions comprising such polypeptides and/or an immunoglobulin single variable domains, and therapeutic uses of the polypeptides and/or an immunoglobulin single variable domains or methods of treatment comprising the use of the polypeptides and/or an immunoglobulin single variable domains.

DESCRIPTION OF THE FIGURES

Figure 1: Nanobody A was expressed in two different *Pichia* clones generated as described in Example 1.2. One *Pichia* clone contained 1 copy of the Nanobody A expression cassette in the genome. The second clone contained more than one copy of the Nanobody A expression cassette inserted in the genome. Equal volumes of supernatant were compared from the different clones on SDS-PAGE gel. Densitometry analysis for relative quantification of the bands corresponding to intact Nanobody product was performed. Quantification of band volumes was done using Imagequant software (GE Healthcare). An inverse correlation between copy number and yield was observed.

Figure 2: Clones transformed with Nanobody A and the auxiliary protein library were tested for improved expression levels of Nanobody A. Equal volumes of supernatant were compared from the different clones on SDS-PAGE gel. Densitometry analysis for relative

quantification of the bands corresponding to intact Nanobody product was performed. Quantification of band volumes was done using Imagequant software (GE Healthcare).

4 clones (6H1, 4C2, 5A6, 9C4) were found to secrete significantly higher levels of Nanobody A in shake flask compared to their corresponding reference clones without auxiliary protein (Ref CN=1 and Ref CN>1). Ref: reference clone with indicated copy number of the Nanobody A expression cassette inserted in the genome.

Figure 3: Shake flask expression of the reference clone with more than 1 copy of the Nanobody A expression cassette transformed with one of the auxiliary proteins Kar2p, RPP0, Hac1 splice variant or PDI1. Nanobody yields were analysed on SDS-PAGE and compared to the Nanobody yield by the reference clone (without auxiliary protein). Densitometry analysis for relative quantification of the bands corresponding to intact Nanobody product was performed. Quantification of band volumes was done using Imagequant software (GE Healthcare).

DETAILED DESCRIPTION OF THE INVENTION

Definitions

Unless indicated or defined otherwise, all terms used have their usual meaning in the art, which will be clear to the skilled person. Reference is for example made to the standard handbooks, such as Sambrook et al. 1989 (Molecular Cloning: A Laboratory Manual, 2nd Ed., Vols. 1-3, Cold Spring Harbor Laboratory Press), Ausubel et al. 1987 (Current protocols in molecular biology, Green Publishing and Wiley Interscience, New York), Lewin 1985 (Genes II, John Wiley & Sons, New York, N.Y.), Old et al. 1981 (Principles of Gene Manipulation: An Introduction to Genetic Engineering, 2nd Ed., University of California Press, Berkeley, CA), Roitt et al. 2001 (Immunology, 6th Ed., Mosby/Elsevier, Edinburgh), Roitt et al. 2001 (Roitt's Essential Immunology, 10th Ed., Blackwell Publishing, UK), and Janeway et al. 2005 (Immunobiology, 6th Ed., Garland Science Publishing/Churchill Livingstone, New York), as well as to the general background art cited herein.

Unless indicated otherwise, all methods, steps, techniques and manipulations that are not specifically described in detail can be performed and have been performed in a manner known per se, as will be clear to the skilled person. Reference is for example again made to the standard handbooks and the general background art mentioned herein and to the

further references cited therein; as well as to for example the following reviews Presta 2006 (Adv. Drug Deliv. Rev. 58: 640), Levin and Weiss 2006 (Mol. Biosyst. 2: 49), Irving et al. 2001 (J. Immunol. Methods 248: 31), Schmitz et al. 2000 (Placenta 21 Suppl. A: S106), Gonzales et al. 2005 (Tumour Biol. 26: 31), which describe techniques for protein engineering, such as affinity maturation and other techniques for improving the specificity and other desired properties of proteins such as immunoglobulins.

When a nucleotide sequence or amino acid sequence is said to “comprise” another nucleotide sequence or amino acid sequence, respectively, or to “essentially consist of” another nucleotide sequence or amino acid sequence, this may mean that the latter nucleotide sequence or amino acid sequence has been incorporated into the first mentioned nucleotide sequence or amino acid sequence, respectively, but more usually this generally means that the first mentioned nucleotide sequence or amino acid sequence comprises within its sequence a stretch of nucleotides or amino acid residues, respectively, that has the same nucleotide sequence or amino acid sequence, respectively, as the latter sequence, irrespective of how the first mentioned sequence has actually been generated or obtained (which may for example be by any suitable method described herein). By means of a non-limiting example, when a polypeptide of the invention is said to comprise an immunoglobulin single variable domain, this may mean that the immunoglobulin single variable domain sequence has been incorporated into the sequence of the polypeptide of the invention, but more usually this generally means that the polypeptide of the invention contains within its sequence the sequence of the immunoglobulin single variable domains irrespective of how said polypeptide of the invention has been generated or obtained. Also, when a nucleic acid or nucleotide sequence is said to comprise another nucleotide sequence, the first mentioned nucleic acid or nucleotide sequence is preferably such that, when it is expressed into an expression product (e.g. a polypeptide), the amino acid sequence encoded by the latter nucleotide sequence forms part of said expression product (in other words, that the latter nucleotide sequence is in the same reading frame as the first mentioned, larger nucleic acid or nucleotide sequence).

By “essentially consist of” is meant that the immunoglobulin single variable domain used in the method of the invention either is exactly the same as the polypeptide of the invention or corresponds to the polypeptide of the invention which has a limited number of amino acid residues, such as 1-20 amino acid residues, for example 1-10 amino acid residues

and preferably 1-6 amino acid residues, such as 1, 2, 3, 4, 5 or 6 amino acid residues, added at the amino terminal end, at the carboxy terminal end, or at both the amino terminal end and the carboxy terminal end of the immunoglobulin single variable domain.

A nucleic acid or amino acid is considered to be “(in) (essentially) isolated (form)” - for example, compared to the reaction medium or cultivation medium from which it has been obtained - when it has been separated from at least one other component with which it is usually associated in said source or medium, such as another nucleic acid, another protein/polypeptide, another biological component or macromolecule or at least one contaminant, impurity or minor component. In particular, a nucleic acid or amino acid is considered “(essentially) isolated” when it has been purified at least 2-fold, in particular at least 10-fold, more in particular at least 100-fold, and up to 1000-fold or more. A nucleic acid or amino acid that is “in (essentially) isolated form” is preferably essentially homogeneous, as determined using a suitable technique, such as a suitable chromatographical technique, such as polyacrylamide-gel electrophoresis.

The terms “expression (of)” or “expressing” a polypeptide such as an immunoglobulin single variable domain and/or polypeptide of the invention or an auxiliary protein is the process by which information from a gene is used in the synthesis of a functional gene product (i.e. the immunoglobulin single variable domain and/or polypeptide of the invention or the auxiliary protein). When an auxiliary protein or a polypeptide such as an immunoglobulin single variable domain and/or polypeptide of the invention is said to be “expressed from” a nucleic acid, genetic construct or chromosome, it is synthesized through a process by which information from a nucleic acid, genetic construct or chromosome is used for the synthesis of the functional gene product (i.e. the immunoglobulin single variable domain and/or polypeptide of the invention or the auxiliary protein). When proteins are said to be “co-expressed”, said proteins are expressed simultaneously. “Enhancing the expression” of a gene means that the production of the gene product (e.g. the auxiliary protein) is increased compared to the production of the gene product without enhancing the expression of the gene. As is further explained, expression of a particular gene can be enhanced by various means including the use of suitable control sequences, e.g. a strong promoter, and/or increasing the gene dose, e.g. by increasing the copy number of the respective gene.

The term “yield” as used in the present invention, relates to the amount of immunoglobulin single variable domain and/or polypeptide of the invention being produced in functional form upon expression in the *Pichia* host. The yield is expressed as g (gram) immunoglobulin single variable domain and/or polypeptide of the invention per L (litre) medium.

Immunoglobulin single variable domain

Unless indicated otherwise, the term “immunoglobulin sequence” - whether used herein to refer to a heavy chain antibody or to a conventional 4-chain antibody - is used as a general term to include both the full-size antibody, the individual chains thereof, as well as all parts, domains or fragments thereof (including but not limited to antigen-binding domains or fragments such as V_{HH} domains or V_H/V_L domains, respectively). In addition, the term “sequence” as used herein (for example in terms like “immunoglobulin sequence”, “antibody sequence”, “variable domain sequence”, “ V_{HH} sequence” or “protein sequence”), should generally be understood to include both the relevant amino acid sequence as well as nucleic acids or nucleotide sequences encoding the same, unless the context requires a more limited interpretation.

The term “immunoglobulin single variable domain”, interchangeably used with “single variable domain”, defines molecules wherein the antigen binding site is present on, and formed by, a single immunoglobulin domain. This sets immunoglobulin single variable domains apart from “conventional” immunoglobulins or their fragments (such as Fabs, scFvs, etc.), wherein two immunoglobulin domains, in particular two variable domains, interact to form an antigen binding site. Typically, in conventional immunoglobulins, a heavy chain variable domain (V_H) and a light chain variable domain (V_L) interact to form an antigen binding site. In this case, the complementarity determining regions (CDRs) of both V_H and V_L will contribute to the antigen binding site, i.e. a total of 6 CDRs will be involved in antigen binding site formation.

In contrast, the binding site of an immunoglobulin single variable domain is formed by a single V_H or V_L domain. Hence, the antigen binding site of an immunoglobulin single variable domain is formed by no more than three CDRs.

The term “immunoglobulin single variable domain” and “single variable domain” hence does not comprise conventional immunoglobulins or their fragments which require

interaction of at least two variable domains for the formation of an antigen binding site. However, these terms do comprise fragments of conventional immunoglobulins wherein the antigen binding site is formed by a single variable domain.

Generally, single variable domains will be amino acids that essentially consist of 4 framework regions (FR1 to FR4 respectively) and 3 complementarity determining regions (CDR1 to CDR3 respectively); or any suitable fragment of such an amino acid (which will then usually contain at least some of the amino acid residues that form at least one of the CDR's). Such single variable domains and fragments are most preferably such that they comprise an immunoglobulin fold or are capable for forming, under suitable conditions, an immunoglobulin fold. As such, the single variable domain may for example comprise a light chain variable domain sequence (e.g. a VL-sequence) or a suitable fragment thereof; or a heavy chain variable domain sequence (e.g. a VH-sequence or VHH sequence) or a suitable fragment thereof; as long as it is capable of forming a single antigen binding unit (i.e. a functional antigen binding unit that essentially consists of the single variable domain, such that the single antigen binding domain does not need to interact with another variable domain to form a functional antigen binding unit, as is for example the case for the variable domains that are present in for example conventional antibodies and scFv fragments that need to interact with another variable domain – e.g. through a VH/VL interaction – to form a functional antigen binding domain).

In one embodiment of the invention, the immunoglobulin single variable domains are light chain variable domain sequences (e.g. a VL-sequence), or heavy chain variable domain sequences (e.g. a VH-sequence); more specifically, the immunoglobulin single variable domains can be heavy chain variable domain sequences that are derived from a conventional four-chain antibody or heavy chain variable domain sequences that are derived from a heavy chain antibody.

For example, the single variable domain or immunoglobulin single variable domain (or an amino acid that is suitable for use as an immunoglobulin single variable domain) may be a (single) domain antibody (or an amino acid that is suitable for use as a (single) domain antibody), a “dAb” or dAb (or an amino acid that is suitable for use as a dAb) or a Nanobody (as defined herein, and including but not limited to a VHH); other single variable domains, or any suitable fragment of any one thereof. For a general description of (single) domain antibodies, reference is also made to the prior art cited herein, as well as to EP 0368684. For

the term “dAb’s”, reference is for example made to Ward et al. 1989 (Nature 341: 544), to Holt et al. 2003 (Trends Biotechnol. 21: 484); as well as to for example WO 04/068820, WO 06/030220, WO 06/003388 and other published patent applications of Domantis Ltd. It should also be noted that, although less preferred in the context of the present invention because they are not of mammalian origin, single variable domains can be derived from certain species of shark (for example, the so-called “IgNAR domains”, see for example WO 05/18629).

In particular, the immunoglobulin single variable domain may be a Nanobody (as defined herein) or a suitable fragment thereof. *[Note: Nanobody, Nanobodies and Nanoclone are registered trademarks of Ablynx N.V.]* For a general description of Nanobodies, reference is made to the prior art cited herein, such as e.g. described in WO 08/020079 (page 16).

The amino acid sequence and structure of an immunoglobulin sequence, in particular a Nanobody can be considered - without however being limited thereto - to be comprised of four framework regions or “FR’s”, which are referred to in the art and herein as “Framework region 1” or “FR1”; as “Framework region 2” or “FR2”; as “Framework region 3” or “FR3”; and as “Framework region 4” or “FR4”, respectively; which framework regions are interrupted by three complementary determining regions or “CDR’s”, which are referred to in the art as “Complementarity Determining Region 1” or “CDR1”; as “Complementarity Determining Region 2” or “CDR2”; and as “Complementarity Determining Region 3” or “CDR3”, respectively.

The total number of amino acid residues in a Nanobody can be in the region of 110-120, is preferably 112-115, and is most preferably 113. It should however be noted that parts, fragments, analogs or derivatives (as further described herein) of a Nanobody are not particularly limited as to their length and/or size, as long as such parts, fragments, analogs or derivatives meet the further requirements outlined herein and are also preferably suitable for the purposes described herein.

For a further description of VHH’s and Nanobodies, reference is made to the review article by Muyldermans 2001 (Rev. Mol. Biotechnol. 74: 277), as well as to the following patent applications, which are mentioned as general background art: WO 94/04678, WO 95/04079 and WO 96/34103 of the Vrije Universiteit Brussel; WO 94/25591, WO 99/37681, WO 00/40968, WO 00/43507, WO 00/65057, WO 01/40310, WO 01/44301, EP 1134231 and

WO 02/48193 of Unilever; WO 97/49805, WO 01/21817, WO 03/035694, WO 03/054016 and WO 03/055527 of the Vlaams Instituut voor Biotechnologie (VIB); WO 03/050531 of Algonomics N.V. and Ablynx N.V.; WO 01/90190 by the National Research Council of Canada; WO 03/025020 (= EP 1433793) by the Institute of Antibodies; as well as WO 04/041867, WO 04/041862, WO 04/041865, WO 04/041863, WO 04/062551, WO 05/044858, WO 06/40153, WO 06/079372, WO 06/122786, WO 06/122787 and WO 06/122825, by Ablynx N.V. and the further published patent applications by Ablynx N.V. Reference is also made to the further prior art mentioned in these applications, and in particular to the list of references mentioned on pages 41-43 of the International application WO 06/040153, which list and references are incorporated herein by reference. As described in these references, Nanobodies (in particular VHHs and partially humanized Nanobodies) can in particular be characterized by the presence of one or more “Hallmark residues” in one or more of the framework sequences. A further description of the Nanobodies, including humanization and/or camelization of Nanobodies, as well as other modifications, parts or fragments, derivatives or “Nanobody fusions”, multivalent constructs (including some non-limiting examples of linker sequences) and different modifications to increase the half-life of the Nanobodies and their preparations can be found e.g. in WO 08/101985 and WO 08/142164.

Thus, in the meaning of the present invention, the term “immunoglobulin single variable domain” or “single variable domain” comprises polypeptides which are derived from a non-human source, preferably a camelid, preferably a camelid heavy chain antibody. They may be humanized, as previously described. Moreover, the term comprises polypeptides derived from non-camelid sources, e.g. mouse or human, which have been “camelized”, as e.g. described in Davies and Riechmann 1994 (FEBS 339: 285), 1995 (Biotechnol. 13: 475) and 1996 (Prot. Eng. 9: 531) and Riechmann and Muyldermans 1999 (J. Immunol. Methods 231: 25).

The term “immunoglobulin single variable domain” encompasses immunoglobulin sequences of different origin, comprising mouse, rat, rabbit, donkey, human and camelid immunoglobulin sequences. It also includes fully human, humanized or chimeric immunoglobulin sequences. For example, it comprises camelid immunoglobulin sequences and humanized camelid immunoglobulin sequences, or camelized immunoglobulin single variable domains, e.g. camelized dAbs as described by Ward et al (see for example WO 94/04678 and Davies and Riechmann 1994, 1995 and 1996).

The invention may be used for the expression or production of any of the immunoglobulin single variable domains described herein. The invention may also be used for the expression or production of polypeptides of the invention (i.e. polypeptides that comprise or essentially consist of such an immunoglobulin single variable domain). In a particular aspect, the invention may be used for the expression or production of immunoglobulin single variable domains that comprise two disulfide bridges. The invention may also be used for the expression or production of polypeptides of the invention (i.e. polypeptides that comprise or essentially consist of such an immunoglobulin single variable domain) with two disulfide bridges. The invention may also be used for the expression or production of polypeptides of the invention (i.e. polypeptides that comprise or essentially consist of such an immunoglobulin single variable domain with two disulfide bridges).

It is known that all VHH's contain at least one disulfide bridge, between the cysteine residue at position 22 and the cysteine residue at position 92 (numbering according to Kabat, see the patent applications of Ablynx N.V. and Muyldermans and Lauwereys 1999, J. Mol. Recognit. 12: 131). Although most VHH's contain only this single disulfide bridge, it is also known that some VHH's contain a total of two (or in exceptional cases three) disulfide bridges. For example, a class of VHH's (and Nanobodies) referred to as "VHH-1 type", "VHH-1 class", or the like (as further defined herein) commonly has a second disulfide bridge between the cysteine residue in CDR2 at position 50 and a cysteine residue present in CDR3 (or in exceptional cases in CDR1 or CDR2). Also, some VHH's derived from camels or dromedaries often have a disulfide bridges between a cysteine residue present in CDR1 (or at position 45 in FR2) and a cysteine residue present in CDR3 (Vu et al. 1997, Mol. Immunol. 34: 1121; Muyldermans and Lauwereys 1999). Some VHH's derived from llamas sometimes have a disulfide bridge between cysteine residues present in CDR1 (such as at position 33) and a cysteine residue present in CDR3 (Vu et al., 1997).

In one specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in one of the framework regions and a cysteine residue in one of the CDR regions.

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine

residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in one of the framework regions and a cysteine residue in CDR3.

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in framework two (FR2) and a cysteine residue in CDR3.

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue at position 45 in framework two (FR2) and a cysteine residue in CDR3 (as in some VHH's derived from camels and dromedaries).

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in one CDR and a cysteine residue in another CDR.

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in CDR3 and a cysteine residue in another CDR (in particular in CDR1, as in some VHH's derived from camels, dromedaries and llamas).

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in CDR1 and a cysteine residue in another CDR.

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in CDR1 and a cysteine residue in CDR1.

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a

cysteine residue in CDR1 and a cysteine residue in CDR3 (as in some VHH's derived from camels, dromedaries and llamas).

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine at position 33 and a cysteine residue in CDR3 (as in some VHH's derived from camels, dromedaries and llamas).

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in CDR2 and a cysteine residue in another CDR.

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in CDR2 and a cysteine residue in CDR2.

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue in CDR2 and a cysteine residue in CDR3 (as in some VHH's derived from llamas).

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue at position 50 and a cysteine residue in another CDR, such as CDR1, CDR2 or CDR3 (as in VHH's and Nanobodies of the VHH1 type).

In another specific but non-limiting aspect, the immunoglobulin single variable domain comprises a disulfide bridge between the cysteine residue at position 22 and the cysteine residue at position 92, and further comprises a disulfide bridge that is formed between a cysteine residue at position 50 and a cysteine residue in CDR3 (as in VHH's and Nanobodies of the VHH1 type).

In a preferred but non-limiting embodiment, the immunoglobulin single variable domain may be a "VHH1 type immunoglobulin single variable domain". An amino acid such

as e.g. an immunoglobulin single variable domain or polypeptide of the invention is said to be a “VHH1 type immunoglobulin single variable domain” or “VHH type 1 sequence”, if said VHH1 type immunoglobulin single variable domain or VHH type 1 sequence has 85% identity (using the blastp algorithm with standard setting, i.e. blosum62 scoring matrix) to the VHH1 consensus sequence (SEQ ID NO: 46):

QVQLVESGGGLVQPGGSLRLSCAASGFTLDYYAIGWFRQAPGKEREGVSCISSDGSTYYADSVKGRFTIS
RDNAKNTVY LQMNSLKPEDTAVYYCAA) and mandatorily has a cysteine in position 50, i.e. Cys50 (using Kabat numbering). These VHH1 type immunoglobulin single variable domains generally have (or are capable of forming) a disulfide bridge between Cys50 and a cysteine residue in CDR3 (or in exceptional cases CDR1 or CDR2).

An amino acid sequence such as e.g. an immunoglobulin single variable domain or polypeptide of the invention is said to be a “VHH2 type immunoglobulin single variable domain” or “VHH type 2 sequence”, if said VHH2 type immunoglobulin single variable domain or VHH type 2 sequence has 85% identity (using the blastp algorithm with standard setting, i.e. blosum62 scoring matrix) to the VHH2 consensus sequence (SEQ ID NO: 47):

QVQLVESGGGLVQAGGSLRLSCAASGSIFSINAMGWYRQAPGKQRELVAAITSGGSTNYADSVKGRFTIS
RDNAKNTVY LQMNSLKPEDTAVYYCNA.

An amino acid sequence such as e.g. an immunoglobulin single variable domain or polypeptide according to the invention is said to be a “VHH3 type immunoglobulin single variable domain” or “VHH type 3 sequence”, if said VHH3 type immunoglobulin single variable domain or VHH type 3 sequence has 85% identity (using the blastp algorithm with standard setting, i.e. blosum62 scoring matrix) to the VHH3 consensus sequence (SEQ ID NO: 48):

QVQLVESGGGLVQAGGSLRLSCAASGRTFSSYAMGWFRQAPGKEREFVAAISWSGGSTYYADSVKGRF
TISRDNKNTVY LQMNSLKPEDTAVYYCAA.

Although the invention is particularly suited for expression of VHH1 (where the presence of two disulfide bridges is very common), it should be noted that it can also be applied to the expression of VHH2 or VHH3 (which may or may not also comprise two disulfide bridges, although this is less common).

For a general description and for some non-limiting examples of Nanobodies (and of polypeptides comprising the same) that are directed against c-Met and that can be

expressed/produced using the methods described herein, reference is made to WO 2012/042026 and WO 2013/045707.

For a general description and for some non-limiting examples of Nanobodies (and of polypeptides comprising the same) that are directed against TNF and that can be expressed/produced using the methods described herein, reference is made to US provisional application US 62/254,375 of Ablynx NV (see also PCT/EP2016/077595).

The inventors also expect that this teaching is not only particularly applicable to VHH's and Nanobodies with two or more disulfide bridges such as VHH-1's, but also to other immunoglobulin single variable domains that comprise two or more disulfide bridges (such as (single) domain antibodies, dAb's, IgNAR domains from sharks, etc.).

Polypeptide of the invention

The immunoglobulin single variable domains prepared in the method of the invention may form part of a protein or polypeptide (referred to herein as "polypeptide of the invention"), which may comprise or essentially consist of one or more (at least one) immunoglobulin single variable domains and which may optionally further comprise one or more further amino acid sequences (all optionally linked via one or more suitable linkers). The term "immunoglobulin single variable domain" may also encompass such polypeptides of the invention. The one or more immunoglobulin single variable domains may be used as a binding unit in such a protein or polypeptide, which may optionally contain one or more further amino acids that can serve as a binding unit, so as to provide a monovalent, multivalent or multispecific polypeptide of the invention, respectively (for multivalent and multispecific polypeptides containing one or more VHH domains and their preparation, reference is also made to Conrath et al. 2001 (J. Biol. Chem. 276: 7346), as well as to for example WO 96/34103, WO 99/23221 and WO 2010/115998).

The polypeptides of the invention may comprise or essentially consist of one immunoglobulin single variable domain, as outlined above. Such polypeptides are also referred to herein as monovalent polypeptides.

The polypeptides of the invention may also encompass constructs comprising two or more antigen binding units in the form of single variable domains, as outlined above. For example, two (or more) immunoglobulin single variable domains with the same or different antigen specificity can be linked to form e.g. a bivalent, trivalent or multivalent construct. By

combining immunoglobulin single variable domains of two or more specificities, bispecific, trispecific, etc. constructs can be formed. For example, an immunoglobulin single variable domain according to the invention may comprise two or three immunoglobulin single variable domains directed against the same target, or one or two immunoglobulin single variable domains directed against target A, and one immunoglobulin single variable domain against target B. Such constructs and modifications thereof, which the skilled person can readily envisage, are all encompassed by the term polypeptide of the invention as used herein.

Moreover, also prepared in the method of the present invention are fused immunoglobulin sequences, comprising tags or other functional moieties, e.g. toxins, labels, radiochemicals, etc.

In another aspect, the polypeptide of the invention that comprises or essentially consists of one or more immunoglobulin single variable domains (or suitable fragments thereof), may further comprise one or more other groups, residues, moieties or binding units. Such further groups, residues, moieties, binding units or amino acid sequences may or may not provide further functionality to the immunoglobulin single variable domain (and/or to the polypeptide in which it is present) and may or may not modify the properties of the immunoglobulin single variable domain.

For example, such further groups, residues, moieties or binding units may be one or more additional amino acids, such that the compound, construct or polypeptide is a (fusion) protein or (fusion) polypeptide. In a preferred but non-limiting aspect, said one or more other groups, residues, moieties or binding units are immunoglobulins. Even more preferably, said one or more other groups, residues, moieties or binding units are chosen from the group consisting of domain antibodies, amino acids that are suitable for use as a domain antibody, single domain antibodies, amino acids that are suitable for use as a single domain antibody, "dAb"'s, amino acids that are suitable for use as a dAb, or Nanobodies.

Alternatively, such groups, residues, moieties or binding units may for example be chemical groups, residues, moieties, which may or may not by themselves be biologically and/or pharmacologically active. For example, and without limitation, such groups may be linked to the one or more immunoglobulin single variable domain so as to provide a "derivative" of the immunoglobulin single variable domain.

In a preferred but non-limiting aspect, said further residues may be effective in preventing or reducing binding of so-called “pre-existing antibodies” to the polypeptides of the invention. For this purpose, the polypeptides and constructs of the invention may contain a C-terminal extension (X)_n (in which n is 1 to 10, preferably 1 to 5, such as 1, 2, 3, 4 or 5 (and preferably 1 or 2, such as 1); and each X is an (preferably naturally occurring) amino acid residue that is independently chosen, and preferably independently chosen from the group consisting of alanine (A), glycine (G), valine (V), leucine (L) or isoleucine (I), for which reference is made to the following co-pending US provisional applications, all entitled “Improved immunoglobulin variable domains”: US 61/994552 filed May 16, 2014; US 61/014,015 filed June 18, 2014; US 62/040,167 filed August 21, 2014; and US 62/047,560, filed September 8, 2014 (all assigned to Ablynx N.V.) as well as the International application WO 2015/173325 which was based on these provisional applications and which was published on November 19, 2015.

Accordingly, in the method of the present invention, the polypeptide may further comprise a C-terminal extension (X)_n, in which n is 1 to 5, such as 1, 2, 3, 4 or 5, and in which X is a naturally occurring amino acid, preferably no cysteine.

In a preferred aspect, the polypeptide of the invention, expressed in the method of the invention, comprises or essentially consists of SEQ ID NO: 55. In a particular aspect, such polypeptide consists of SEQ ID NO: 55.

In the polypeptides described above, the one or more immunoglobulin single variable domains and the one or more groups, residues, moieties or binding units may be linked directly to each other and/or via one or more suitable linkers or spacers. For example, when the one or more groups, residues, moieties or binding units are amino acids, the linkers may also be an amino acid, so that the resulting polypeptide is a fusion protein or fusion polypeptide.

In one specific aspect of the invention, a polypeptide of the invention is prepared that has an increased half-life, compared to the corresponding immunoglobulin single variable domain. Polypeptides of the invention that comprise such half-life extending moieties for example include, without limitation, polypeptides in which the immunoglobulin single variable domain is suitably linked to one or more serum proteins or fragments thereof (such as (human) serum albumin or suitable fragments thereof) or to one or more binding units

that can bind to serum proteins (such as, for example, domain antibodies, amino acids that are suitable for use as a domain antibody, single domain antibodies, amino acids that are suitable for use as a single domain antibody, "dAb"'s, amino acids that are suitable for use as a dAb, or Nanobodies that can bind to serum proteins such as serum albumin (such as human serum albumin), serum immunoglobulins (such as IgG), or transferrin); polypeptides in which the immunoglobulin single variable domain is linked to an Fc portion (such as a human Fc) or a suitable part or fragment thereof; or polypeptides in which the one or more immunoglobulin single variable domain(s) are suitably linked to one or more small proteins or peptides that can bind to serum proteins (such as, without limitation, the proteins and peptides described in WO 91/01743, WO 01/45746 or WO 02/076489).

Generally, the polypeptides of the invention with increased half-life preferably have a half-life that is at least 1.5 times, preferably at least 2 times, such as at least 5 times, for example at least 10 times or more than 20 times, greater than the half-life of the corresponding immunoglobulin single variable domain or polypeptide of the invention per se.

In a preferred, but non-limiting aspect, such polypeptides of the invention have a serum half-life that is increased with more than 1 hour, preferably more than 2 hours, more preferably more than 6 hours, such as more than 12 hours, or even more than 24, 48 or 72 hours, compared to the corresponding immunoglobulin single variable domain or polypeptide of the invention per se.

In another preferred, but non-limiting aspect, such polypeptides of the invention exhibit a serum half-life in human of at least about 12 hours, preferably at least 24 hours, more preferably at least 48 hours, even more preferably at least 72 hours or more. For example, polypeptides of the invention may have a half-life of at least 5 days (such as about 5 to 10 days), preferably at least 9 days (such as about 9 to 14 days), more preferably at least about 10 days (such as about 10 to 15 days), or at least about 11 days (such as about 11 to 16 days), more preferably at least about 12 days (such as about 12 to 18 days or more), or more than 14 days (such as about 14 to 19 days).

The method of the present invention is especially suited for immunoglobulin single variable domains and/or polypeptides of the invention that are not easily expressible in *Pichia*, such as *P. pastoris*, or in very low yields, when expressed under expression conditions applicable for use with these hosts; and for which thus, upon expression of these

immunoglobulin single variable domains or polypeptides of the invention in *Pichia*, such as *P. pastoris*, not sufficient amounts can be obtained. Accordingly, the method of the invention is especially suited for immunoglobulin single variable domains and/or polypeptides of the invention for which, upon expression in a *Pichia* host (such as *P. pastoris*) under standard *Pichia* expression conditions (as defined herein), a low yield is obtained. A “low yield” as used in the present invention means that the yield of the immunoglobulin single variable domain and/or polypeptide of the invention obtained is 0.5 g/L or lower, such as 0.4 g/L or lower, 0.3 g/L or lower, 0.2 g/L or lower, 0.1 g/L or lower, 0.05 g/L or lower, 0.01 g/L or lower, or even less [expressed as g (gram) immunoglobulin single variable domain or polypeptide of the invention per L (litre) medium].

The method of the invention is also especially suited for immunoglobulin single variable domains and polypeptides of the invention for which, upon expression in a *Pichia* host (such as *P. pastoris*) under standard *Pichia* expression conditions (as defined herein), the yield obtained shows an inverse correlation (as defined herein) with the copy number of the nucleic acid encoding said immunoglobulin single variable domain and/or polypeptide. A yield showing “an inverse correlation” means that the yield of said immunoglobulin single variable domain and/or polypeptide obtained is higher when said immunoglobulin single variable domain and/or polypeptide is expressed (under standard *Pichia* expression conditions as defined herein) in a *Pichia* host that has one copy of the nucleic acid encoding said immunoglobulin single variable domain and/or polypeptide than the yield of said immunoglobulin single variable domain and/or polypeptide obtained when said immunoglobulin single variable domain and/or polypeptide is expressed (under standard *Pichia* expression conditions as defined herein) in a *Pichia* host that has more than one copy of the nucleic acid encoding said immunoglobulin single variable domain and/or polypeptide.

Preferred polypeptides for use in the method of the invention include SEQ ID NO's: 49 to 54. This sequence also forms a separate aspect of the invention. Accordingly, the present invention also relates to a polypeptide with SEQ ID NO: 49, 50, 51, 52, 53, 54, 55 or 56.

Other specifically preferred polypeptides for use in the method of the invention comprise or essentially consist of an immunoglobulin single variable domain that essentially consists of 4 framework regions (FR1 to FR4, respectively) and 3 complementarity determining regions (CDR1 to CDR3, respectively), in which CDR1 is SEQ ID NO: 58, CDR2 is

SEQ ID NO: 60 and CDR3 is SEQ ID NO: 62, such as e.g. immunoglobulin single variable domains with the amino acid sequence of SEQ ID NO's: 55 or 56.

Auxiliary protein

The term “auxiliary protein” as used herein refers to proteins that assist other molecular structures and/or proteins in performing their biological function, but do not themselves occur in the structure of these other molecular structures and/or proteins when these other molecular structures and/or proteins are performing their normal biological functions. Auxiliary proteins may e.g. modify the biophysical, pharmacological and/or expression properties of the other molecular structure and/or protein. Without being limiting, auxiliary proteins may stabilize the other molecular structure and/or protein (e.g. through complex formation), they may modulate the activity of the other molecular structure and/or protein, they may increase the (surface) expression of the other molecular structure and/or protein, and/or they could assist in folding and/or assembly.

The auxiliary protein of which the expression is enhanced in the method of the invention, is a functional HAC1 protein. The HAC1 protein may originate from any species, but is preferably from yeast origin, most preferably a yeast from the Saccharomycetes, such as a yeast from the genus *Saccharomyces*, *Komagataella* or *Pichia* (*Hansenula*), such as *Saccharomyces cerevisiae* or *Pichia pastoris*. In a specific aspect, the functional HAC1 protein is “HAC1spliced” or “HAC1spliced protein” (both terms are used interchangeably herein). HAC1spliced is the HAC1 protein obtained following the splicing event on the *HAC1* mRNA (removal of the intron) as described in Guerfal et al. 2010 (Microbial Cell Factories 9: 49). In a more specific aspect, the HAC1spliced protein is from *Pichia* origin. In a preferred aspect, the HAC1spliced protein has the sequence as described in Guerfal et al. 2010 (Microbial Cell Factories 9: 49; SEQ ID NO: 14).

In the method of the present invention, in addition to HAC1 protein, optionally expression of one or more additional auxiliary proteins selected from protein disulfide isomerase (PDI1; EC 5.3.4.1), Kar2p and Conserved ribosomal protein P0 (RPP0) is enhanced. These auxiliary proteins may originate from any species as long as the enhancement of their expression in a *Pichia* host provides for increased yield of the immunoglobulin single variable and/or polypeptide of the invention. In a preferred aspect, the auxiliary protein originates from a fungus, such as a yeast; preferably a yeast from the Saccharomycetes, such as a yeast

from the genus *Saccharomyces*, *Komagataella* or *Pichia* (*Hansenula*), such as *Saccharomyces cerevisiae* or *Pichia pastoris*.

In a preferred aspect, the auxiliary protein is selected from the following:

- *P. pastoris* HAC1spliced

MPVDSSHKTASPLPPRKRAKTEEEKEQRRVERILRNRRAAHASREKKRRHVEFLENHVVDLESALQES
AKATNKLKEIQDIIVSRLEALGGTVSDLDLTVPEVDFPKSSDLEPMSDLSTSSKSEKASTSTRRLTEDLDE
DDVAEYDDEEEDEELPRKMKVLNDKNKSTSİKQEKLNELPSPLSSDFSDVDEEKSTLTHLKLQQQQQQ
PVDNYVSTPLSLPEDSVDFINPGNLKIESDENFLLSSNTLQIKHENDTDYITTAPSGSINDFFNSYDISESN
RLHHPAAPFTANAFDLNDFVFFQE (SEQ ID NO: 14);

and optionally one or more of:

- *P. pastoris* protein disulfide isomerase (PDI1):

MQFNWNIKTVASILSALTQAQSDQEAIAPEDSHVVKLTEATFESFITSNPHVLAEFFAPWCGHCKKL
PELVSAAEILKDNEQVKIAQIDCTEEKELCQGYEIKGYPTLKVFHGEVEVPSDYQGQRQSQSIVSYMLK
QSLPPVSEINATKDLDDTIAEAKEPVIVQVLPEDASNLESNTTFYGVAGTLREKFTFVSTKSTDYAKKYTS
DSTPAYLLVRPGEEPSVYSGEELDETHLVHWIDIESKPLFGDIDGSTFKSYAEANIPLAYFYENEEQRAA
AADIIKPFKEQRGKINFVGLDAVKFGKHAKNLNMDEEKLPLFVIHDLVSNKKFGVPQDQELTNKDVT
ELIEKFIAGEAEPIVKSEPIPEIQEEKVFKLVGKAHDEVVFDESKDVLVKYYAPWCGHCKRMAPAYEELA
TLYANDEDASSKVVIKLDHTLNDVDNVDIQQGYPTLILYPAGDKSNPQLYDGSRDLESIAEFVKERGTH
KVDALALRPVEEEKEAEEEAASEADAHDEL (SEQ ID NO: 5);

- *P. pastoris* Kar2p:

MLSLKPSWLTLAALMYAMLLVVVPFAKPVRADDVESYGTVIGIDLGTYSVCVGMKSGRVEILANDQ
GNRITPSYVSFTEDERLVGDAAKNLAASNPKNTIFDIKRLIGMKYDAPEVQRDLKRLPYTVKSKNGQP
VSVEYKGEEKSFTPEEISAMVLGKMKLIAEDYLGKKVTHAVVTVPAYFNDAQRQATKDAGLIAGLTVL
RIVNEPTAAALAYGLDKTGEERQIIVYDLGGGTDFVSLLSIEGGAFEVLATAGDTHLGGEDFDYRVVRH
FVKIFKKKHNDISNNDKALGKLKREVEKAKRTLSSQMTTRIEIDSFVDGIDFSEQLSRAKFEEINIELFKKT
LKPVEQVLKDAGVKKSEIDDIVLVGGSTRIPKVQQLLDYFDGKKASKGINPDEAVAYGAAGVQAGVLS
GEEGVDDIVLLDVNPLTLGIETTGGVMTTLINRNTAIPTKKSQIFSTAADNQPTVLIQVYEGERALAKDN
NLLGKFELTGIPPAPRGTPQVEVTFVLDANGILKVSATDKGTGKSESITINNDRGRLSKEEVDRMVVEEA
EKYAAEDAALREKIEARNALENYAHSLRNQVTDDSETGLGSKLDEDDKETLTDAIKDTLEFLEDNFDTA
TKEELDEQREKLSKIAYPITSKLYGAPEGGTPPGGQGFDDDDGDFDYDYDHDDEL (SEQ ID NO: 4);

and

- *P. pastoris* 60S acidic ribosomal protein P0 (RPP0):

MGGINEKKAIEYFNKLRELLESYKSIFIVGVDNVSSQQMHEVRQTLRGKAVILMGKNTMVRKALRDFV
 EELPVFEKLLPFVRGNIGFVFTNEDLKTIRDVIIENRVAAPARPGAIAPLDVFIPAGNTGMEPGKTSFFQ
 ALGVPTKISRGTIEITSDVKVVEKDSRVGPSEAQLLNMLNISPFITYGLTVVQVFDDGQVFPANILDITDD
 ELLSHFTSAISTIAQISLAAGYPTLPSVGHSVNVNHYKNVLAVSIATDYSFEGSEAIKDRLANPEAYAAAAP
 AAGEASAGAEETAAAAEEDEESEDMDMGFGLFD (SEQ ID NO: 6).

Method of the invention

The invention relates to a method for expressing and/or producing immunoglobulin single variable domains and/or polypeptides comprising the same in a *Pichia* host. In the method of the invention the expression of HAC1spliced protein in the *Pichia* host is enhanced. The method of the invention comprises the following steps:

- a) expressing, in a *Pichia* host, a nucleic acid encoding an immunoglobulin single variable domain and/or a polypeptide of the invention; and
- b) enhancing, in said *Pichia* host, the expression of a nucleic acid(s) encoding HAC1spliced protein;

optionally followed by:

- c) isolating and/or purifying the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained.

As such, the method of the present invention comprises the steps of:

- a) cultivating a *Pichia* host under conditions that are such that said *Pichia* host will multiply;
- b) maintaining the *Pichia* host under conditions that are such that said *Pichia* host expresses and/or produces the immunoglobulin single variable domain and/or polypeptide of the invention; and
- c) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein;

optionally followed by:

- d) isolating and/or purifying from the medium the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained.

To produce/obtain expression of the immunoglobulin single variable domains and/or polypeptide of the invention, the transformed *Pichia* host may generally be kept, maintained and/or cultured under conditions that are such that the (desired) immunoglobulin single

variable domain and/or polypeptide of the invention is expressed/produced. Suitable conditions will be clear to the skilled person and will usually depend upon the *Pichia* host strain used, as well as on the regulatory elements that control the expression of the (relevant) nucleotide sequence of the invention.

Generally, suitable conditions may include the use of a suitable medium, the presence of a suitable source of food and/or suitable nutrients, the use of a suitable temperature, and optionally the presence of a suitable inducing factor or compound (e.g. when the nucleotide sequence(s) of the invention are under the control of an inducible promoter); all of which may be selected by the skilled person. Again, under such conditions, the immunoglobulin single variable domain and/or polypeptide of the invention may be expressed in a constitutive manner, in a transient manner, or only when suitably induced.

Culturing conditions for the recombinant production of heterologous proteins in *Pichia* are e.g. described by Higgins and Cregg 1998 (Eds, Methods in Molecular Biology, *Pichia* protocols, Volume 103, 2nd Ed., Humana Press) and by InvitrogenTM (InvitrogenTM *Pichia* Expression Kit; For Expression of Recombinant Proteins in *Pichia pastoris*; Catalog no. K1710-01). Production of immunoglobulin single variable domains in *P. pastoris* has been extensively described in WO 94/25591, WO 2010/125187, WO 2012/056000 and WO 2012/152823. The contents of these references are explicitly referred to in the connection with general culturing techniques and methods, including suitable media and conditions. The contents of these documents are incorporated by reference. The present invention also relates to specific conditions described in the art, for example the general culturing methods described in WO 94/25591, Gasser et al. 2006 (Biotechnol. Bioeng. 94: 535); Gasser et al. 2007 (Appl. Environ. Microbiol. 73: 6499), or Damasceno et al. 2007 (Microbiol. Biotechnol. 74: 381).

Pichia, in particular *P. pastoris*, is typically cultured at 30 °C in fed-batch fermentations using glycerol as carbon source. Such a medium generally comprises a buffering agent, glycerol, trace elements and ammonium hydroxide. Examples of buffering agents include (without being limiting) H₃PO₄, CaSO₄.2H₂O, K₂SO₄, MgSO₄.7H₂O, and KOH. A common growth medium (e.g. basal salt medium) consists of (per L) 26.7 mL 85% H₃PO₄, 0.93 g CaSO₄.2H₂O, 18.2 g K₂SO₄, 14.9 g MgSO₄.7H₂O, 4.13 g KOH, 40 g glycerol, 2 mL trace elements [composed of 6 g/L cupric sulfate.5H₂O; 0.8 g/L potassium iodide; 3 g/L manganese sulfate.H₂O; 0.2 g/L sodium molybdate.2H₂O; 0.2 g/L boric acid; 0.5 g/L copper sulfate; 20

g/L zinc chloride; 65 g/L ferrous sulfate.7H₂O; 0.2 g/L biotin and 5 mL concentrated sulfuric acid]. Ammonium hydroxide (NH₄OH) is used for pH control (e.g. pH 5) and as a nitrogen source. During the fed-batch phase, glycerol (e.g. 50% v/v) is fed at a feed rate of e.g. 15 mL/L/h for several hours. The AOX1 promoter is used to drive expression of the gene of interest encoding the desired immunoglobulin single variable domain and/or polypeptide of the invention. Expression of the immunoglobulin single variable domain and/or polypeptide of the invention is carried out at 30°C with a methanol feed rate of 4-10 mL/L/h (such as e.g. 4 mL/L/h). These conditions are also referred to herein as “standard *Pichia* expression conditions”.

Expression of an auxiliary protein can be enhanced in the *Pichia* host by commonly known means, including e.g. the use of suitable control sequences, e.g. a strong promoter, and/or increasing the gene dose, e.g. by increasing the copy number of the respective gene. The copy number can be increased e.g. by introducing genetic constructs (plasmids or vectors) suitable for expression of the auxiliary protein. The additional presence of a plasmid or vector will increase the overall copy number. Moreover, genetic constructs that can multiply independently of the *Pichia* host genome and are present in multiple copies in the *Pichia* host can be used. For example, multi copy plasmids or vectors may be present in copy numbers between 5 and 50 in the *Pichia* host cell.

In addition to plasmid or vector transformation of the *Pichia* host, chromosomal transformation of the *Pichia* host is also encompassed by the present invention. A strong (inducible) promoter (instead of the native promoter of a native auxiliary protein) may be introduced on the chromosome of the host; or another copy of an auxiliary protein gene sequence under control of another (strong) promoter may be introduced into the chromosome. The skilled person will know a multitude of possibilities of enhancing the expression of the auxiliary protein, all of which are encompassed by the present invention.

The auxiliary protein(s) and the immunoglobulin single variable domain and/or polypeptide of the invention can be expressed from the same or different nucleic acids. Co-expression of the two or more proteins can be accomplished by expression of the two or more proteins on the same genetic construct (plasmid or vector or integrated into the chromosome of the host); or by expression of the two or more proteins on different genetic constructs (plasmids or vectors or integrated into the chromosome of the host).

When expressed on the same genetic construct, the nucleic acids encoding said two or more proteins are preferably located next to each other. The transcription of the nucleic acid encoding said two or more proteins may be controlled by one promoter (located in front of both genes); or each nucleic acid encoding one of the two or more proteins may be controlled by a separate promoter, which may be the same or different promoters.

When expressed from different genetic constructs, the transcription of the nucleic acid encoding the polypeptide of the invention and the transcription of the nucleic acid encoding one or more auxiliary protein(s) may be controlled by two separate promoters which may be the same or different.

The promoter can be a constitutive promoter or an inducible promoter. In a preferred aspect, the promoter is an inducible promoter.

The number of auxiliary proteins of which expression is enhanced in the method of the present invention may be one (HAC1spliced protein), or may be more than one such as e.g. two, three, four, five or more. In a preferred aspect, the number of auxiliary proteins of which expression is enhanced in the method of the present invention is one (HAC1spliced protein).

In another preferred aspect, the number of auxiliary proteins of which expression is enhanced in the method of the present invention is two, three, four or even more. In this preferred aspect the expression of HAC1spliced protein is enhanced and additionally the expression of one or more additional auxiliary proteins is enhanced. The additional auxiliary protein can be any auxiliary protein available and/or known in the art. Preferably the additional auxiliary protein is selected from any of PDI1, Kar2p and RPP0.

Accordingly, also encompassed in the method of the present invention is the expression and/or production of immunoglobulin single variable domains and/or polypeptides comprising the same in a *Pichia* host wherein the expression of HAC1spliced protein and one or more additional auxiliary proteins selected from PDI1, Kar2p and RPP0 is enhanced. Accordingly, the present invention also relates to a method comprising the following steps:

- a) expressing, in a *Pichia* host, a nucleotide sequence encoding an immunoglobulin single variable domain and/or a polypeptide of the invention;

- b) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein and the expression of one, two, three (or more) nucleic acids encoding an auxiliary protein selected from PDI1, Kar2p and RPP0;

optionally followed by:

- c) isolating and/or purifying the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained.

In a specific aspect, the method of the present invention comprises the steps of:

- a) cultivating a *Pichia* host under conditions that are such that said *Pichia* host will multiply;
- b) maintaining the *Pichia* host under conditions that are such that said *Pichia* host expresses and/or produces the immunoglobulin single variable domain and/or polypeptide of the invention; and
- c) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein and the expression of one, two, three (or more) nucleic acids encoding an auxiliary protein selected from PDI1, Kar2p and RPP0;

optionally followed by:

- d) isolating and/or purifying from the medium the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained.

Accordingly, in this particular aspect of the method of the present invention, expression of following combinations of auxiliary protein(s) can be enhanced:

- PDI1 and HAC1spliced;
- Kar2p and HAC1spliced;
- RPP0 and HAC1spliced;
- PDI1, Kar2p and HAC1spliced;
- PDI1, RPP0 and HAC1spliced;
- Kar2p, RPP0 and HAC1spliced; and
- PDI1, Kar2p, RPP0 and HAC1spliced.

When the number of auxiliary proteins is two or more, the auxiliary proteins can be expressed by the action of the same genetic construct, such as expression of the two or more auxiliary proteins on one plasmid or vector; or by expression of the two or more auxiliary proteins on different plasmids or vectors. In addition to plasmid or vector transformation of the *Pichia* host, chromosomal transformation of the *Pichia* host also is

encompassed by the present invention. A strong (inducible) promoter (instead of the native promoter of a native auxiliary protein) may be introduced on the chromosome of the host; another copy of an auxiliary protein gene sequence under control of another (strong) promoter may be introduced into the chromosome. When expressed on the same genetic construct, the two or more auxiliary proteins may be controlled by the same promoter or by a different promoter. When expressed from different genetic constructs, the two or more auxiliary proteins may be controlled by two separate promoters which may be the same or different. The promoter may be a constitutive or an inducible promoter.

By use of the above methods, the present inventors were able to increase the unexpectedly low yield sometimes observed with immunoglobulin single variable domains and/or polypeptides comprising the same. Low yields were particularly observed for immunoglobulin single variable domains (and/or polypeptides comprising the same) that comprise two disulfide bridges, immunoglobulin single variable domains (and/or polypeptides comprising the same) that are VHH1 type immunoglobulin single variable domains, and/or immunoglobulin single variable domains (and/or polypeptides comprising the same) for which, upon expression in a *Pichia* host under standard *Pichia* expression conditions, the yield obtained shows an inverse correlation with the copy number of the nucleic acid encoding said immunoglobulin single variable domain (and/or polypeptide comprising the same).

Accordingly, the present invention also provides a method for increasing the (expression and/or production) yield of immunoglobulin single variable domains and/or polypeptides comprising the same, comprising the following steps:

- a) expressing, in a *Pichia* host, a nucleic acid encoding an immunoglobulin single variable domain and/or a polypeptide of the invention; and
- b) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein;

optionally followed by:

- c) isolating and/or purifying the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained.

As such, the method for increasing the (expression and/or production) yield of immunoglobulin single variable domains and/or polypeptides comprises the steps of:

- a) cultivating a *Pichia* host under conditions that are such that said *Pichia* host will multiply;
- b) maintaining the *Pichia* host under conditions that are such that said *Pichia* host expresses and/or produces the immunoglobulin single variable domain and/or polypeptide of the invention; and
- c) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein;

optionally followed by:

- d) isolating and/or purifying from the medium the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained.

In a preferred aspect, the method of the invention provides a yield of immunoglobulin single variable domain and/or polypeptide of the invention which is 2 or more (preferably 3 or 4 or more, more preferably 5, 7.5, 10, 15, 20, 25, 30, 40, 50 or even more) times the yield of the same immunoglobulin single variable domains and/or polypeptide of the invention obtained in a method wherein the expression of HAC1spliced protein is not enhanced.

Accordingly, in a preferred aspect, the present invention also provides a method for increasing the (expression and/or production) yield of immunoglobulin single variable domains and/or polypeptides comprising the same, comprising the following steps:

- a) expressing, in a *Pichia* host, a nucleic acid encoding an immunoglobulin single variable domain and/or a polypeptide of the invention; and
- b) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein;

optionally followed by:

- c) isolating and/or purifying the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained,

wherein said method provides a yield of immunoglobulin single variable domain and/or polypeptide of the invention which is 2 or more (preferably 3 or 4 or more, more preferably 5, 7.5, 10, 15, 20, 25, 30, 40, 50 or even more) times the yield of the same immunoglobulin single variable domains and/or polypeptide of the invention obtained in a method wherein the expression of HAC1spliced protein is not enhanced.

As such, the method for increasing the (expression and/or production) yield of immunoglobulin single variable domains and/or polypeptides comprises the steps of:

- a) cultivating a *Pichia* host under conditions that are such that said *Pichia* host will multiply;
- b) maintaining the *Pichia* host under conditions that are such that said *Pichia* host expresses and/or produces the immunoglobulin single variable domain and/or polypeptide of the invention; and
- c) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein;

optionally followed by:

- d) isolating and/or purifying from the medium the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained,

wherein said method provides a yield of immunoglobulin single variable domain and/or polypeptide of the invention which is 2 or more (preferably 3 or 4 or more, more preferably 5, 7.5, 10, 15, 20, 25, 30, 40, 50 or even more) times the yield of the same immunoglobulin single variable domains and/or polypeptide of the invention obtained in a method wherein the expression of HAC1spliced protein is not enhanced.

In another preferred aspect, the method of the invention provides a yield of immunoglobulin single variable domain and/or polypeptide of the invention which is 1 g/L or more, more preferably of 1.5 g/L or more, 2 g/L or more, or even 2.5 g/L or more.

Accordingly, in another preferred aspect, the present invention also provides a method for increasing the (expression and/or production) yield of immunoglobulin single variable domains and/or polypeptides comprising the same, comprising the following steps:

- a) expressing, in a *Pichia* host, a nucleic acid encoding an immunoglobulin single variable domain and/or a polypeptide of the invention; and
- b) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein;

optionally followed by:

- c) isolating and/or purifying the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained,

wherein said method provides a yield of immunoglobulin single variable domain and/or polypeptide of the invention which is 1 g/L or more, more preferably of 1.5 g/L or more, 2 g/L or more, or even 2.5 g/L or more.

As such, the method for increasing the (expression and/or production) yield of immunoglobulin single variable domains and/or polypeptides comprises the steps of:

- a) cultivating a *Pichia* host under conditions that are such that said *Pichia* host will multiply;
- b) maintaining the *Pichia* host under conditions that are such that said *Pichia* host expresses and/or produces the immunoglobulin single variable domain and/or polypeptide of the invention; and
- c) enhancing, in said *Pichia* host, the expression of a nucleic acid encoding HAC1spliced protein;

optionally followed by:

- d) isolating and/or purifying from the medium the immunoglobulin single variable domain and/or polypeptide of the invention thus obtained,

wherein said method provides a yield of immunoglobulin single variable domain and/or polypeptide of the invention which is 1 g/L or more, more preferably of 1.5 g/L or more, 2 g/L or more, or even 2.5 g/L or more.

The immunoglobulin single variable domains and/or the polypeptides of the invention are produced extracellular, and are isolated from the medium in which the *Pichia* host cell is cultivated.

Normally, but not necessarily, the immunoglobulin single variable domains and/or polypeptides of the invention will have at least a transport signal which directs the immunoglobulin single variable domains and/or polypeptides of the invention to the periplasm. In the present invention, the *Pichia* host can be removed from the culture medium by routine means. For example, the *Pichia* host can be removed by centrifugation or filtration. The solution obtained by removal of the *Pichia* host from the culture medium is also referred to as culture supernatant, or clarified culture supernatant.

It will also be clear to the skilled person that the immunoglobulin single variable domains and/or polypeptides of the invention may (first) be generated in an immature form (as mentioned above), which may then be subjected to post-translational modification, depending on the *Pichia* host used. Also, the immunoglobulin single variable domains and/or polypeptides of the invention may be glycosylated, again depending on the *Pichia* host cell used.

The immunoglobulin single variable domains and/or the polypeptides of the invention can subsequently be isolated from the *Pichia* host and/or from the medium in which said *Pichia* host was cultivated by standard methods. Standard methods include, but are not limited to chromatographic methods, including size exclusion chromatography, hydrophobic chromatography, ion exchange chromatography, and affinity chromatography. These methods can be performed alone or in combination with other purification methods, e.g. differential precipitation techniques, gel electrophoresis, affinity techniques (e.g. using a specific, cleavable amino acid sequence fused with the immunoglobulin single variable domains and/or polypeptides of the invention) and/or preparative immunological techniques (i.e. using antibodies against the immunoglobulin single variable domains and/or polypeptides of the invention to be isolated). The skilled person can devise suitable combinations of purification methods for immunoglobulin single variable domains on the basis of common general knowledge. For specific examples the art cited herein is referred to.

Immunoglobulin single variable domains and/or polypeptides comprising the same can be purified from culture supernatant by a combination of affinity chromatography on Protein A, ion exchange chromatography and size exclusion chromatography. Reference to any "step of purification", includes, but is not limited to these particular methods. More specifically, immunoglobulin single variable domains and/or polypeptides comprising the same can be purified from culture supernatant using a process wherein the clarified supernatant (obtained by centrifugation) is captured on a Protein A resin; followed by a SOURCE 15S (GE Healthcare) cation exchange chromatography step and a Superdex 75 (GE Healthcare) SEC step.

After removal of the *Pichia* host, the immunoglobulin single variable domain and/or polypeptide of the invention may be present in a wide range of suitable buffers. Examples include, but are not limited to PBS, Tris-HCl, histidine or phosphate buffer. The immunoglobulin single variable domains and/or polypeptides of the invention may also be present in physiological saline.

Generally, for pharmaceutical use, the immunoglobulin single variable domains and/or polypeptides of the invention may be formulated as a pharmaceutical preparation or compositions comprising at least one immunoglobulin single variable domain and/or polypeptide of the invention and at least one pharmaceutically acceptable carrier, diluent or

excipient and/or adjuvant, and optionally one or more further pharmaceutically active polypeptides and/or compounds. By means of non-limiting examples, such a formulation may be in a form suitable for oral administration, for parenteral administration (such as by intravenous, intramuscular or subcutaneous injection or intravenous infusion), for topical administration, for administration by inhalation, by a skin patch, by an implant, by a suppository, etc. Such suitable administration forms - which may be solid, semi-solid or liquid, depending on the manner of administration - as well as methods and carriers for use in the preparation thereof, will be clear to the skilled person.

Nucleic acid and genetic construct of the invention

The present invention also relates to nucleic acids encoding an immunoglobulin single variable domain and/or a polypeptide of the invention and HAC1spliced. These nucleic acids are also referred to herein as “nucleic acid(s) of the invention”.

The nucleic acid of the invention may be in the form of, be present in and/or be part of a genetic construct, as will be clear to the person skilled in the art. Such genetic constructs generally comprise at least one nucleic acid of the invention that is optionally linked to one or more elements of genetic constructs known per se, such as for example one or more suitable regulatory elements (such as suitable promoter(s), enhancer(s), terminator(s), etc.) and the further elements of genetic constructs referred to herein. Such genetic constructs comprising at least one nucleic acid of the invention will also be referred to herein as “genetic construct(s) of the invention”. As such, a genetic construct of the invention at least encodes an immunoglobulin single variable domain and/or polypeptide of the invention and HAC1spliced.

The number of auxiliary proteins encoded by the nucleic acid(s) and genetic construct(s) of the invention may be one (i.e. HAC1spliced protein) or may be more than one, such as two, three, four, five, or more. In a preferred aspect, the number of auxiliary proteins encoded by the nucleic acid and genetic construct of the invention is one (i.e. HAC1spliced protein). In another preferred aspect, the number of auxiliary proteins encoded by the nucleic acid(s) and genetic construct(s) of the invention is two, or more.

Accordingly, the present invention also encompasses nucleic acids or genetic constructs encoding an immunoglobulin single variable domain and/or a polypeptide of the invention and two or more auxiliary proteins (including HAC1spliced protein).

As discussed above, the immunoglobulin single variable domain and/or a polypeptide of the invention and the auxiliary protein(s) can be co-expressed from one single nucleic acid and/or genetic construct; or from different (separate) nucleic acids and/or genetic constructs (possibly including expression from the chromosome of the host). All these nucleic acids and/or genetic constructs encoding the immunoglobulin single variable domain and/or a polypeptide of the invention and/or encoding the auxiliary protein(s) (as one construct or as separate constructs) are encompassed within the terms “nucleic acid(s) of the invention” and “genetic construct(s) of the invention”.

A nucleic acid of the invention can be in the form of single or double stranded DNA or RNA, and is preferably in the form of double stranded DNA. For example, the nucleic acid of the invention may be genomic DNA, cDNA or synthetic DNA (such as DNA with a codon usage that has been specifically adapted for expression in the *Pichia* host cell).

According to one embodiment of the invention, the nucleic acid of the invention is in essentially isolated form, as defined herein. The nucleic acid of the invention may also be in the form of, be present in and/or be part of a vector, such as for example a plasmid, cosmid or YAC, which again may be in essentially isolated form.

The nucleic acids of the invention can be prepared or obtained in a manner known per se, based on the information on the immunoglobulin single variable domain and/or polypeptide of the invention to be expressed and the auxiliary protein(s) used for co-expression. Also, as will be clear to the skilled person, to prepare a nucleic acid of the invention, also several nucleotide sequences, such as at least one nucleotide sequence encoding an immunoglobulin single variable domain and/or polypeptide of the invention and at least one nucleotide sequence encoding the auxiliary protein can be linked together in a suitable manner.

Techniques for generating the nucleic acids of the invention will be clear to the skilled person and may for instance include, but are not limited to, automated DNA synthesis, site-directed mutagenesis, combining two or more naturally occurring and/or synthetic sequences (or two or more parts thereof), introduction of mutations that lead to the expression of a truncated expression product, introduction of one or more restriction sites (e.g. to create cassettes and/or regions that may easily be digested and/or ligated using suitable restriction enzymes), and/or the introduction of mutations by means of a PCR reaction using one or more “mismatched” primers. These and other techniques will be clear

to the skilled person, and reference is again made to the standard handbooks, such as Sambrook et al. and Ausubel et al., mentioned herein, as well as the Examples below.

The genetic constructs of the invention may be DNA or RNA, and are preferably double-stranded DNA. The genetic constructs of the invention may also be in a form suitable for transformation of the *Pichia* host, in a form suitable for integration into the genomic DNA of the *Pichia* host cell or in a form suitable for independent replication, maintenance and/or inheritance in the *Pichia* host. For instance, the genetic constructs of the invention may be in the form of a vector, such as for example a plasmid, YAC, a viral vector or transposon. In particular, the vector may be an expression vector, i.e. a vector that can provide for expression in the *Pichia* host.

In a preferred but non-limiting embodiment, a genetic construct of the invention comprises:

- a) at least one nucleic acid of the invention; operably connected to
- b) one or more regulatory elements, such as a promoter and optionally a suitable terminator;

and optionally also

- c) one or more further elements of genetic constructs known per se;

in which the terms “regulatory element”, “promoter”, “terminator” and “operably connected” have their usual meaning in the art (as further described herein); and in which said “further elements” present in the genetic constructs may for example be 3’- or 5’-UTR sequences, leader sequences, selection markers, expression markers/reporter genes, and/or elements that may facilitate or increase (the efficiency of) transformation or integration. These and other suitable elements for such genetic constructs will be clear to the skilled person, and may for instance depend upon the type of construct used, the *Pichia* host strain, the manner in which the nucleotide sequences of the invention of interest are to be expressed (e.g. via constitutive, transient or inducible expression), and/or the transformation technique to be used. For example, regulatory sequences, promoters and terminators known per se for the expression and production of antibodies and antibody fragments (including but not limited to (single) domain antibodies and ScFv fragments) may be used in an essentially analogous manner.

Preferably, in the genetic constructs of the invention, said at least one nucleic acid of the invention and said regulatory elements, and optionally said one or more further

elements, are “operably linked” to each other, by which is generally meant that they are in a functional relationship with each other. For instance, a promoter is considered “operably linked” to a coding sequence if said promoter is able to initiate or otherwise control/regulate the transcription and/or the expression of a coding sequence (in which said coding sequence should be understood as being “under the control of” said promoter). Generally, when two nucleotide sequences are operably linked, they will be in the same orientation and usually also in the same reading frame. They will usually also be essentially contiguous, although this may also not be required. In one aspect of the invention, the nucleotide sequence encoding the immunoglobulin single variable and/or polypeptide of the invention is operably linked to the nucleotide sequence encoding the auxiliary protein(s). They may be under control of the same promoter, or they may each be under control of a separate (same or different) promoter.

Methods of designing, creating or obtaining nucleic acid sequences for expression, of constructing appropriate vectors, inserting nucleic acid sequences into vectors, choosing appropriate *Pichia* host strains, introducing vectors into the *Pichia* host strain, causing or allowing expression of polypeptides or proteins, isolating nucleic acids from *Pichia* host strains or identifying nucleic acid sequences and corresponding protein sequences are standard methods (Sambrook et al. 1989) which are well known to anyone of ordinary skill in the art. The skilled person can also devise suitable genetic constructs for expression of immunoglobulin single variable domains and/or polypeptides of the invention in the *Pichia* host on the basis of common general knowledge. The present invention also refers to genetic constructs described in the art, for example the plasmids, promoters and leader sequences described in WO 94/25591, Cereghino and Cregg 2000 (Curr. Opin. Biotechnol. 10: 422), Gasser et al. 2006 (Biotechnol. Bioeng. 94: 535), Gasser et al. 2007 (Appl. Environ. Microbiol. 73: 6499) or Damasceno et al. 2007 (Microbiol. Biotechnol. 74: 381).

Preferably, the regulatory and further elements of the genetic constructs of the invention are such that they are capable of providing their intended biological function in the *Pichia* host.

For instance, a promoter, enhancer or terminator should be “operable” in the *Pichia* host by which is meant that (for example) said promoter should be capable of initiating or otherwise controlling/regulating the transcription and/or the expression of a nucleotide sequence - e.g. a coding sequence - to which it is operably linked (as defined herein).

Some preferred, but non-limiting examples of suitable promoters, terminator and further elements include those that can be used for the expression in the *Pichia* host; and in particular those mentioned herein and/or those used in the Examples below.

Some particularly preferred promoters include, but are not limited to, promoters known per se for the expression in the *Pichia* host; and in particular those mentioned herein and/or those used in the Examples. The specific sequence of the promoter determines the strength of the promoter (a “strong promoter” leads to a high rate of transcription initiation). When the expression of a nucleic acid encoding an auxiliary protein is said to be controlled by a “strong promoter”, it is meant that the expression of the nucleic acid encoding said auxiliary protein is controlled by a promoter which leads to higher rate of transcription initiation than the native promoter which controls transcription initiation of the native auxiliary protein. In addition to sequences that “promote” transcription, a promoter may include additional sequences known as operators that control the strength of the promoter. For example, a promoter may include a binding site for a protein that attracts or obstructs the RNA binding to the promoter. The presence or absence of this protein will affect the strength of the promoter. Such a promoter is known as a regulated promoter.

The *P. pastoris* alcohol oxidase I (*AOX1*) promoter is one of the strongest, most regulated promoters known. On the contrary, the *P. pastoris* second alcohol oxidase (*AOX2*) is controlled by a much weaker promoter. The *P. pastoris* glyceraldehyde-3-phosphate dehydrogenase (*GAP*) promoter provides a constitutively high level of expression on glucose, glycerol, and methanol media (Waterham et al. 1997, Gene 186: 37). The *P. pastoris* formaldehyde dehydrogenase (*FLD1*) promoter can be induced either by methanol or methylamine and its expression levels are comparable to those obtained with the *AOX1* promoter in methanol (Shen et al. 1998, Gene 216: 93). The peroxin 8 (*PEX8*) promoter gives low expression on glucose and is induced modestly (about 10-fold) when cells are shifted to methanol (Johnson et al. 1999, Genetics 151: 1379).

Strong promoters in *H. polymorpha* include elements derived from the methanol oxidase (*MOX*), formate dehydrogenase (*FMD*), and dihydroxyacetone synthase (*DHAS*) gene (Song et al. 2006, Biotechnol. Lett. 25: 1999). The glyceraldehyde-3-phosphate dehydrogenase (*GAP1*) promoter (Sohn et al. 1999, Appl. Microbiol. Biotechnol. 51: 800) and *PMA1* promoter (Cox et al. 2000, Yeast 16: 1191) in *H. polymorpha* are constitutive

elements. The *PMA1* promoter competes with the *MOX* promoter in terms of high expression levels.

The *P. methanolica* alcohol oxidase (*AUG1*) promoters, P(MOD1) and P(MOD2), are strong and tightly regulated by methanol (P(MOD1) and P(MOD2)) and glycerol (only P(MOD1) (Nakagawa et al. 2006, Yeast 23: 15).

Strong promoters from *C. boidinii* include the alcohol oxidase (*AOD1*) promoter and the dihydroxy acetone synthase (*DAS1*) promoter (Yurimoto et al. 2000, Biochim. Biophys. Acta 1493: 56). Both *DAS1* and formate dehydrogenase (*FMD*) promoters are available in *C. boidinii* (Sakai et al. 1995, Appl. Microbiol. Biotechnol. 42: 860; 1996, Biochim. Biophys. Acta 1308: 81) and *H. polymorpha* (Hollenberg and Gellissen 1997, Curr. Opin. Biotechnol. 8: 554).

A selection marker should be such that it allows - i.e. under appropriate selection conditions – *Pichia* host cells that have been (successfully) transformed with the nucleic acids of the invention to be distinguished from *Pichia* host cells that have not been (successfully) transformed. Some preferred, but non-limiting examples of such markers are genes that provide resistance against antibiotics (such as e.g. Zeocin, blasticidin, geneticin (G418), phleomycin, kanamycin or ampicillin), genes that provide for temperature resistance, or genes that allow the *Pichia* host to be maintained in the absence of certain factors, compounds and/or (food) components in the medium that are essential for survival of the non-transformed cells.

A leader sequence should be such that - in the *Pichia* host - it allows for the desired post-translational modifications and/or for secretion of the expression product from said cell. As such, the leader sequence may be any pro-, pre-, or prepro-sequence operable in the *Pichia* host. However, normally, but not necessarily, the immunoglobulin single variable domains and/or polypeptides of the invention will have at least a transport signal which directs the immunoglobulin single variable domain and/or protein of the invention to the periplasm. For example, leader sequences known per se for the expression and production of antibodies and antibody fragments (including but not limited to single domain antibodies and ScFv fragments) may be used in an essentially analogous manner.

Some preferred, but non-limiting secretory sequences include the *S. cerevisiae* derived α -mating factor signal sequence, the *P. pastoris* derived acid phosphatase (*PHO1*) signal sequence, *P. pastoris* derived phosphatase (*pho1*) leader sequence, the secretion signal of yeast invertase (*Suc*), the Human Serum Albumin signal peptide, the *S. occidentalis* derived

GAM1 signal sequence, and *Carcinus maenas* derived hyperglycemic hormone (CHH) sequences, etc.

The skilled person may also envisage the use of predicted signal peptides derived from genome sequencing experiments. These predicted signal peptide sequences may originate from any species, but are preferably from yeast origin, most preferably from a yeast from the Saccharomycetes, such as a yeast from the genus *Saccharomyces*, *Komagataella* or *Pichia* (*Hanensula*), such as *Saccharomyces cerevisiae* or *Pichia pastoris*. Some preferred, but non-limiting predicted signal peptides derived from *P. pastoris* are described in De Schutter et al. 2009 (Nature Biotech 27(6): 561-566). Further reference is made to WO 2012/152823 for the use of such predicted signal peptides for the production of immunoglobulin single variable domains.

Known or predicted secretory sequences may also be modified to improve the properties of produced immunoglobulin single variable domains and/or polypeptides of the invention. Such modifications may for example improve the purity of immunoglobulin single variable domains and/or polypeptides of the invention by improving processing efficiency. Modification of the α -mating factor signal sequence for improved processing efficiency is described for example in WO 2012/152823.

An expression marker or reporter gene should be such that - in the *Pichia* host - it allows for detection of the expression of (a gene or nucleotide sequence present on) the genetic construct. Such reporter genes may also be expressed as a protein fusion with the immunoglobulin single variable domain and/or polypeptide of the invention. Some preferred, but non-limiting examples include fluorescent proteins such as GFP and luciferase (LUC).

The genetic constructs of the invention may generally be provided by suitably linking the nucleic acids and/or nucleotide sequence(s) of the invention to the one or more further elements described above, for example using the techniques described in the general handbooks such as Sambrook et al. and Ausubel et al., mentioned above.

Often, the genetic constructs of the invention will be obtained by inserting a nucleic acid or nucleotide sequence of the invention in a suitable (expression) vector known per se. Some preferred, but non-limiting examples of suitable expression vectors are those used in the Examples below, as well as those mentioned herein.

Some preferred, but non-limiting vectors for use in the genetic constructs of the invention include vectors for expression in yeast or other fungal cells such as pYES2 (Invitrogen), pUR3515 and pUR3501 (Sierkstra et al. 1991, Curr. Genet. 19: 81) and *Pichia* expression vectors, such as e.g. (without being limiting) the pPICZ vectors, pPIC3.5, pPIC3.5K, pPIC6a, pPIC9, pPIC9K, pHIL-D2, pHIL-S1 for *P. pastoris* expression, pMET, pMETalpha for *P. methanolica* expression provided by Invitrogen. For a non-exhaustive list of *Pichia* expression vectors reference is also made to Daly and Hearn 2004 (J. Mol. Recognition 18: 119), *Pichia* Protocols 2007 (Ed. Cregg, 2nd Ed., Humana Press, NJ) and Gelissen 2000 (Appl. Microbiol. Biotechnol. 54: 741).

Also encompassed in the present invention are methods for the preparation of the nucleic acid and genetic construct of the invention, comprising the step of cloning the auxiliary protein gene(s) and/or the nucleotide sequence encoding the immunoglobulin single variable domain and/or a polypeptide of the invention in a suitable vector.

The nucleic acids of the invention and/or the genetic constructs of the invention may be used to transform a *Pichia* host, i.e. for expression and/or production of the immunoglobulin single variable domain and/or of a polypeptide of the invention. Suitable techniques for transforming a *Pichia* host will be clear to the skilled person. Reference is again made to the handbooks and patent applications mentioned above. For more detail on transformation procedures reference is made to Invitrogen's EasySelect™ *Pichia* Expression manual, to Wu and Letchworth 2004 (BioTechniques 36: 152), to *Pichia* Protocols 2007 (Ed. Cregg, 2nd Ed., Humana Press, NJ) and to Faber 1994 (Curr. Genet. 25: 305). After transformation, a step for detecting and selecting those *Pichia* host cells that have been successfully transformed with the nucleotide sequence/genetic construct of the invention may be performed. This may for instance be a selection step based on a selectable marker present in the genetic construct of the invention or a step involving the detection of the amino acid sequence of the invention, e.g. using specific antibodies.

Accordingly, the invention further relates to the preparation of a *Pichia* host cell comprising the genetic constructs or nucleic acids of the invention. The skilled person can introduce the nucleic acids or genetic constructs of the invention into the *Pichia* host by routine measures, e.g. by transformation. The skilled person can then select suitable *Pichia* host cells comprising the nucleic acids or genetic construct, e.g. by monitoring the expression of the auxiliary protein on the nucleic acid and/or protein level. A strain with a

satisfactory level of expression will be selected. A high expression of auxiliary protein is desirable, however, it should not be so high as to result in competition with expression of the immunoglobulin single variable domain and/or polypeptide of the invention. This can be determined by routine methods.

The transformed *Pichia* host cell (which may be in the form of a stable cell line) forms a further aspect of the present invention.

Pichia host

Accordingly, the present invention also relates to a *Pichia* host comprising such genetic constructs or nucleic acids as described above. The terms “*Pichia* host” and “*Pichia* host cells” are used interchangeably and refer to the *Pichia* (*Hansenula* and *Hyphopichia* are obsolete synonyms) genus of yeasts in the family Saccharomycetaceae with spherical, elliptical or oblong acuminate cells. *Pichia* is a teleomorph, and forms during sexual reproduction hat-shaped, hemispherical or round ascospores. The anamorphs of some *Pichia* species are *Candida* species. The asexual reproduction is by multilateral budding.

The present invention relates to *Pichia* hosts without limitation, provided that they are suitable for the production of an immunoglobulin single variable domain and/or polypeptide comprising the same. For the purpose of the present invention, the term “*Pichia* host” also includes *Hansenula* and *Candida* species.

The *Pichia* host of the present invention will be capable of producing the immunoglobulin single variable domain and/or the polypeptide of the invention and the HAC1spliced protein, and optionally one or more additional auxiliary protein(s) such as e.g. protein disulfide isomerase (PDI1), Kar2p or Conserved ribosomal protein P0 (RPP0). It will typically be genetically modified such that it comprises one or more nucleic acids encoding the immunoglobulin single variable domain and/or polypeptide of the invention and that expression of HAC1spliced protein is enhanced. Non-limiting examples of genetic modifications comprise the transformation e.g. with a plasmid or vector, or the transduction with a viral vector. Some hosts can be genetically modified by fusion techniques. Genetic modifications include the introduction of separate nucleic acid molecules into a host, e.g. plasmids or vectors, as well as direct modifications of the genetic material of the host, e.g. by integration into a chromosome of the host, e.g. by homologous recombination. Oftentimes a combination of both will occur, e.g. a host is transformed with a plasmid,

which, upon homologous recombination will (at least partly) integrate into the host chromosome. The skilled person knows suitable methods of genetic modification of the host to enable the host to produce immunoglobulin single variable domains and/or polypeptides of the invention.

Suitable *Pichia* hosts will be clear to the skilled person, and may for example be a yeast, including but not limited to *Pichia*, *Hansenula*, or *Candida*, such as methylotrophic yeasts including *Pichia pastoris*, *Pichia methanolica*, *Hansenula polymorpha* (*Pichia angusta*) and *Candida boidinii*. For a non-exhaustive list of *Pichia* strains reference is e.g. made to Gelissen 2000 (Appl. Microbiol. Biotechnol. 54: 741). Without being limiting, *P. pastoris* strains are listed in Daly and Hearn 2004 (J. Mol. Recognition, 18: 119) and *Pichia* Protocols 2007 (Ed. Cregg, 2nd Ed., Humana Press, NJ). Examples of *P. pastoris* strains include (without being limiting) X33, GS115, KM71, KM71H, SMD1163, SMD1165, SMD1168, SMD1168H, NRRL-Y 11430, GS200 provided by Invitrogen, described by Cereghino and Cregg 2004 (FEMS Microbiol. Rev. 24: 45), by Macauley-Patrick et al. 2005 (Yeast 22: 249) and/or Damasceno et al. 2007 (Appl. Microbiol. Biotechnol. 74: 381).

Examples of *Hansenula polymorpha* strains include (without being limiting) A16 (Veale et al. 1992, Yeast 8: 361), GF16 (Faber 1994, Proc. Natl. Acad. Sci. USA 91: 12985), CBS4732 (CCY38-22-2; ATCC34438, NRRL-Y-5445), DL-1 (NRRL-Y-7560; ATCC26012), and strain NCYC495 (CBS1976; ATAA14754, NRLL-Y-1798).

Examples of *P. methanolica* strains include (without being limiting) PMAD11 and PMAD16 provided by Invitrogen. Strains of *P. methanolica* (IAM12901 and IAM12481) and *C. boidinii* (IAM12875) are described by Nakagawa et al. 1996 (J. Fermentation Bioeng. 81: 498).

Reference is also made to the general background art cited hereinabove, as well as to for example WO 94/29457, Frenken et al. 1998 (Res. Immunol. 149: 589), van der Linden 2000 (J. Biotechnol. 80: 261), Joosten et al. 2003 (Microb. Cell Fact. 2: 1), and the further references cited herein.

For production on industrial scale, preferred heterologous hosts for the (industrial) production of immunoglobulin single variable domain-containing protein therapeutics include strains of *Pichia pastoris* that are suitable for large scale expression/production/fermentation, and in particular for large scale pharmaceutical expression/production/fermentation. Suitable examples of such strains will be clear to the

skilled person. Such strains and production/expression systems are also made available by companies such as Avecia Biologics (Billingham, North East England, UK), BIOMEVA GmbH (Heidelberg, Germany), PharmedArtis GmbH (Aachen, Germany), Richter-Helm (Hamburg, Germany), and CMC Biologics (Copenhagen, Denmark).

The invention also includes further generations, progeny and/or offspring of the *Pichia* host cell of the invention, which may for instance be obtained by cell division.

Pharmaceutical preparation

The present invention also relates to immunoglobulin single variable domains and/or polypeptides of the invention obtainable by the methods of the invention as described herein.

Accordingly, the present invention also relates to pharmaceutical preparations and other compositions comprising an immunoglobulin single variable domain and/or a polypeptide of the invention obtainable by the methods of the present invention. The present invention also relates to the medical use of the immunoglobulin single variable domain and/or polypeptide of the invention obtainable by the method of the present invention.

The skilled person can readily formulate pharmaceutically suitable formulations on the basis of common general knowledge. Moreover, the references specifically dealing with immunoglobulin single variable domains and/or Nanobodies, which are cited herein, are explicitly referred to. Without limitation, formulations for standard routes of application can be prepared, including formulations for nasal, oral, intravenous, subcutaneous, intramuscular, intraperitoneal, intravaginal, rectal application, topical application or application by inhalation.

Based on the present invention, the skilled person can also readily devise suitable methods of treatment characterized by the use of a therapeutically effective amount of the immunoglobulin single variable domain and/or polypeptide of the invention obtainable by the method of the invention.

The present invention is further illustrated by the following Examples, which in no way should be construed as further limiting. The entire contents of all of the references (including literature references, issued patents, published patent applications, and co-

pending patent applications) cited throughout this application are hereby expressly incorporated by reference, in particular for the teaching that is referenced hereinabove.

EXAMPLES

The experimental section describes the surprising observation of low yields upon expression of monovalent and multivalent immunoglobulin single variable domains in *Pichia pastoris*. We also observed that yield is further reduced when more than 1 copy of the expression cassette is present in the genome of *P. pastoris*. Also described is a method for increasing the expression yield of said immunoglobulin single variable domains by enhancing the expression of HAC1spliced.

Example 1: Identification of auxiliary proteins that increase the expression of Nanobodies in *Pichia pastoris*

1.1 Construction of expression vectors

Nanobody A, previously described in WO 2013/045707 as SEQ ID NO: 7, is a bivalent Nanobody consisting of two sequence optimized variable domains of a heavy-chain llama antibody. The N-terminal subunit in Nanobody A is a VHH1 type immunoglobulin single variable domain and is specific for binding c-Met, while the C-terminal subunit binds to human serum albumin (HSA). The subunits are fused head-to-tail with a 9G/S linker. The sequence of Nanobody A (SEQ ID NO: 49) is depicted in Table A-1. Nanobody A was previously shown to give very low yields (0.2 g/L or lower) upon fermentation in *P. pastoris*.

DNA fragments containing the coding information of Nanobody A were cloned into the multiple cloning site of a *Pichia* expression vector (derivative of pPIC6a, Invitrogen) that contains a blasticidin™ resistance gene marker, such that the Nanobody sequence was downstream of and in frame with the aMF signal peptide sequence. To generate *Pichia* clones with more than 1 copy number of the expression cassette in the genome, a unique BglII site in the *Pichia* expression vector was used to introduce a second expression cassette of Nanobody A.

Coding sequences of the auxiliary proteins depicted in Table A-2 were cloned (using the restriction enzymes BstBI and NotI) into a *Pichia* expression vector (derivative of the pPICZa, Invitrogen) containing the Zeocin™ resistance gene marker. Auxiliary proteins

containing a BstBI site in their coding sequence were cloned using the restriction enzymes AfeI and NotI. The Nanobody and auxiliary protein in the pPIC6a and the pPICZa vectors were both under the control of the AOX1 methanol inducible promoter.

1.2 Transformation of the Nanobody coding sequences and expression and secretion of the Nanobody in *Pichia pastoris*

Transformation and expression studies of wild type *Pichia* X33 were performed by standard techniques and in accordance with the 'User manual for pPicZalphaA, B and C' (version D, 110801, Manual part no. 25-0148; Invitrogen) and Methods in Molecular Biology 2007 (Humana Press Inc.). Firstly, the *P. pastoris* strain was transformed with the appropriate expression vector with single or double expression cassette of Nanobody A. Transformants were grown on selective medium containing blasticidin™. A number of individual colonies were characterized by qPCR to select clones having 1 copy of the expression vector integrated into the genome and clones having more than one copy of the expression cassette integrated into the genome. Expression and secretion into the medium of the Nanobody was verified (Figure 1).

1.3 Transformation of the auxiliary protein coding sequences and expression and secretion of the Nanobody in *Pichia pastoris*

Once a suitable Nanobody expressing colony was identified, its inoculum was propagated and prepared as competent cells. These cells were then transformed with a library of expression vectors containing the 22 auxiliary proteins depicted in Table A-2. Transformants were grown on selective medium containing a different selection marker (Zeocin™) and this way, co-transformants containing both the Nanobody of interest and one or more of the auxiliary proteins of Table A-2 were obtained. Shake-flask expression was performed in 5 mL cultures in BMCM medium and induced by the addition of methanol as has been described in *Pichia* protocols (see e.g. Methods in molecular biology 2007, Humana Press Inc.).

In each setup 1128 clones were screened for improved expression and compared to their corresponding reference clone which only contained 1 copy or more than one copy of the Nanobody expression cassette integrated into the genome but without the expression vector coding for one or more auxiliary proteins. For each setup we found 2 clones with

yields significantly higher than their reference clones. Clones 6H1 and 4C2 had one copy of the Nanobody coding sequence integrated in the genome (copy number = 1) and clones 5A6 and 9C4 had more than one copy of the Nanobody coding sequence integrated in the genome (copy number >1) (Figure 2).

1.4 Identification of auxiliary proteins that have positive effect on the expression yield

Identification of the auxiliary proteins that exert a positive effect on the expression yield of the Nanobody in *P. pastoris* was done by means of genomic DNA PCR using sequence-specific PCR primers. The list of primers used is shown in Table A-3.

The identified auxiliary proteins are shown in Table 1.

Table 1: The auxiliary proteins present in 4 clones (6H1, 4C2, 5A6, 9C4) that showed increased expression levels of Nanobody A were identified with specific PCR primers. The black boxes indicate the presence of a specific auxiliary protein in the corresponding clone. Auxiliary proteins that were present in both clones 4C2 and 6H1 or in both clones 5A6 and 9C4 are indicated with arrows.

		Copy number = 1		Copy number > 1	
		4C2	6H1	5A6	9C4
	Fkpa				
➡	Kar2p				
➡	PDI1				
➡	RPPO				
	BMH2				
	Cct2				
	Gim4				
	Mdj1				
➡	HAC1spliced				
	Gas1				
	Pma1				

SSe1				
Uso1				
Ydj1				

1.5 Determination of the expression yield of the different clones

Expression yields of the Nanobody/auxiliary protein(s) co-transformants were compared to expression yields of controls (Nanobody transformants without enhancement of the expression of one or more auxiliary protein(s)) in expression experiments by quantification of the yields of expressed and secreted Nanobody in the medium. Standard fed batch fermentations conditions were used. Glycerol fed batches were performed and induction was initiated by the addition of methanol. The productions were performed at 2L scale at pH6, 30°C in complex medium with a methanol feed rate of 4 ml/L*h.

Samples were subjected to SDS-PAGE analysis. Relative quantifications of the proteins were done by means of Coomassie stained SDS-PAGE densitometry scan measurements (Table 2).

Table 2: Expression yields of Nanobody A with and without enhanced expression of auxiliary protein(s). Yield was estimated using SDS-PAGE/Coomassie staining and quantification of bandvolume was done using Imagequant software (GE Healthcare).

Fermenter	Clone	Yield determined on gel (g/L)	Improvement over reference clone
R5/130529	Reference clone (Copy number=1)	0.2	-
R6/130529	4C2	1.5	7.5 times
R7/130529	6H1	0.9	4.5 times
R8/130529	Reference clone (Copy number >1)	0.1	-
R9/130529	5A6	2.2	22 times
R10/130529	9C4	2.7	27 times

Clone 4C2 and 6H1 showed a remarkable increase in expression compared to their reference clone (1 copy number of the expression cassette integrated into the genome). This increase in expression was likely the result of the co-expression of PDI1 present in both clones. Similarly clones 5A6 and 9C4 showed a vast increase in yield compared to their reference clone. The auxiliary proteins that were both expressed in clones 5A6 and 9C4 are Kar2p, RPP0 and HAC1spliced. Most likely those auxiliary proteins are involved in improved expression of Nanobody A. Interestingly, the expression level of clone 4C2 is remarkably higher than clone 6H1 which is likely the result of the co-expression of Kar2p and HAC1spliced which are also present in the highest expressing clones 5A6 and 9C4. The clones that had the highest yield all co-expressed HAC1spliced (Tables 1 and 2).

Example 2: Evaluation of Nanobody A yields when the expression of individual auxiliary proteins is enhanced

The individual auxiliary proteins PDI1, Kar2p, RPP0 and HAC1spliced were transformed into the Reference clone with more than 1 copy of the Nanobody expression cassette in the genome as described in Example 1.3. Transformants were grown on selective medium containing ZeocinTM. Co-transformants containing both Nanobody A and a specific auxiliary protein were obtained. Shake-flask expression was performed in 5 mL cultures in BMCM medium and induced by the addition of methanol as has been described in Pichia protocols (see e.g. Methods in molecular biology 2007, Humana Press Inc.). Relative quantifications of the proteins were done by means of Coomassie stained SDS-PAGE densitometry scan measurements (Figure 3). All clones co-expressing one of the auxiliary proteins showed a significant increase in yield of Nanobody A. Again, we observed that the clone containing HAC1spliced showed the largest improvement in yield.

Example 3: Evaluation of Nanobody B yields when expression of the individually auxiliary proteins is enhanced

3.1 Construction of expression vectors

Nanobody B is a trivalent Nanobody consisting of three sequence optimized variable domains of a heavy-chain llama antibody. The subunits in Nanobody B are not of the VHH1 type immunoglobulin single variable domain and do not bind to human serum albumin

(HSA). The subunits are fused head-to-tail with 35 G/S linkers. Nanobody B was previously shown to give very low yields (0.29 g/L or lower) upon fermentation in *P. pastoris*.

DNA fragments containing the coding information of Nanobody B were cloned into the multiple cloning site of a *Pichia* expression vector (derivative of the pPpT4_Alpha_S; Näätäsaari et al. 2012, PLoS One 7: e39720) that contains a Zeocin™ resistance gene marker, such that the Nanobody sequence was downstream of and in frame with the aMF signal peptide sequence. Coding sequences of the auxiliary proteins HAC1spliced, Kar2p, PDI1 and RPP0 were cloned into *Pichia* expression vectors containing the Blasticidin™ resistance gene marker. The Nanobody and auxiliary protein were both under the control of the AOX1 methanol inducible promoter.

3.2 Transformations

The individual auxiliary proteins PDI1, Kar2p, RPP0 and HAC1spliced were transformed into the *Pichia pastoris* strain NRRL Y-11430 (ATCC number 76 273). Transformants were grown on selective medium containing Blasticidin™. Single clones were isolated and subsequently transformed with Nanobody B.

3.3 Determination of the expression yields of the individual clones

Expression analysis was done as described in Example 1.5. Relative quantifications of the proteins were done by means of Coomassie stained SDS-PAGE densitometry scan measurements (Table 3). Only the clone co-expressing HAC1spliced auxiliary protein showed a large increase in yield of Nanobody B, which again shows that enhanced expression of the auxiliary protein HAC1spliced most efficiently improves the yield.

Table 3: Expression yields of Nanobody B with and without enhanced expression of auxiliary protein. Yield was estimated using SDS-PAGE/Coomassie staining and quantification of bandvolume was done using Imagequant software (GE Healthcare).

	Strain	Yield determined on gel (g/L)	Improvement over reference clone
Nanobody B	Reference	0.29	-
	HAC1spliced	3.0	10.3 times
	Kar2p	0.8	2.8 times

	PDI1	0.16	-
	RPP0	0.22	-

Example 4: Evaluation of the yields of the good expressing Nanobody C when expression of the individual auxiliary proteins is enhanced

Nanobody C is a bivalent Nanobody consisting of two sequence optimized variable domains of a heavy-chain llama antibody. The subunits in Nanobody C are not of the VHH1 type immunoglobulin single variable domain. The C-terminal subunit binds to human serum albumin (HSA). The subunits are fused head-to-tail with a 35 G/S linker. The individual auxiliary proteins PDI1, Kar2p, RPP0 and HAC1spliced were cloned as described in Example 3 and transformed into the *Pichia pastoris* strain NRRL Y-11430. Transformants were grown on selective medium containing BlasticidinTM. Single clones were isolated and subsequently transformed with Nanobody C. Expression analysis was done as described in Example 1.5. Relative quantifications of the proteins were done by means of Coomassie stained SDS-PAGE densitometry scan measurements (Table 4).

Only the clone co-expressing HAC1spliced auxiliary protein showed a significant increase in yield of Nanobody C. This again shows that enhanced expression of the auxiliary protein HAC1spliced is most effective to improve Nanobody yield. This illustrates that enhanced expression of the auxiliary protein HAC1spliced can also further increase the yield of good expressing Nanobodies.

Table 4: Expression yields of Nanobody C with and without enhanced expression of auxiliary protein. Yield was estimated using SDS-PAGE/Coomassie staining and quantification of bandvolume was done using Imagequant software (GE Healthcare).

	Strain	Yield determined on gel (g/L)	Improvement over reference clone
Nanobody C	Reference	1.4	-
	HAC1spliced	4.8	3.4 times
	Kar2p	0.9	-
	PDI1	1.7	1.2 times
	RPP0	2	1.4 times

Example 5: Evaluation of the expression yields of different Nanobodies (Nanobody D, Nanobody E, Nanobody F, Nanobody G, and Nanobody H) when expression of the auxiliary protein HAC1spliced is enhanced

Nanobody D is a bivalent Nanobody consisting of two sequence optimized variable domains of a heavy-chain llama antibody. The N-terminal subunit in Nanobody D is a VHH1 type immunoglobulin single variable domain and is specific for binding c-Met, while the C-terminal subunit binds to human serum albumin (HSA). The subunits are fused head-to-tail with a 9G/S linker and contain a C-terminal Flag3-His6 epitope tag. The sequence of Nanobody D (SEQ ID NO: 50) is depicted in Table A-1. Nanobody E is a trivalent Nanobody consisting of three sequence optimized variable domains of a heavy-chain llama antibody. The C-terminal subunit in Nanobody E is a VHH1 type immunoglobulin single variable domain, while the central subunit binds to human serum albumin (HSA). The subunits are fused head-to-tail with G/S linkers and contain a C-terminal Flag3-His6 epitope tag. The sequence of Nanobody E (SEQ ID NO: 51) is depicted in Table A-1. Nanobody F is a trivalent Nanobody consisting of three sequence optimised variable domains of a heavy-chain llama antibody. The C-terminal subunit in Nanobody F is a VHH1 type immunoglobulin single variable domain. The subunits in Nanobody F do not bind to human serum albumin (HSA). The subunits are fused head-to-tail with 35 G/S linkers and contain a C-terminal Flag3-His6 epitope tag. The sequence of Nanobody F (SEQ ID NO: 52) is depicted in Table A-1. Nanobodies G and H are tetravalent Nanobodies consisting of four sequenced optimised variable domains of a heavy-chain llama antibody. The C-terminal subunit in Nanobodies G and H is a VHH1 type immunoglobulin single variable domain, while one of the central subunits binds to human serum albumin (HSA). The subunits are fused head-to-tail with 35G/S linkers. Sequences of Nanobodies G and H (SEQ ID NO: 53 and SEQ ID NO: 54, respectively) are depicted in Table A-1. Nanobody I is a monovalent Nanobody that specifically binds TNF. Nanobody I essentially consists of one sequence optimized variable domain, further comprising one alanine residue as C-terminal extension. Nanobody I is not a VHH1 type immunoglobulin single variable domain. The sequence of Nanobody I (SEQ ID NO: 55) is depicted in Table A-1.

The individual auxiliary protein HAC1spliced was cloned as described in 3.1 and transformed into the *Pichia pastoris* strain NRRL Y-11430. Transformants were grown on selective medium containing BlasticidinTM. Single clones were isolated and subsequently

transformed with Nanobody D, E, F, G or H. Expression analysis was done as described in Example 1.5. Relative quantifications of the proteins were done by means of Coomassie stained SDS-PAGE densitometry scan measurements (Table 5). Enhancing the expression of HAC1spliced auxiliary protein improved the yield of all Nanobodies. This again demonstrates that enhancement of the expression of the auxiliary protein HAC1spliced effectively improves yield of Nanobodies.

Table 5: Expression yields of Nanobody D, E, F, G or H. with and without enhanced expression of HAC1spliced auxiliary protein. Yield was estimated using SDS-PAGE/Coomassie staining and quantification of bandvolume was done using Imagequant software (GE Healthcare).

	Strain	Yield determined on gel (g/L)	Improvement over reference clone
Nanobody D	Reference	0.4	4.3 times
	HAC1spliced	1.7	
Nanobody E	Reference	0.8	2.0 times
	HAC1spliced	1.6	
Nanobody F	Reference	0.9	2.2 times
	HAC1spliced	2.0	
Nanobody G	Reference	0.19	6.3 times
	HAC1spliced	1.2	
Nanobody H	Reference	0.25	6.0 times
	HAC1spliced	1.5	
Nanobody I	Reference	2.5	3.2 times
	HAC1spliced	8.0	

TABLES

Table A-1. Immunoglobulin single variable domain sequences

SEQ ID NO	Reference	Amino acid sequence
46	VHH1 consensus sequence	QVQLVESGGGLVQPGGSLRLSCAASGFTLDYYAIGWFRQAPGKEREGVSCISSSDGSTYYADSVKGRFTISRDNANKNTVYLMNSLKPEDTAVYYCAA
47	VHH2 consensus sequence	QVQLVESGGGLVQAGGSLRLSCAASGSIFSINAMGWYRQAPGKQRELVAAITSGGSTNYADSVKGRFTISRDNANKNTVYLMNSLKPEDTAVYYCNA
48	VHH3 consensus sequence	QVQLVESGGGLVQAGGSLRLSCAASGRTFSSYAMGWFRQAPGKEREFVAAISWSGGSTYYADSVKGRFTISRDNANKNTVYLMNSLKPEDTAVYYCAA
49	Nanobody A	EVQLVESGGGLVQPGGSLRLSCAASGFILDYYAIGWFRQAPGKEREGVLCIDASDDITYYADSVKGRFTISRDNANKNTVYLMNSLKPEDTGVYYCATPIGLSSSCLLEYDYDYWGQGTTLVTVSSGGGGSGGGGSEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLMNSLRPEDTAVYYCTIGGSLSRSSQGTTLVTVSS
50	Nanobody D	EVQLVESGGGLVQPGGSLRLSCAASGFLLNYFEIVWFRQAPGKEREGIIICISNSDDKTTYVDSVKGRFTFSRDVAKNTVYLMNSLKR EDTADYYCATNL YGTCHTTLKADDMAYWGKGTTLVTVSSGGGGSGGGGSEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLMNSLRPEDTAVYYCTIGGSLSRSSQGTTLVTVSSGAADYKDHDGDYKDHDIDYKDDDDKGAAHHHHH
51	Nanobody E	EVQLLESGGGLVQPGGSLRLSCAASGFTLDDYAIWFRQAPGKGREGVSGIDSGDGSAYYADSVKGRFTISSDNSKNTVYLMNSLRPEDTAVYYCARVRTGWGLNAPDYAMDYWGQGTTLVTVSSGGGGSGGGGSGGGGSEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLMNSLRPEDTAVYYCTIGGSLSRSSQGTTLVTVSSGGGGSGGGGSEVQLLES GGGLVQP GGSLRLSCAASGFTLDYLAIGWFRQAPGKGREGVSCVSSSGQYTYYADSVKGRFTISRDNSESTVYLMNSLRPEDTAVYYCATDPECYRVRGYNAEYDYWGQGTTLVTVSS
52	Nanobody F	EVQLVESGGGLVGTGGSLRLSCAASGNIADLGVMGWYRQAPAKKGELVATMPRTGSKWYQDSVKGRFTIHRDNSKSTVDLEMGSLKPEDTAVYYCVASRMFQTILKPNYWQGTTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLVESGGGLVQPGESLRLSCVASGFTFSSTDMSWLRQATGKGP EWLSSINS GGSSTRYAESVKGRFTVSRDNTKNTLYLMDSLQPEDTAKYYCARGWTPTGRAGPGTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLVESGGGLVQAGGSLRLSCAASGFTLDDYDMSWFRQAPGKE REMISCISSDGRPYEDSVKGRFTVTSDNAKNTVYLMNSLKPEDTAVY YCAAGAKIFAVPGSLCSVRNAHWGQGTTLVTVSSGAADYKDHDGDYKDHDIDYKDDDDKGAAHHHHH
53	Nanobody G	DVQLVESGGGLVQPGGSLRLSCAASGLTFSTNPMYWYRQAPGKQRELVASISSRGITNYADSVKGRFTISRDNSKNTVYLMNSLRPEDTAVYYCRLASLSSGTVYWQGTTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLVESGGGLVQP GGSLRLSCAASGSTRSVNPMWFRQAPGKQREWVATISRSGYATYADSVKGRFTISRDNSKNTVYLMNSLRPEDTAVYYCVTGTYWGQGTTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLVESGGGLVQP GGSLRLSCAASGFTFRSFGMSWVRQAPGKGPEWVSSISGSGSDTLYADSVKGRFTISRDNSKNTLYLMNSLRPEDTAVYYCTIGGSLSRSSQGTTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSEVQLVESGGGLVQP GGSLRLSCAASGSTLDYYAIGWFRQAPGKEREGVSCTSNSGSTYYADSVKGRFTASRDNSKNTVYLMNSLRPEDTAVYYCVATIGCATLGGTLDVQRYYYRGQGTTLVTVSSA
54	Nanobody H	DVQLVESGGGLVQP GGSLRLSCAASGLTFSTNPMYWYRQAPGKQRELVASISSRGITNYADSVKGRFTISRDNSKNTVYLMNSLRPEDTAVYYCRLASLSSGTVYWQGTTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGG

		GSEVQLVESGGGLVQPGGSLRLSCAASGRIFSIINRMGWYRQAPGKQRELV AGVTINAITNYADSVKGRFTISRDN SKNTVY LQMNSLRPEDTAVYYCHAW ARSSGSAPYSQNWGQGTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGG GGSGGGGSEVQLVESGGGLVQPGGSLRLSCAASGFTFRSFGMSWVRQAPG KGPEWVSSIISGSGSDTLYADSVKGRFTISRDN SKNTLY LQMNSLRPEDTA VYYCTIGGSLSRSSQGTLVTVSSGGGGSGGGGSGGGGSGGGGSGGGGSGG GGSGGGGSEVQLVESGGGLVQPGGSLRLSCAASGSTLDYYAIGWFRQAPG KEREGVSCTSNSGSTYYADSVKGRFTASRDN SKNTVY LQMNSLRPEDTAV YYCVATIGCATLGGTLDVQRYYYRGQGTLVTVSSA
55	Nanobody I	DVQLVESGGGVVQPGGSLRLSCTASGFTFSTADMGWFRQAPGKGREFVAR ISGIDGTTYDEPVKGRFTISRDN SKNTVY LQMNSLRPEDTALYYCRSPR YADQWSAYDYWGQGTLVTVSSA
56	Nanobody J	DVQLVESGGGVVQPGGSLRLSCTASGFTFSTADMGWFRQAPGKGREFVAR ISGIDGTTYDEPVKGRFTISRDN SKNTVY LQMNSLRPEDTALYYCRSPR YADQWSAYDYWGQGTLVTVSS

Table A-2. Auxiliary proteins that were screened for improving the expression of Nanobody A

Auxiliary protein	SEQ ID NO:	Protein ID/ Reference	Sequence
Fkpa	1	BAI37926	MKSLFKVTLLATTMAVALHAPITFAAEAAKPATTADSKAAAFKN DDQKSAYALGASLGRYMENSLKEQEKLGIKLDKQDLIAGVQDA FADKSKLSDQEIEQTLQAFEARVKSSAQAKMEKDAADNEAKGK EYREKFAKEKGVKTSSTGLVYQVVEAGKGGEAPKDSDTVVVNYK GTLIDGKEFDNSYTRGEPLSFRLDGVIPGWTEGLKNIKKGGKI KLVI PPDELAYGKAGVPGIPPNSTLVFDVELLDVKPAPKADAKP EADAKAADSACK
SKP	2	BAI34178.1	MKKWLLAAGLGLALATSAQAADKIAIVNMGSLFQQVAQKTGVS NTLENEFKGRASELQRMETDLQAKMKKLQSMKAGSDRTKLEKD VMAQRQTFAQKAQAFEQDRARRSNEERGKLVTRIQTAVKSVAN SQDIDLVDANAVAYNSSDVKIDITADV LKQVK
EroO	3	CAY67364.1	MRIVRSVAIAIAIACHCITALANPQIPFDGNYTEIIVPDTEVNIG QIVDINHEIKPKLVELVNTDFFKYKLN LWKPCPFWNGDEGFC KYKDCSVDFITDWSQVPDIWQPDQLGKLGDNTVHKDKGQDENE LSSNDYCALDKDDDEDLVYVNLIDNPERFTGYGGQQSESIWTA VYDENC FQPNEGSQLGQVEDLCLEKQIFYRLVSGLHSSI THL TNEYLN LKNGAYEPNLKQFMIKVG YFTERIQNLHLNYVLV LKS LIK LQEYNVIDNLPLDDSLKAGLSGLISQGAQGINQSSDDYLF NEKVL FQNDQND DLKNEFRDKFRNVTRLMDCVHCERCKLWGKL QTTGYGTALKILFDLKNPNDSINLKRVELVALVNTFHRLSKSV ESIENFEKLYKIQPPTQDRASASSESLGLFDNEDEQNLLNSFS VDQAVISSKEAPEEIKSKPVGKAAYKQNSCPSLGSKSIKEAFH EELHAFIDAIGFILNSYRTL PKLLYTLFLVKSSSELWDIFIGTQ RHRDTTYRVDL
Kar2p	4	AAX77226.1	MLSLKPSWLT LAALMYAMLLVVVPFAKPV RADDVESYGTVIGI DLGTTYSCVGV MKSGRVEILANDQGNRITPSYVSFTEDERLVG DAAKNLAASNPKNTIFDIKRLIGMKYDAPEVQRDLKRLPYTVK SKNGQP VVSVEYKGEEKSFTPEEISAMVLGKMKLIAEDYLGKK VTHAVVTVPAYFNDAQRQATKDAGLIAGLTVLRIVNEPTAAAL AYGLDKTGEERQII VYDLGGGTDFDVSLLSIEGGAFEVLATAGD THLGGEDFDYRVVRHFVKIFKKKHNIDISNNDKALGKLKREVE KAKRTLSSQMTTRIEIDSFVDGIDFSEQLSRAKFEEINIELFK KTLKPVEQVLKDAGVKKSEIDDIVLVGGSTRIPKVQQLEDYF DGKKASKGINPDEAVAYGA AVQAGVLSGEEGVDDIVLLDVNPL TLGIETTGGVMTTLINRNTAIPTKKSQIFSTAADNQPTVLIQV YEGERALAKDNNLLGKFELTGIPAPRGTPQVEVTFVLDANGI LKVSATDKGTGKSESITINNDRGRLSKEEVDRMV EEA EKYAAE DAALREKIEARNALENYAHSLRNQVTDDSETGLGSKLDEDDKE TLTDAIKDTLEFLEDNFD TATKEELDEQREKLSKIAYPITSKL YGAPEGGTPPGGQGFD DDDGDYDYDYDHDEL
PDI1	5	ACF17572.1	MQFNWNIKTVASILSALT LAQASDQEAIAPEDSHVVKLTEATF ESFITSNPHVLAEFFAPWCGHCKKLGPELVSA AEILKDNEQVK IAQIDCTEEKELCQGYEIKGYPTLKVFHGEVEVPSDYQGQRQS QSIVSYMLKQSLPPVSEINATKDLDDTIAEAKEPVIVQVLPED ASNLESNTTFYGVAGTLREKFTFVSTKSTDYAKKYTSDSTPAY LLVRPGEEPSVYSGEELDETHLVHWIDIESKPLFGDIDGSTFK SYAEANIPLAYYFYENEEQRAAAADI IKPFAKEQRGKINFVGL DAVKFGKHAKNLNMDEEKLPLFVIHDLVSNKKFGVPQDQELTN KDVTE LIEKF IAGEAEPIVKSEPIPEIQEEKVFKLVGKAHDEV VFDESKDVLVKYYAPWCGHCKRMAPAYEELATLYANDEDASSK VVI AKLDHTLNDVDNVDIQGYPTLILYPAGDKSNPQLYDGSRD LES LAEFVKERGTHKVDALALRPVEEEKEAE EEA ESEADAHDE L

RPP0	6	CAY67120.1	MGGINEKKAEYFNKLRELLESYKSIFIVGVNDNVSSQQMHEVRQ TLRGKAVILMGKNTMVRKALRDFVEELPVFEKLLPFVVRGNIGF VFTNEDLKTIRDVIIENRVAAPARPGAIAPLDVFI PAGNTGME PGKTSFFQALGVPTKISRGTIEITS DVKVVEKDSRVGPSEAQL LNMLNIS PFTYGLTVVQVFDDGQVFPANILDITDDELLSHFTS AISTIAQISLAAGYPTLPSVGHSVNNHYKNVLAVSIATDYSFE GSEAIKDRLANPEAYAAAAPAAGEASAGAEETAAAAEEDEES EDDDMGFGLFD
BFR2	7	CAY70333	MARKTLAETLAELSQPASGDFDIEDQEGGAVLDYGDNSSFGSE SEEDKSNHYVKVGKSRIRENAVKLGGQYEGKKSSRADVFGDED DEEEDDEDVEHSETEDALSVSGSESESEDEKNSDQSQGDSESEE ESNSGEDLDYKRSKLQQLISSERKTIVNQLSTSNKQDALKGFA VLNQQKQYDQLVDLRIKLQKGLVASNGLPINKEYYEQNKAPKS SKHLDKLQDKLYNLLDVTLELRGKLLNKSIVSQEFPPIPSKK RSLQHYLEESSKLDNIVNEYRRNVLVKWSQKVQNASGATALSS SKFKAINQDSSTQVDNYLADMDRLIKRTRLNRRSVVPLGYTET EEVVDDDELIDNDKDNNETKYFSNIDRSLKENKYIYDDDDFYR VLLNDLVDKKVS DTQKLTSTSTVITFSKSKLHKS YERKATKGR KLRYTVQDPLLNF EASNPHAYKWNDYQIDEFFASLFGQKVMNN EDEHNEEVQGESEGEDILKDDIKLFG
BMH2	8	CAY68707.1	MSREDSVYLAKLAEQAERYEEMVENMKT VASSGLELSVEERNL LSVAYKNVIGARRASWRIVSSIEQKEEAKGNQSQVSLIREYRS KIETELANICEDILSVLSEHLIPSARTGESKVFFYFKMKG DYHR YLAEFVAVGDKRKEAANLSLEAYKSASDVAVTELPPTHPIRLGL ALNFSVFYYEILNSPDRACHLAKQAFDDAIAELET LSEESYKD STLIMQLLRDNLTLWTS DMSETGQEESSNSQDKTEAAPKDEE
Cct2	9	CAY71348.1	MSVNI LGDQVSEERAENARLSAFVGAIAVGDLVKTTLGPKGMD KLLTSASSGQSI VTNDGATILKSIPLDNPAAKVLVNLSKVQDD EVGDGTTSVTVLASELLREAEKLVDRKIHPQTIIIEGFRIASKA ALEALDKVAVDNSHDDAAFRKDLINIAKTTLSSKILAQDRDKF AEIAVSAILRLRGSTSLERIQLIKIIGGQLSDSYLDDGFILNK KFGLDQPKKIKDASILIAN TSMDTDKVKIFGAKFKVDSTSKLA QLEKAEKDKMKAKVEKIKNFNINCFVNRQLIYDWPEQ L LADS NINTIEHADFDGVERLALVTGGEVVSTFDYPGKVKL GKCDLIEE VIIGEEVMTRFSGVSEGA ACTIILRGATEQVLDEAERSLHDAL SVLSQTTKETRTV LGGGCSEMIMSN AVDTQAQNQEGKKQLAVE AFARALRQLPTILADNAGYDSSELVARLRSAIYSGLTTSGLNL SNGTVGDMRQLGVMESYKLKRAVVNSASEAAEVLLRVDNIIRA KPRTADRNR
Erj5	10	CAY67194.1	MKLHLVILCLITAVYCFSAVDREIFQLNHEL RQEYGDNFNFYE WLKL PKGPSSTFEDIDNAYKKLSRKLHPDKIRQKKLSQE QFEQ LKKKATERYQQLSAVGSILRSESKERYDYFVKHGF P VYKNDY TYAKFRPSVLLTIFILFALATLTHFVFIRLSAVQSRKRLSSLI EENKQLAWPQGVQDVTQVKDVKVYNEHLRKWFLVCFDG SVHYV ENDKTFHVDPEEVELPSWQDTLP GKLI VKLI PQ LARKPRSPKE IKKENLDDKTRKTKKPTGDSKTL PNGKTIYKATKSGGRRRK
Gim4	11	XP002491325	MSEGKPNPNQELFQKQYDEFQETLEALNNKIGQLQGDIEEHNI VLKTIT TAPKDRKCFHMIGGV LIEKTAGEVEPTLKT NVTKMND AVENLKNEIQNTHKQFEDWKKKTGVKIVSANE

KIN2	12	CAY70388.1	MDREQGILPQDPFSNSVHVPKLRASSGGQPQKPVIQNSAPATA RMLRNASSSTSAAALLKELNTHEHSQRQHTPQKQPSLDAPAALV PVESATKQFHRTSIGDWEFSNTIGAGSMGKVKVAKHRVTHEVC AIKIVIRSAKIWQRNHQNDPEPETEEKRKKLRDEYKKELERDE RTVREAALGKIMYHPNICRLFECYTMSNHYMLFEIVQGVQLL DYIVSHGKLKETRVRQFARSIASALDYCHSNNIVHRDLKIENI MINNKGEIKLIDFGLSNMYDRRNLLKTFCGSLYFAAPELLSCR PYIGPEIDVWSFGVVLVFLVSGKVPFDDDSVPKLHAKIKRGKV EYPEFISPLCHSLLSQMLLVNPDHRVTLKAAMEHPWMTLG FAG PPSNYLPQRSPIVLPLDLSVVREIANLGLGNEEQIARDITNLI SSREYEACVERWKLDQQKANIKGYSARDDSAIIAFHPLLSTYY LVDEMRRKRLAKGALKGQTSVLDTVKVSPDIPKTPAIPQKLET TDVEQPLLATVPPAYTSPHGQPAELEAMIEPAQPLSSAHPFEM DMTQQQHASRKTHIKHAPERQDRGGYNVHKNNSSGGLNSLFRRL SGKRPHKNEAEWEPSSPPPQVHPFSVNDADRTSVRGVSPITQP AAVKNVTSNNSKNYLDPVDDSKLVRRVGSRLRITNKEKQQVTS D FPRLPNFTIPEQPPKNAPIPIHAQPTTTGTTFQSNDEIKKKL QASTSPNEQRGPPTLAPSQQRR LHPTARAKSLGHSRKQSLNFK FGGPANNQLPALPTKENYDVFEDAQITDNNLLNPEGKYSANTN VHIKPMTESQILFEAEHAPPGTMPSVEYPRTLFLKGFFSVQTT SSKPLPVIRYNIIAALCKLNIQFTEVNGGFVCVYRK TENLQIG DIRSPVIESRVTD DTDSDVANSSKLSSSSTANTRVNVIEDDSS SPSSARLKHRRKFSLGNGILNHIRKPTLDGTEFDDYDATVNT P VTPAPANVHSRSSSYHTESDNESMESLHDIRGGS DMILKNVPE R RNARQIDTVKEEETDDDDLGSINEGSTHRTPLKFEIHIVKVPL VGLYGVRFKKILGNAWIYKRLASKLLQELNL
Mdj1	13	CAY69583.1	MSQRF LQGMNRR LPHLVWLRTKQPLLSCAFQRHPLSKYQARGF HGSAARLISDPYKTLNVDRNASTSDIKKAYYKLAKQYHPDINK EKGAEKKFHDIQAAYEILSDTEKKQQFDQFGTVFDS DGNPMGG SGGRGGPGNPFAGGNPFGAGNPFGNAAGGFSFNLEDLFGDAFN GANRQGGRRAGGAAYMEQYQGNDVEILKTI SFKESIFGTNASV NYNVLDGCNTCEGTGLKKGRKKSTCSTCNGSGASVHYLQGFQM SSTCNACGGTGVTISKDDQCGHCHGNGVGQSSKTTEVKLPCGI RDGTRLRVSGAGDAPNVTKGPNVRTVKGDLIIRVRVKPDPRYS RDGNDIVYNCEIPMTTAALGGQVEIPTLDDTKLRLKVP IGTQH GRTVSI PGQGVPIHGSLSNRGALKVQFNVKVL RPDNATQTALL EALADTFNDTTAKKVNPSWKPFENSAPPAEGEDSDHPSRLKKI ESFLS DAFKRITNKKDDCK
HAC1spliced	14	Guerfal et al. 2010, Microbial Cell Factories 9: 49, Figure 2 (PpHac1)	MPVDSSHKTASPLPPRKRKAT EEEKEQRRVERILNRRAAHAS REKKRRHVEFLENHVVDLESALQESAKATNKLKEIQDIIVSRL EALGGTVSDLDLTVPEVDFPKSSDLEPMSDLSTSSKSEKASTS TRRSLTEDLDEDDVAEYDDEEEDEELPRKMKVLNDKNKSTSIK QEKLNELPSPLSSDFSDVDEEKSTLTHLKLQOQQQQQPV DNYVS TPLSLPEDSVDFINPGNLKIESDENFLLSSNTLQIKHENDTDY ITTAPSGSINDFFNSYDISESNRLHHPAAPFTANAFDLNDFVF FQE
Def1	15	CAY67433.1	MSERSSKKGPKGGAKRSSQGS SGGLESTKLATLTELFDPWTAQ DLEPVLEEYPDEDLNVIIENIISGKINKWTDPSAKKEKKKREE SFNASEELSTPSYHQTPNSAKKEYPKKEVKAKSKKSQPRSTTS TTTASTKAQLTPSSNPSTKSSWAAALHQKQEDKPSSTVTP TTE TETPNGENASQSPVAETKSEQEESFAPAAVVETS AKPKSWAAM VAQSAKPKKKILKRPEQAAKPSSNEELSQQNGEIQDEQQSLQT QAETQAEQPIQSIELQQTNEQISQQEQKPVQEPKPLERKQQQQ QQQQPVVLPSAVNLDSIGGISFGSLSLNEKEASSAQQAAQQASQ PTSQVQAQTQNNQQYQRYENQYYNNNRQFYQDGKQVNYDSFVRQ QQQQQQHQQQQYWAHPQAQAQGVASAGGSDLNSASPAASNALP QGQPQGT PSASNANPVNAYNNPQFYTPYVYYPYGQYYQNPQLY SGYMGYGAGQPQTQPHQPQVPPTASPSQQTQQVQPTSGQVPNQ QLAGFQGYQQPYQQAYLNKNGYPLYQQYPQQQQQQVGGQGQSQ PQGKEVEEPKPQQQGQQAGQHQQQAQLPQQYPGHPGQYFGQQ ALGAQQTPYTEYPVYPNSNDYNNTNAKGWI

Gas1	16	XP002489568.1	MFKSLCMLIGSCLLSSVLAADFPTIEVTGNKFFYSNNGSQFYI KGVAYQKDTSGLSSTATFVDPLADKSTCERDIPYLEELGTNVI RVYAVDADADHDDCMQMLQDAGIYVIADLSQPNNSIITTDPEW TVDLYDGYTAVLDNLQKYDNILGFFAGNEVITNKSNTDTAPFV KAAIRDMKTYMEDKGYRSIPVGYSANDELTRVASADYFACGD SDVKADFYGINMYEWCGKATFSNSGYKDRTAEFKNLSIPVFFS EYGCNEVQPRLFTEVQSPLYGDDMTDVWSGGIVYMYFEETNNYG LVTIKSDGDVSTLEDNFNNLKTELASISPSIATQSEVSATATEI DCPATGSNWKASTDLPPVPEQAACQCMADALSCVVSEDVDTD YSDLFSYVCENVSSCDGVSADSESGEYGSYSFCSSKEKLSFLL NLYYSENGAKSSACDFSGSATLVSGTTASECSSIILSAAGTAGT GSITGITGSVEAATQSGSNSGSSKSSSASQSSSSNAGVGGGAS GSSWAMTGLVSIISVALGMIMSF
LHS1	17	CCA36228.1	MRTQKIVTVLCLLLNTVLGALLGIDYGQEFKAVLVAPGVPFE VILTPDSKRKDNSMMAIKENSKGEIERYYGSSASSVCIRNPET CLNHLKSLIGVSIIDVSTIDYKKYHSGAEMVPSKNNRNTVAFK LGSSVYPVEEILAMSLDDIKSRAEDHLKHAVPGSYSVISDAVI TVPTFFTQSQRALALKDAAEISGLKVVGVLVDDGISVAVNYASSR QFNGDKQYHMIYDMGAGSLQATLVSISSSDDGGIVIDVEAIAY DKSLGGQLFTQSVYDILLQKFLSEHPSFSESDFNKNSKMSKL WQAAEKAKTILSANTDTRVSVESLYNDIDFRATIARDEFEDYN AEHVHRITAPIIEALSHPLNGNLTSFFPLTSLSSVILTGGSTR VPMVKKHLESLLGSELIAKNVNADES AVFGSTLRGVTLSQMFK AKQMTVNERSVYDYCLKVGSSEINVFVPGTPLATKKVVELENV DSENQLTIGLYENGQLFASHEVTDLKKSIKSLTQEGKECSNIN YEATVELSESRLLSLTRLOAKCADEAEYLPPVDTESEDTKSEN STTSETIEKPNKKLFYPVTIPTQLKSVHVKPMGSSTKVSSSLK IKELNKKDAVKRSIEELKNQLESKLYRVRSYLEDEEVVEKGPA SQVEALSTLVAENLEWLDYDSDDASAKDIREKLNSVSDSVAFI KSYIDLNDVTFDNNLFTTIYNTTLNSMQNVQELMLNMSEDALS LMQQYEKEGLDFAKESQKIKIKSPPLSDKELDNLFNTVTEKLE HVRMLTEKDTISDLPREELFKLYQELQNYSSRFEAIMASLEDV HSQRINRLTDKLRKHIERVSNEALKAAALKEAKRQQEEEEKSHEQ NEGEEQSSASTSHTNEDIEEPSESPKVQTSHDEL
Pma1	18	XP002489633.1	MSAEEPTKEKIPINHSDDDEDIDQLIEDLQSVHGFDDDEEEEE HHEGATAKPVPEELLQTDPAYGLTTDEVHKRRKRFGENKMAEE KENLLVKFCMFFVGPIQFVMEAAAILAAGLEDWVDFGVILALL FLNASVGFIQEYQAGSIVDELKKTLAN SATVIRDGQVVDILAD EVVPGDILKLEDGVVIPADGRLVSEECFLQVDQSAITGESLAV DKKTGDSTYSSSTVKRGEAYMVVTATGDSTFVGRAAALVNKAS AGQGHFTEVLNGIGTILLVLVIATLLVVWVACFYRTSPIVRIL RFTLAITIVGVPVGLPAVVTTTMAVGASYLAKKQAIQKLSAI ESLAGVEILCSDKTGTLTKNKLSLHEPYTVEGVEADDLMLTAC LAASRKKKGLDAIDKAFLKSLISYPRAKAALT KYKVIEFQPF PVS KKV TAYVESPEGERIICVKGAPL FVLKTVEEDHPIPEDVH DNYENKVAEFASRGFRSLGVARKRGQGHWEILGIMPCMDPPRD DTAQTVNEATHLGLRVKMLTGDAVGIAKETCRQLGLGTNIYNA ERLGLGGAGDMPGSEIAD FVENADGFAEVFPQHKNVVEILQQ RGYLVAMTGDGVNDAPSLKKADTGIAVEGASDAARSAADIVFL APGLSAIIDALKTSRQIFHRMYSYVVYRIALSLHLELFLGLWI AIMNRS LNIDL VVFIAIFADVATLAIAYDNAPYSPKPTKWNLP RLWGMSIILGIILAIGTWITLT TMLLPRGGIIQNF GSV DGVLF LEISLTENWLIFITRAAGPFWSSCPSWELAGAVIIVDIIATMF TLFGWWSQNWTDIVTVVRVWIFSF GVF CVMGGAYYLMSESEGF DRLMNGKPRKEPPPQRSMEDFIVAMQRVSTQHEKSG

SSE1	19	CAY67046.1	MSVPFGVDLGNNNTVIGVARNRGIDILVNEVSNRQTPSIVGFG AKSRAIGESGKTQQNSNLKNTVEHLVRI LGLPADSPDYEIEKK FFTSPLIEKDNEILSEVNFQGGKTTFTPIQLVAMYLNKIKNTA IKETKGKFTDICLAVPVWFTEKQRSAASDACKVAGLNPVRIVN DITAAAVGYGVFKTDLPEDEPKKVAIVDIGHSTYSVLIAAFKK GELKVLGSASDKHFGGRDFDYAITKHFAEEFKSKYKIDITQNP KAWSRVYTAAERLKKVLSANTTAPFNVESVMNDVDVSSSLTRE ELEKLVQPLLDRAHIPVERALAMAGLKAEDVDTVEVVG GCTRV PTLKATLSEVFGKPLSFTLNQDEAIARGAA FICAMHSPTLRVR PFKFEDVNPYSVSYWDKDPAAEDDDHLEVFPVGGSF PSTKVI TLYRSQDFNIEARYTDKNALPAGTQEFIGRWSIKGVV VNEGED TIQTKIKLRNDPSGFHIVESAYTVEKKT IQEPIEDPEADED AE PQYRTVEKLVKKN DLEITGQTLHLPDELLNSYLETEAALE VQD KLVADTEERKNALEEYIYELRGKLEDQYKEFASEQEKT KLTAK LEKAEEWLYDEGYDSTKAKYIAKYEELASIGNVIRGRYLAKEE EKKQAI REKEESKKASAI AEKMAAERASREAAGSTNEQAQKNE ENTK DADGDVSMNQDELD
Sti1	20	XP002491431.1	MSSEEFKAQGNQAFQAKDYEKAVSFFTQAIEASPTPNHILFSN RSAAYASLGQYQDALDDANKCVEINGSWAKGYNRVGA AHYGRG EWDEAHKAYSKALELDPANKMAKEGLNETEIARDAGNDVK NIF SDAGMVEKCLKKNPKTAELMKDPELVAKVQKLQTDPKSMSQELF SDPRLMTVMGAMLGVDLGVQPSQQSAPQEDTPVPDAYPEPSSK PETNTTSAKNAAPEPEKEATPEPVDNSKEEADNLKQQANQLY KKRQFDEAIELYNKAWETFQDITYLNNRAAAEF EKGDYDATIE TCENAVEKGREL RADYKLVAKSFARLGSAYLKKDDL PNAIKFF EKSLTEHRSPDVLSKLRAAEADLKKKEAEEYIDPEKAEEEARLQ GKDFFTKGDWPAAVKAYTEMINRAPKDARGYSNRAAALAKLMS FPDAVKDCDKAIELDPSFVRAYIRKATALIAMKDFNKAMTTLE EARTVDADTNEGKAANEINGLYYKASSQRFAAIDGETPEQTFE RASKDPEVSAI LQDPVMNSILOQ QARENPAALQE HMKNPEVAKK INILIAAGVIRTR
Uso1	21	XP002493742.1	MTTPIAQIQLEQEASKNPPKQHTRLSDLVEKTKGTSWVSPFR TDAKAASPKRESYPPQIVADV KPEDVDNAEEETILDHDDANAT VDPIESESVL D ASDISIKGSTAEDNQEEQPEPATDVLPQDAEE EVADKDTQSGDIPQDEGSQAEQEEEQAPEAQEEQVSESQEAKE DDKVDNVEAKKDVADKKVTKQTQQA IKDTEEGAKAVKEAQAKL KEAELKLLKEPVVITPDLLQPPAEDDAEKT LKDKPLLLNRYKQ NKEIAESSLQKKDVENPDQVVDLGGG LLLTQAQIYSIAQARVK PLLGKIDKQVDLNLKADELKKRQTEQQYHEQKDLQQSKNLEKY QTQLTRENNIIVARFDTDIAALSSTILSNATLLEEFATQTRKE IDDLGTKALAEEEKLAEEHETNKTKLEENAKQYKEDLET KLLN ATTGQEDEKTKIEELKVKVEEEKAIADDLEEKAFDKNEALNAK RAELEELVAEEAKLQATVDESEQFQKECDAKAAALSVDHTKST KKLEKLQSHVSALGSAIEKHASKIGFLTGA AVASREVKRKHNE SLKSEWLAEKARIRSEVAKANERKTLEAE LERERLAKEKEIER QQKEEQYAQEKLDRAEE EKRLKEDVAELQRVKQLKKEKSKLSK KLASTGSFFAGGVATGAAIGAATGAAAGSAAGAAASGAGAAAS GASKVVSSSTNTASKGASDAAQVGNGAKKTADIKRNESFASNS PEIKIDDET LNKDAKPLFTEVVEDVPTTTSKADEDI KKKNRLS FLGSIKRKASLGSKKEPEKKEPATGVVPASSSIAKDND DGEYE EVSTLETISDAEYEAHKDDPNYFIVDPK
Ydj1	22	XP002492146.1	MVRETKLYDILGVSPDATDAQLKKAYRVGALKNHDPKNPSPEA AETFKGMSHAYEVLSDPQKREIYDQYGE EGLNGGGAGPGGMGE DIFSQFFGGMFPGGGQPTGPPQRGKDIKHSISCTLEELYKGR TA KLALNKT V LCKECDGKGKGNVKKCSACNGQGLRFVTRQIGPMI QRAQVRCDVCN GEGDIISGADRCKACSGKKITNERKILEVNIE RGMRHGQKVVFSGESDQAPDVI PGDVIFVVDEKPHKDFSRKGD DLYYEAKIDLLTALAGGELAIKHISGEY LKITIIPGEVISPGS VKVIVGKGMPVRKSSSYGNLYVKFEIDFP PKNFTTAENLQ LLE QVLPARTPVSI PADAEVDEVVLADV DPTQQQRQGGRGGSYDS DDEEQGGQGVQCASQ

Table A-3: Primers used in genomic DNA PCR for identification of the auxiliary proteins that exert a positive effect on the expression yield of Nanobody A in *P. pastoris*

General forward primer	SEQ ID NO	Sequence
FW-AOX promoter	23	GACTGGTTCCAATTGACAAGC
Specific reverse primers		Sequence
RV-FkpA	24	GTCGTGGGCGCGCCT TTTTGGCGCTATCTGCGG
RV-SKP	25	GTCGTGGGCGCGCCT TTTGACTTGCTTCAGCACGT
RV-Ero	26	AGCTGGCGGCCGC TTACAAGTCTACTCTATATGTGGTA
RV-Kar2p	27	AGCTGGCGGCCGC CTACAACATCATGATCATAGTCA
RV-PDI1	28	AGCTGGCGGCCGC TTAAAGCTCGTCGTGAGCGTCTGCC
RV-RPP0	29	AGCTGGCGGCCGC TTAATCAAACAAACCGAATCCCATG
RV-BFR2	30	AGCTGGCGGCCGC TTATCCAAACAGTTTGATATCATCC
RV-BMH2	31	AGCTGGCGGCCGC TCACTCTTCATCTTTGGGAGCAGCT
RV-Cct2	32	AGCTGGCGGCCGC TCAACGATTACGGTCGGCAGTGCGT
RV-Erj5	33	AGCTGGCGGCCGC TTATTTCTTCTACGTCCACCGGAT
RV-Gim4	34	AGCTGGCGGCCGC CTACTCATTAGCACTCACAATCTTG
RV-KIN2	35	AGCTGGCGGCCGC CTATAAATTCAATTCTTGTAGCAGC
RV-Mdj1	36	AGCTGGCGGCCGC CTATTTACAGTCGTCCTTCTTATTG
RV-HAC1spliced	37	ATGCATTAGCGGTAAATGGTGCTGCTGGATGATGCAACCGATTG
RV-Def1	38	AGCTGGCGGCCGC TTAAATCCACCCTTTAGCATTTG
RV-Gas1	39	AGCTGGCGGCCGC TTAGAATGACATAATCATTCCA
RV-LHS1	40	AGCTGGCGGCCGC CTACAACATCATGGGATGT
RV-Pma1	41	AGCTGGCGGCCGC TTAACCAGACTTCTCGTGCTGA
RV-SSE1	42	AGCTGGCGGCCGC TTAATCTAGCTCATCTTGGTTC
RV-Sti1	43	AGCTGGCGGCCGC TTAACGAGTACGAATGACACC
RV-Uso1	44	AGCTGGCGGCCGC TTATTTGGGATCGACGATGAAA
RV-Ydj1	45	AGCTGGCGGCCGC TTACTGAGAAGCACATTGGACAC

Table A-4: CDRs and framework sequences of TNF binding Nanobody I. CDR1, CDR2 and CDR3 were determined according to Kontermann, 2010

	SEQ ID NO	
FR1	57	DVQLVESGGGVVQPGGSLRLSCTAS
CDR1	58	GFTFSTADMG
FR2	59	WFRQAPGKGREFVA
CDR2	60	RISGIDGTTY
FR3	61	YDEPVKGRFTISRDN SKNTVY LQMNSLRPEDTALYYCRS
CDR3	62	PRYADQWSAYDY
FR4	63	WGQGTLVTVSS

CLAIMS

1. A method for the production of a polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain, said method comprising the step of expressing, in a *Pichia* host, said polypeptide and simultaneously enhancing, in said *Pichia* host, the expression of the auxiliary protein HAC1spliced (SEQ ID NO: 14).
2. The method of claim 1, further comprising the step of isolating and/or purifying the polypeptide.
3. The method of any of claims 1 or 2, further comprising the step of formulating the polypeptide in a buffer.
4. The method of any of claims 1 to 3, further comprising the step of formulating the polypeptide as a pharmaceutical composition.
5. A *Pichia* host that expresses, or that under suitable circumstances is capable of expressing, a polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and wherein the expression of the auxiliary protein HAC1spliced (SEQ ID NO: 14) is enhanced.
6. The method according to any of claims 1 to 4 or the *Pichia* host according to claim 5, wherein expression of the HAC1spliced protein is enhanced by introduction, into the *Pichia* host, of:
 - one or more nucleic acid(s) encoding HAC1spliced protein; and/or
 - one or more strong promoter(s) controlling the expression of a nucleic acid encoding HAC1spliced protein.
7. The method or *Pichia* host according to claim 6, wherein the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and the HAC1spliced protein are expressed from the same genetic construct.

8. The method or *Pichia* host according to claim 7, wherein the nucleic acid encoding the HAC1spliced protein is located on the genetic construct downstream of the nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain.
9. The method or *Pichia* host according to any of claims 7 or 8, wherein transcription of the nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and transcription of the nucleic acid encoding the HAC1spliced protein is controlled by the same promoter.
10. The method or *Pichia* host according to any of claims 7 or 8, wherein transcription of the nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and transcription of the nucleic acid encoding the HAC1spliced protein is controlled by a different promoter.
11. The method or *Pichia* host according to claim 6, wherein the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and the HAC1spliced protein are expressed from different genetic constructs.
12. The method or *Pichia* host according to claim 11, wherein transcription of the nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and transcription of the nucleic acid encoding the HAC1spliced protein from the different genetic constructs is controlled by separate promoters which are the same.
13. The method or *Pichia* host according to claim 11, wherein transcription of the nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and transcription of the nucleic acid encoding the HAC1spliced protein from the different genetic constructs is controlled by separate promoters which are different.

14. The method or *Pichia* host according to any of claims 6 to 13, wherein the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and/or the HAC1spliced protein are expressed from the chromosome.
15. The method or *Pichia* host according to claim 14, wherein expression of the HAC1spliced protein is enhanced by introduction of a strong promoter into the chromosome of the *Pichia* host.
16. The method or *Pichia* host according to claim 14, wherein one or more nucleic acids encoding the HAC1spliced protein controlled by a strong promoter are introduced into the chromosome of the *Pichia* host.
17. The method or *Pichia* host according to claim 14, wherein one or more nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain is introduced into the chromosome of the *Pichia* host.
18. The method or *Pichia* host according to any of claims 1 to 17, wherein the copy number of nucleic acid(s) encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain is one.
19. The method or *Pichia* host according to any of claims 1 to 17, wherein the copy number of the nucleic acid(s) encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain is 2 or more.
20. The method or *Pichia* host according to any of claims 1 to 19, wherein the *Pichia* host is *Pichia pastoris*.

21. The method or *Pichia* host according to claim 20, wherein the *Pichia pastoris* strain is selected from X33 and NRRL Y-11430.
22. A nucleic acid encoding a polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and encoding HAC1spliced (SEQ ID NO: 14).
23. The nucleic acid according to claim 22, that is in the form of a genetic construct.
24. A genetic construct that comprises a nucleic acid encoding a polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and a nucleic acid encoding HAC1spliced (SEQ ID NO: 14).
25. The genetic construct according to claim 24, wherein the nucleic acid encoding the HAC1spliced protein is located downstream of the nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain.
26. The genetic construct according to any of claims 24 or 25, wherein expression of the nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and expression of the nucleic acid encoding the HAC1spliced protein is controlled by the same promoter.
27. The genetic construct according to any of claims 24 or 25, wherein expression of the nucleic acid encoding the polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain and expression of the nucleic acid encoding the HAC1spliced protein is controlled by a different promoter.
28. The *Pichia* host according to any of claims 5 to 21, that comprises a genetic construct according to any of claims 24 to 27.

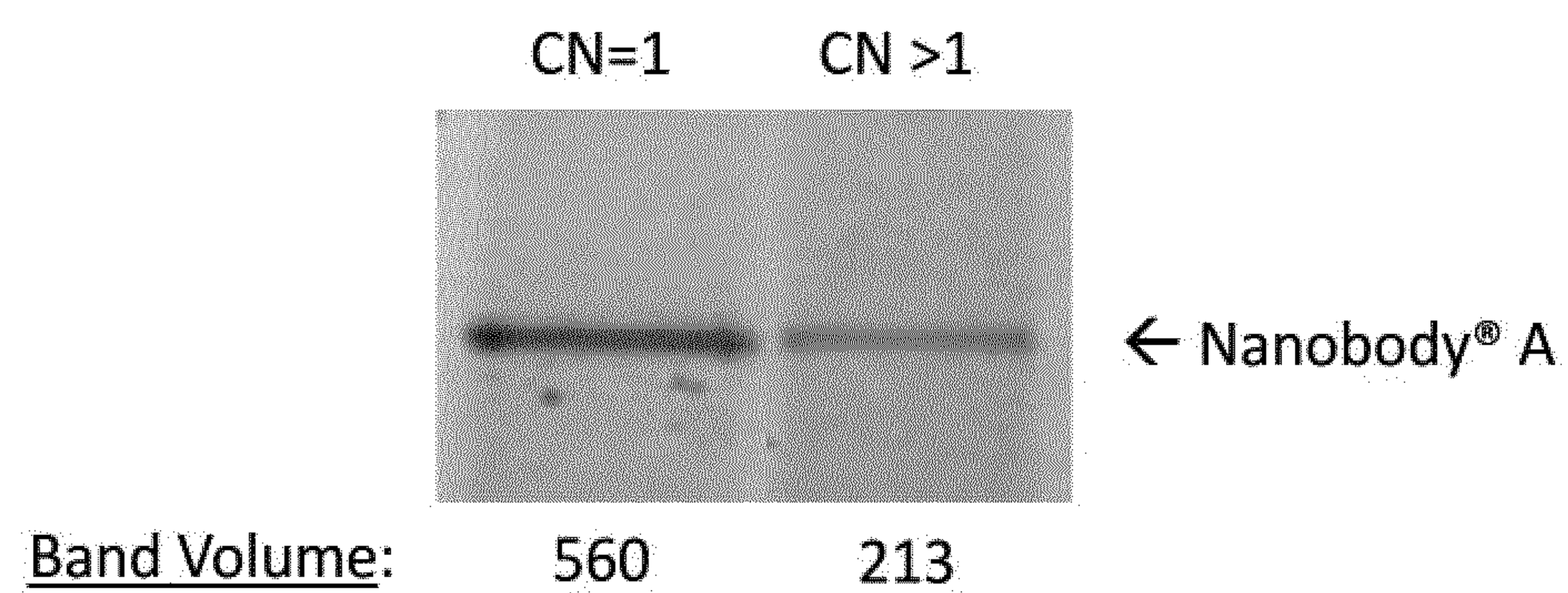
29. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 1 to 28, wherein the expression of one or more additional auxiliary proteins is enhanced.
30. The method, nucleic acid, genetic construct or *Pichia* host according to claim 29, wherein the additional auxiliary protein is selected from PDI1, Kar2p, and RPP0.
31. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims claim 29 or 30, wherein the number of auxiliary proteins is two.
32. The method, nucleic acid, genetic construct or *Pichia* host according to claim 31, wherein the two auxiliary proteins are selected from the following combination of two auxiliary proteins:
- PDI1 (SEQ ID NO: 5) and HAC1spliced (SEQ ID NO: 14);
 - Kar2p (SEQ ID NO: 4) and HAC1spliced (SEQ ID NO: 14); and
 - RPP0 (SEQ ID NO: 6) and HAC1spliced (SEQ ID NO: 14).
33. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims claim 29 or 30, wherein the number of auxiliary proteins is three.
34. The method, nucleic acid, genetic construct or *Pichia* host according to claim 33, wherein the three auxiliary proteins are selected from the following combination of three auxiliary proteins:
- PDI1 (SEQ ID NO: 5), RPP0 (SEQ ID NO: 6) and HAC1spliced (SEQ ID NO: 14);
 - Kar2p (SEQ ID NO: 4), RPP0 (SEQ ID NO: 6) and HAC1spliced (SEQ ID NO: 14); and
 - PDI1 (SEQ ID NO: 5), Kar2p (SEQ ID NO: 4) and HAC1spliced (SEQ ID NO: 14).
35. The method, nucleic acid, genetic construct or *Pichia* host according to claim 29, wherein the expression is enhanced of PDI1 (SEQ ID NO: 5), Kar2p (SEQ ID NO: 4), RPP0 (SEQ ID NO: 6) and HAC1spliced (SEQ ID NO: 14).
36. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 1 to 35, wherein the polypeptide is a monovalent polypeptide.

37. The method, nucleic acid, genetic construct or *Pichia* host according to claim 1 to 36, wherein the immunoglobulin single variable domain essentially consists of 4 framework regions (FR1 to FR4, respectively) and 3 complementarity determining regions (CDR1 to CDR3, respectively), in which CDR1 is SEQ ID NO: 58, CDR2 is SEQ ID NO: 60 and CDR3 is SEQ ID NO: 62.
38. The method, nucleic acid, genetic construct or *Pichia* host according to claim 1 to 37, wherein the immunoglobulin single variable domain is SEQ ID NO: 56.
39. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 1 to 38, wherein the polypeptide further comprises one or more other residues or binding units, optionally linked via one or more peptidic linkers.
40. The method, nucleic acid, genetic construct or *Pichia* host according to claim 39, wherein said one or more other residues or binding units are immunoglobulin single variable domains.
41. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 39 or 40, wherein the polypeptide is a multivalent construct.
42. The method, nucleic acid, genetic construct or *Pichia* host according to claim 41, wherein the polypeptide is a bivalent, trivalent or tetravalent polypeptide.
43. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 39 to 42, wherein said one or more other binding units provide the polypeptide with increased half-life, compared to the polypeptide without said one or more binding units.
44. The method, nucleic acid, genetic construct or *Pichia* host according to claim 43, wherein said one or more other binding units that provides the polypeptide with increased half-life are chosen from the group consisting of binding units that can bind to

serum albumin (such as human serum albumin) or a serum immunoglobulin (such as IgG).

45. The method, nucleic acid, genetic construct or *Pichia* host according to claim 44, wherein the polypeptide comprises at least one immunoglobulin single variable domain that binds serum albumin.
46. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 1 to 45, wherein the polypeptide further comprises a C-terminal extension X(n), in which n is 1 to 5, such as 1, 2, 3, 4 or 5, and in which X is a naturally occurring amino acid, preferably no cysteine.
47. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 1 to 46, wherein the polypeptide comprises or essentially consists of SEQ ID NO: 55.
48. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 1 to 47, wherein the at least one immunoglobulin single variable domain is selected from a domain antibody, a single domain antibody, a "dAb" or a Nanobody.
49. The method, nucleic acid, genetic construct or *Pichia* host according to claim 48, wherein the at least one immunoglobulin single variable domain comprises two or more disulfide bridges.
50. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 39 to 49, wherein the one or more binding units are selected from a domain antibody, a single domain antibody, a "dAb" or a Nanobody.
51. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 48 or 50, wherein the Nanobody is selected from a VHH, a humanized VHH, and a camelized heavy chain variable domain.

52. The method, nucleic acid, genetic construct or *Pichia* host according to claim 51, wherein the VHH belongs to the group of VHH1 type immunoglobulin single variable domains.
53. The method, nucleic acid, genetic construct or *Pichia* host according to any of claims 1 to 52, wherein the polypeptide is selected from the group of polypeptides for which, upon expression in a *Pichia* host under standard *Pichia* expression conditions, the yield obtained shows an inverse correlation with the copy number of the nucleic acid encoding said polypeptide in the *Pichia* host.
54. A polypeptide comprising or essentially consisting of at least one immunoglobulin single variable domain, obtainable by any of the methods of claims 1 to 4, 6 to 21 or 29 to 53.
55. Pharmaceutical composition comprising the polypeptide of claim 54.

Figure 1

2/3

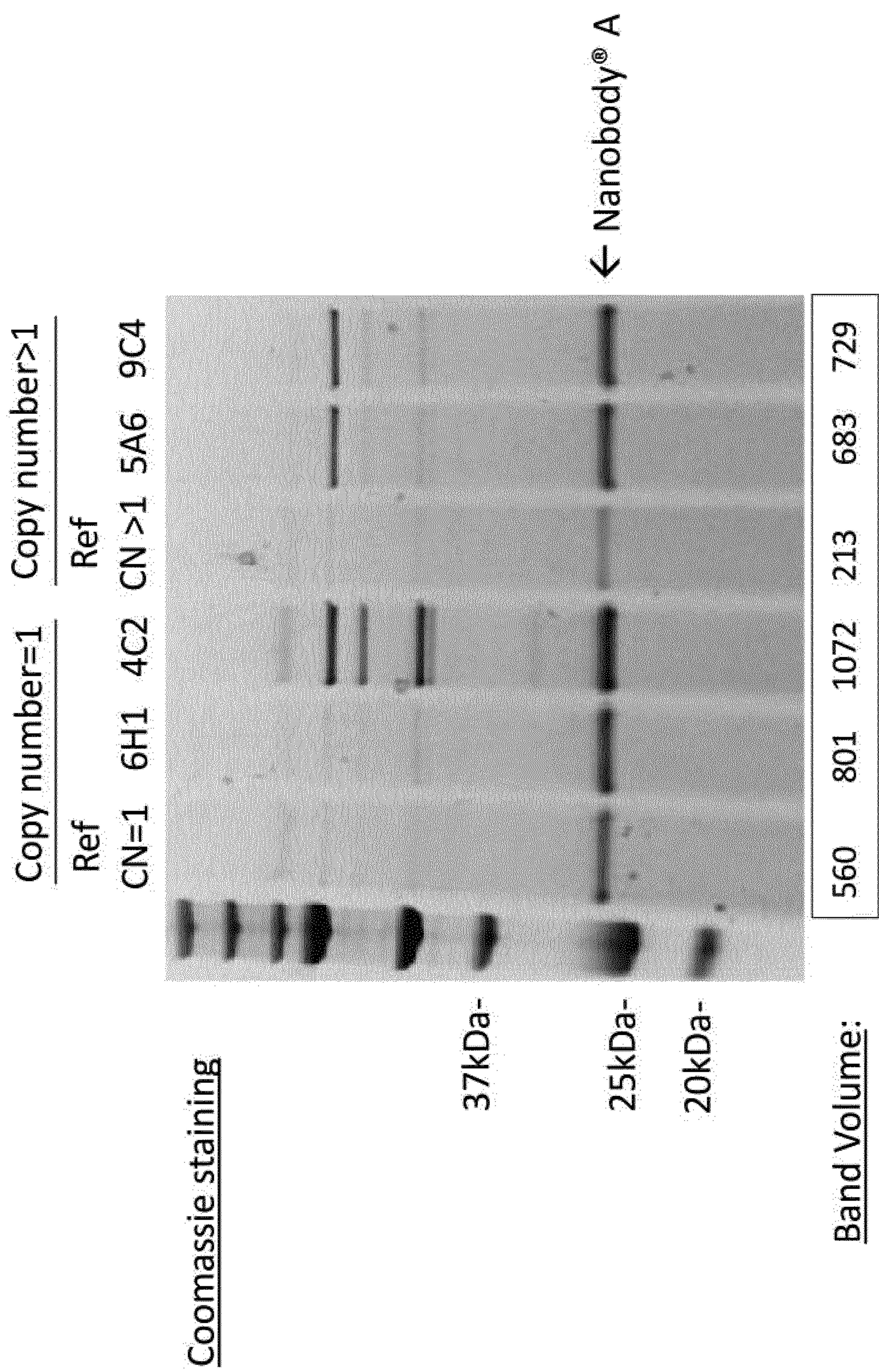
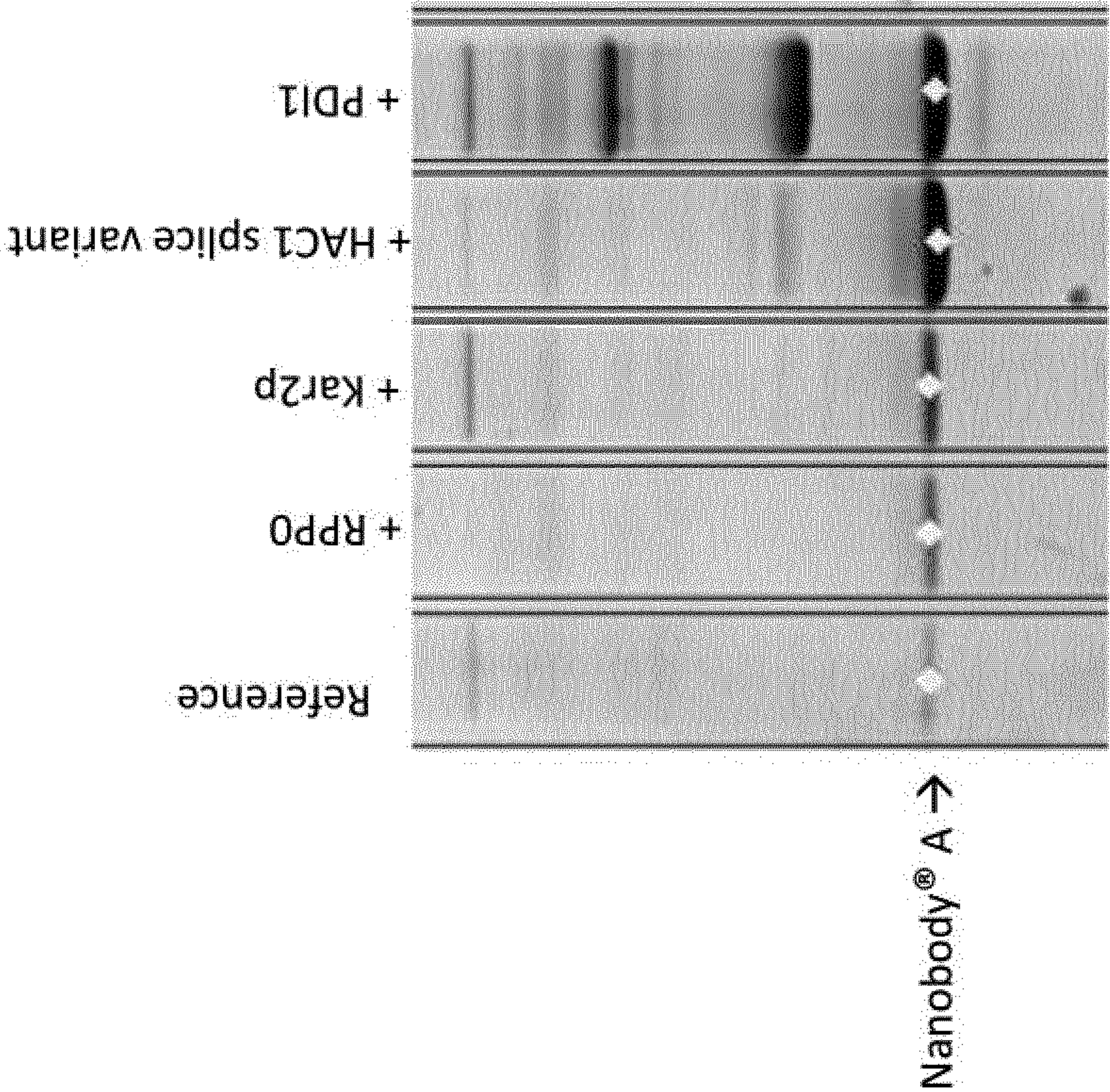


Figure 3



	Strain	Yield
Nanobody® A	Reference	0.1
	HAC1splice variant	1.5
	KAR2p	0.5
	PDI1	1.4
	RPP0	0.3