TRANSAMIDATION EMPLOYING SORTASE A IN DEEP EUTECTIC SOLVENTS

Herein is reported a transamidation reaction using sortase in deep eutectic solvents.

Background of the Invention

Deep eutectic solvents (DESs) are water free solutions with melting points below 100°C. DESs are based on hydrogen bonds. In DES the properties of water and organic solvents are combined. Most DESs are liquid at suitable temperatures for biocatalysis. DESs are more hydrophobic than water however in some cases they do not lead to hydrophobicity induced inactivation of the biocatalyst (Gorke, J., et al., Biotechnol. Bioprocess E 15 (2010) 40-53). Additionally deep eutectics solvents can be produced water free and are therefore inert for hydrolysis. Ionic liquids have similar properties as DES, but are more difficult to produce and more harmful to the environment. Proteases have shown activity in DES (Zhao, H., et al., J. Mol. Catal. B-Enzym 72 (2011) 163-167).

Additionally, some studies suggest that DES are more enzyme compatible solvents than the conventional organic ones (Gorke, J. T., et al., Chem. Commun. (2008) 1235-1237; Lindberg, D., et al., J. Biotechnol 147 (2010) 169-171).

The DES which showed the best performance for several enzyme catalyzed reactions is a mixture of choline chloride and glycerol in the molar ratio of 1:2 (Gorke, J. T., et al., Chem. Commun. (2008) 1235-1237; Zhao, H., et al., J. Mol. Catal. B-Enzym 72 (2011) 163-167).Sortase A (SrtA) is a membrane bound enzyme which attaches proteins covalently to the bacterial cell wall. The specific recognition motif on the SrtA substrate is LPXTG, whereby the enzyme cleaves between the residues threonine and glycine. The recognition motif on the peptidoglycan is a pentaglycine motif. It has been shown that a triglycine and even a diglycine motif on the N-terminus is sufficient to support the SrtA reaction (Clancy, K.W., et al., Peptide science 94 (2010) 385-396). The reaction proceeds through a thioester acyl-enzyme intermediate, which is resolved by the attack of an amine nucleophile from the oligoglycine, covalently linking peptidoglycan to a protein substrate and regenerating SrtA. SrtA can be used to covalently conjugate chemically synthetized peptides to recombinantly expressed proteins.

Applicable Sortases for technical bioconjugation are limited. The most wildly used Staphylococcus aureus Sortase A (St.au. SrtA) shows suitable conversion kinetics

for technical application but has a limited substrate spectrum, only recognizing LPXTG sortase-motives. The St.au. SrtA, that lacks the N-terminal membrane-anchoring motif, has been used for cell-surface protein labeling, covalent protein immobilization and incorporation of novel functionality into proteins. For orthogonal/dual labeling strategies, sortases with new substrate spectra are needed. The same holds true for standard sortase mediated bioconjugation approaches where a LPXTG motive in the product has e.g. negative effects on its structure and/or function. Therefore sortases with recognition sequences different from LPXTG would be of high value. The Staphylococcus pyogenes SrtA (St.py. SrtA) recognizes a LPXTA sortase-motives, however the conversion kinetic parameter of the enzyme turn it in to a not suitable sortase on a technical scale.

Sortases that accept sortase-motives different from LPXTG are reported in literature. Thereunder are wild-types e.g. Sortase A from Streptococcus pyogenes (St.py. SrtA) and Sortase A from Clostridium difficile (Cl.di. SrtA) as well as engineered sortase. Beside the St.py. SrtA none of the reported sortase recognizes a LPXTA motif (see e.g. van Leeuwen, H. C., et al., FEBS Lett. 588 (2014) 4325-4333; Dorr, B. M., et al., Proc. Natl. Acad. Sci. USA 111 (2014) 13343-13348; Bentley, M. L., et al., J. Biol. Chem. 282(2007) 6571-6581; Race, P. R., et al., J. Biol. Chem. 284 (2009) 6924-33; Antos, J. M., et al., J. Am. Chem. Soc. 131 (2009) 10800-10801).

Sortase reactions are performed in aquatic solutions which has several drawbacks. One is that many compounds have a low solubility in water - this is especially true for many fluorophores. Additionally sortases have a very low Km making hydrophobic substrate not accessible for sortase reactions (Chen, I., et al., Proc. Natl. Acad. Sci. USA 108 (2011) 11399-11404). However, antibody/binding-proteins linked to a fluorophore are probably the most important application in bioconjugation (Shreve, P. and Aisen, A. M., Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 3 (1986) 336-340; Drapkin, R. L., et al., Am. J. Hematol. 7 (1979) 163-172). Another issue is the enzyme intermediate which is formed during the reaction. It can be hydrolyzed which leads to product loss. This makes water free and organic solvents an interesting alternative to water. However sortases are not stable in organic solvents e.g. more than 20 % Dimethyl sulfoxide (DMSO) diminishes sortase activity (Pritz, S. (2008) Enzymatische Ligation von Peptiden, Peptidnucleinsäuren und Proteinen).

In WO 2010/087994 methods for ligation and uses thereof are reported. Recombinant approaches to IgG-like bispecific antibodies are reported by Marvin, J.S., et al. (Acta Pharmacol. Sinica 26 (2005) 649-658). In WO 2013/003555 the use of sortases to install click chemistry handles for protein ligation is reported.

Strijbis, K. et al (Traffic 13 (2012) 780-789) report protein ligation in living cells using sortase. It has been stated by them that the Ca^{2+}-dependent S. aureus sortase A is not functional intracellularly, but that the Ca^{2+}-independent S. pyogenes sortase A is functional in the cytosol and endoplasmic reticulum (ER) lumen of both Saccharomyces cerevisiae and mammalian HEK293T cells.

Levary, D.A., et al., report protein-protein fusion catalyzed by Sortase A (PLOS ONE 6 (2011)). Engineering of an anti-epidermal growth factor receptor antibody to single chain format and labeling by sortase A-mediated protein ligation is reported by Madej, M.P., et al. (Biotechnol. Bioeng. 109 (2012) 1461-1470). Ta, H.T., et al., report enzymatic single-chain antibody tagging as a universal approach to targeted molecular imaging and cell homing in cardiovascular diseases (Cir. Res. 109 (2011) 365-373). Popp, M., et al., report making and breaking peptide bonds - protein engineering using sortase (Angew. Chem. Int. Ed. Eng. 50 (2011) 5024-5032). Engineered proteins with high affinity for DOTA chelates are reported in WO 2010/099536.

Different efforts to block the revers reactions of Sortase have been reported. Yamamura, Y., et al. (Chem. Commun. 47 (2011) 4742-4744) reported enhancement of sortase A-mediated protein ligation by inducing a beta-hairpin structure around the ligation site by introducing a β-hairpin around the recognition site of the substrate.

Sorting of LPXTG peptides by archetypal sortase A, role of invariant substrate residues in modulating the enzyme dynamics and conformational signature of a productive substrate was reported by Biswas, T., et al. (Biochem. 53 (2014) 2515-2524).

Li, Y.M., et al. report irreversible site-specific hydrazinolysis of proteins by use of Sortase (Angew. Chem. Int. Ed. Engl. 53 (2014) 2198-2202).

Ling and co-workers showed the introduction of a thioester via a sortase (Ling, J. J. J., et al., J. Am. Chem. Soc. 134 (2012) 10749-10752).

Lindberg, D., et al. (J. Biotechnol. 147 (2010) 169-171) reported deep eutectic solvents (DESs) are viable co-solvents for enzyme-catalyzed epoxide hydrolysis. Gorke, J.T., et al. reported hydrolase-catalyzed biotransformations in deep eutectic solvents (Chem. Commun. (2008) 1235–1237).

In US 8,247,198 enzymatic processing in deep eutectic solvents is reported.

Bellucci, J.J., et al. report the use of lysine as nucleophile (Angew. Chem. Int. Ed. Engl. 53 (2014) 1-6).

WO 2013/016653 provides methods for detecting the concurrent presence of at least two targets within a biological sample. The method includes contacting said biological sample with a first binding agent, said first binding agent operably linked to a first sortase molecule, wherein said first binding agent specifically binds to a first target; contacting said biological sample with a second binding agent, said second binding agent operably linked to a first sortase recognition sequence peptide, wherein said second binding agent specifically binds to a second target; adding a sortase substrate under conditions where a first sortase-mediated ligation of the sortase substrate to the first sortase recognition sequence will produce a ligation product, and detecting the ligation product, wherein detection of said ligation product indicates the concurrent presence of the first target and the second target in the biological sample.

Schmohl, L. and Schwarzer, D., reported about sortase-mediated ligations for the site-specific modification of proteins (Curr. Opin. Chem. Biol. 22 (2014) 122-128).

Summary

It has been found that a transpeptidation reaction can be performed in deep eutectic solvents. It has further been found that the reduction or even absence of water is not detrimental to the reaction but suppresses hydrolysis side reaction.

As reported herein is a method for the enzymatic production of a polypeptide comprising the following step

- incubating

  i) a first polypeptide (optionally comprising within the 100 C-terminal amino acid residues) the amino acid sequence LPXTG (SEQ ID NO:

01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue),

ii) a second polypeptide comprising i) a glycinyl, an alaninyl, or a cysteinyl compound at its N-terminus (i.e. a compound that comprises a cysteine amino acid residue with free alpha amino group, e.g. as NH_2 or NH_3^+, and a carboxy group, which is part of a peptide bond, at position 1, or a compound that comprises an alanine amino acid residue with free alpha amino group, e.g. as NH_2 or NH_3^+, and a carboxy group, which is part of a peptide bond, at position 1, or a compound that comprises a glycine amino acid residue with free alpha amino group, e.g. as NH_2 or NH_3^+, and a carboxy group, which is part of a peptide bond), at position 1, or ii) an oligoglycine, or oligoalanine, or a cysteine amino acid residue followed by one to three glycine or alanine amino acid residues at its N-terminus, or iii) a lysine amino acid residue within its 5 N-terminal amino acid residues, and

iii) a third polypeptide with sortase A activity (i.e. a third polypeptide that is a sortase A or a catalytically active fragment thereof, i.e. that has sortase A activity),

in a deep eutectic solvent

and thereby producing a polypeptide.

One aspect as reported herein is a method for the enzymatic production of a polypeptide comprising the following step

- incubating

i) a first polypeptide comprising the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue),

ii) a second polypeptide that has an oligoglycine, or oligoalanine, or a cysteine amino acid residue followed by one to three glycine or alanine amino acid residues at its N-terminus, and

iii) a third polypeptide with sortase A activity wherein the third polypeptide

a) has the amino acid sequence of Staphylococcus aureus sortase A or Listeria monocytogenes Sortase A with 0 to 10 amino acid sequence changes, or

b) is a shortened variant of the amino acid sequence of Staphylococcus aureus sortase A or Listeria monocytogenes Sortase A,

in a deep eutectic solvent which is a mixture of choline chloride with glycerol at a molar ratio of 1:2 comprising up to 5 % (v/v) aqueous co-solvent

and thereby producing a polypeptide.

In one embodiment the third polypeptide with sortase A activity is derived from Staphylococcus aureus sortase A, or from Streptococcus pyogenes Sortase A, or from Listeria monocytogenes Sortase A.

In one embodiment the third polypeptide is derived from Staphylococcus aureus sortase A, or from Streptococcus pyogenes Sortase A, or from Listeria monocytogenes Sortase A.

In one embodiment the third polypeptide with sortase A activity is derived from Staphylococcus aureus sortase A, or from Streptococcus pyogenes Sortase A, from Listeria monocytogenes Sortase A, or a catalytically active fragment thereof.

In one embodiment the third polypeptide comprises the amino acid sequence of SEQ ID NO: 05, SEQ ID NO: 06, or SEQ ID NO: 38. In one preferred embodiment the third polypeptide comprises the amino acid sequence of SEQ ID NO: 38.

In one embodiment the third polypeptide comprises additionally a tag at its N- or C-terminus either conjugated directly or via an intervening linker. In one embodiment the third polypeptide is consisting of the amino acid sequence of SEQ ID NO: 38 and the C-terminal tag of SEQ ID NO: 32. In one embodiment the third polypeptide is consisting of the amino acid sequence of SEQ ID NO: 38.

In one embodiment the method is for the enzymatic conjugation of two polypeptides. In one embodiment the method is for the enzymatic transpeptidation of two polypeptides. In one embodiment the first or the second or both polypeptides are insoluble in water (100 % (v/v)).

In one embodiment the deep eutectic solvent comprises choline chloride. In one embodiment the deep eutectic solvent is a mixture of choline chloride with glycerol at a molar ratio of 1:2. In one embodiment the deep eutectic solvent comprises an aqueous co-solvent. In one embodiment the deep eutectic solvent comprises up to 30 % (v/v) aqueous co-solvent. In one embodiment the deep eutectic solvent comprises from (and including) 20 % (v/v) aqueous co-solvent to (and including) 30 % (v/v) aqueous co-solvent. In one preferred embodiment the deep eutectic solvent comprises about 25 % (v/v) aqueous co-solvent. In one embodiment the deep eutectic solvent comprises up to 15 % (v/v) aqueous co-solvent. In one embodiment the deep eutectic solvent is a mixture of choline chloride with glycerol at a molar ratio of 1:2 comprising up to 5 % (v/v) aqueous co-solvent. In one embodiment the deep eutectic solvent is a mixture of choline chloride with glycerol at a molar ratio of 1:2 comprising up to 30 % (v/v) aqueous co-solvent.

In one embodiment the incubating is at a temperature of from 30 °C to 40 °C. In one embodiment the incubating is at a temperature of about 37 °C.

In one embodiment the second polypeptide has at its N-terminus the amino acid sequence GGG, AAA, CGG, CAA, KGG or KAA.

In one embodiment the first polypeptide comprises at its C-terminus the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue). In one embodiment the first polypeptide comprises at its C-terminus the amino acid sequence LPETG (SEQ ID NO: 04) or LPETA (SEQ ID NO: 42).

In one embodiment the first polypeptide and the second polypeptide are independently of each other selected from an antibody variable domain, an

antibody heavy chain Fab-fragment, an antibody Fc-region, a tag, a peptide comprising the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue), a linker, and a non-sortase motif moiety.

One aspect as reported herein is the use of a deep eutectic solvent comprising choline chloride and glycerol at a molar ratio of 1:2, and further comprising up to 5 % (v/v) aqueous co-solvent as solvent in an enzymatic transamidation reaction catalyzed by a sortase, preferably by a Sortase A.

Detailed Description of the Invention

The invention is at least in part based on the finding that deep eutectic solvents can be used to conduct transamidation reactions, especially sortase catalyzed reactions.

The invention is at least in part based on the finding that hardly water soluble substrates can be conjugated by sortase (in low water content) deep eutectic solvents.

I. DEFINITIONS

The term "derived from" denotes that the respective amino acid sequence comprises the same amino acid sequence, or contains 1 to 10 amino acid sequence changes, or is a shortened variant or a fused variant of a parent amino acid sequence.

The term "comprising" when used herein expressly includes the term "consisting of".

The term "a glycinyl, an alaninyl, or a cysteinyl compound" denotes a compound that comprises a glycine or an alanine or a cysteine amino acid residue with free alpha amino group, e.g. as NH_2 or NH_3^+, and a carboxy group that is in/part of a peptide bond with another moiety at position 1, whereby the moiety can be any amino group containing moiety, such as an isolated amino acid residue, a peptide, a polypeptide, a protein, a small molecule, a dye, or a (synthetic or natural) linker.

In the present specification and claims the numbering of the residues in an immunoglobulin heavy chain Fc-region is that of the EU index of Kabat (Kabat, E.A., et al., Sequences of Proteins of Immunological Interest, 5th ed., Public Health

Service, National Institutes of Health, Bethesda, MD (1991), NIH Publication 91-3242, expressly incorporated herein by reference).

The term "alteration" denotes the mutation, addition, or deletion of one or more amino acid residues in a parent amino acid sequence, e.g. of an antibody or fusion polypeptide comprising at least an FcRn binding portion of an Fc-region, to obtain a variant antibody or fusion polypeptide.

The term “amino acid mutation” denotes a modification in the amino acid sequence of a parent amino acid sequence. Exemplary modifications include amino acid substitutions, insertions, and/or deletions. In one embodiment the amino acid mutation is a substitution. The term “amino acid mutations at the position” denotes the substitution or deletion of the specified residue, or the insertion of at least one amino acid residue adjacent the specified residue. The term “insertion adjacent to a specified residue” denotes the insertion within one to two residues thereof. The insertion may be N-terminal or C-terminal to the specified residue.

The term “amino acid substitution” denotes the replacement of at least one amino acid residue in a predetermined parent amino acid sequence with a different “replacement” amino acid residue. The replacement residue or residues may be a “naturally occurring amino acid residue” (i.e. encoded by the genetic code) and selected from the group consisting of: alanine (Ala); arginine (Arg); asparagine (Asn); aspartic acid (Asp); cysteine (Cys); glutamine (Gln); glutamic acid (Glu); glycine (Gly); histidine (His); isoleucine (Ile): leucine (Leu); lysine (Lys); methionine (Met); phenylalanine (Phe); proline (Pro); serine (Ser); threonine (Thr); tryptophan (Trp); tyrosine (Tyr); and valine (Val). In one embodiment the replacement residue is not cysteine. Substitution with one or more non-naturally occurring amino acid residues is also encompassed by the definition of an amino acid substitution herein. A “non-naturally occurring amino acid residue” denotes a residue, other than those naturally occurring amino acid residues listed above, which is able to covalently bind adjacent amino acid residues(s) in a polypeptide chain. Examples of non-naturally occurring amino acid residues include norleucine, ornithine, norvaline, homoserine, alpha-amino isobutyric acid and other amino acid residue analogues such as those described in Ellman, et al., Meth. Enzym. 202 (1991) 301-336. To generate such non-naturally occurring amino acid residues, the procedures of Noren, et al. (Science 244 (1989) 182) and/or Ellman, et al. (supra) can be used. Briefly, these procedures involve chemically activating a suppressor tRNA with a non-naturally occurring amino acid residue followed by in vitro

transcription and translation of the RNA. Non-naturally occurring amino acids can also be incorporated into peptides via chemical peptide synthesis and subsequent fusion of these peptides with recombinantly produced polypeptides, such as antibodies or antibody fragments.

The term “amino acid insertion” denotes the incorporation of at least one additional amino acid residue into a predetermined parent amino acid sequence. While the insertion will usually consist of the insertion of one or two amino acid residues, the present application contemplates larger “peptide insertions”, e.g. insertion of about three to about five or even up to about ten amino acid residues. The inserted residue(s) may be naturally occurring or non-naturally occurring as defined above.

The term "amino acid deletion" denotes the removal of at least one amino acid residue at a predetermined position in an amino acid sequence.

Within this application whenever an amino acid alteration is mentioned it is a deliberated amino acid alteration and not a random amino acid modification.

The term “tag” denotes a sequence of amino acid residues connected to each other via peptide bonds that has specific binding properties. In one embodiment the tag is an affinity or purification tag. In one embodiment the tag is selected from Arg-tag, His-tag, Flag-tag, 3xFlag-tag, Strep-tag, HA-tag, Nano-tag, SBP-tag, c-myc-tag, S-tag, SNUT-Tag, NusA, T7, thioredoxin, calmodulin-binding-peptide, cellulose-binding-domain, chitin-binding-domain, GST-tag, or MBP-tag (see, e.g., Amau, J., et al., Prot. Expr. Purif. 48 (2006) 1-13).

In one embodiment the tag is selected from SEQ ID NO: 07 (RRRRR), or SEQ ID NO: 08 (RRRRRR), or SEQ ID NO: 09 (HHHHHH), or SEQ ID NO: 10 (KDHLIHNVHKEFHAHAHNK), or SEQ ID NO: 11 (DYKDDDDK), or SEQ ID NO: 12 (DYKDHDGDYKDHDIDYKDDDDK), or SEQ ID NO: 13 (AWRHPQFGG), or SEQ ID NO: 14 (WSHPQFEK), or SEQ ID NO: 15 (MDVEAWLGAR), or SEQ ID NO: 16 (MDVEAWLGARVPLVET), or SEQ ID NO: 17 (MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP), or SEQ ID NO: 18 (EQKLISEEDL), or SEQ ID NO: 19 (KETAAAKFERQHMDS), or SEQ ID NO: 20 (KRRWKKNFIAVSAANRFKKISSSGAL), or SEQ ID NO: 21 (cellulose binding domain), or SEQ ID NO: 22 (cellulose binding domain), or SEQ ID NO: 23 (TNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEP SNVPALWQLQ), or SEQ ID NO: 24 (GST-tag), or SEQ ID NO: 25 (MBP-tag), or SEQ ID NO: 32 (MRGSHHHHHHGS).

An "effective amount" of an agent, e.g., a pharmaceutical formulation, refers to an amount effective, at dosages and for periods of time necessary, to achieve the desired therapeutic or prophylactic result.

The term “individual” or “subject” denotes a mammal. Mammals include, but are not limited to, domesticated animals (e.g. cows, sheep, cats, dogs, and horses), primates (e.g., humans and non-human primates such as monkeys), rabbits, and rodents (e.g., mice, rats, and hamsters). In certain embodiments, the individual or subject is a human.

The term "pharmaceutical formulation" refers to a preparation which is in such a form as to permit the biological activity of an active ingredient contained therein to be effective, and which contains no additional components which are unacceptably toxic to a subject to which the formulation would be administered.

A “pharmaceutically acceptable carrier” refers to an ingredient in a pharmaceutical formulation, other than an active ingredient, which is nontoxic to a subject. A pharmaceutically acceptable carrier includes, but is not limited to, a buffer, excipient, stabilizer, or preservative.

The term “position” denotes the location of an amino acid residue in the amino acid sequence of a polypeptide. Positions may be numbered sequentially, or according to an established format, for example the EU index of Kabat for antibody numbering.

As used herein, “treatment” (and grammatical variations thereof such as “treat” or “treating”) refers to clinical intervention in an attempt to alter the natural course of the individual being treated, and can be performed either for prophylaxis or during the course of clinical pathology. Desirable effects of treatment include, but are not limited to, preventing occurrence or recurrence of disease, alleviation of symptoms, diminishment of any direct or indirect pathological consequences of the disease, preventing metastasis, decreasing the rate of disease progression, amelioration or palliation of the disease state, and remission or improved prognosis. In some embodiments, antibodies of the invention are used to delay development of a disease or to slow the progression of a disease.

The term “sortase A enzymatic activity” as used herein denotes a polypeptide that shows transpeptidation activity in a Reporter immobilization assay (i.e. an assay

according to Example 6; and as reported in WO 2016/096740 (incorporated herein by reference).

II. DEEP EUTECTIC SOLVENTS

Deep eutectic solvents (DESs) are mixtures comprising polar components (liquids or solids) that when combined have decreased melting temperatures. DESs can be produced by mixing choline chloride (ChCl) with an organic hydrogen donor, e.g. an amine, amide, alcohol, or carboxylic acid (see e.g. Abbott, A.P., et al., J. Am. Chem. Soc. 126 (2004) 9142–9147; ibid. Chem. Commun. (2003) 70–71). The DES comprises anions, cations and neutral components so that a DES has both, ionic and organic solvent properties (Zhang et al., Angew. Chem. Int. Ed. 48 (2009) 3486–3490).

Proteases have shown activity in DES (Zhao, H., et al., J. Mol. Catal. B-Enzym 72 (2011) 163-167). Likewise lipases have shown activity in DES at 50 °C with 1 % (v/v) water (Zhao, H., et al., J. Mol. Catal. B-Enzymatic 85-86 (2013) 243-247) or in a solution comprising 0.135 mL 2-propanol, 1.215 mL Tris–HCl buffer (50 mmol L-1, pH 8.0) and 0–1.2mol L-1 of DES at 37 °C or combining 0.4-0.5 ml oil, 2.1 mL DES, 0.9 mL methanol, and 300 mg enzyme powder without addition of extra water at 50 °C (Huang, Z.-L., et al., J. Chem. Technol. Biotechnol. 89 (2014) 1975-1981) or at 40 °C to 60 °C (US 8,247,198).

Generally DES can be made by simply warming and stirring the components for a certain time, optionally with a successive drying step. For example, DES can be prepared by a thermal treatment of the previously mixed components (see Abbott (2003) supra; Abbott, A.P., et al., ChemPhysChem 7 (2006) 803) or by a freeze-drying procedure (Gutierrez, M.C., et al., Langmuir 25 (2009) 5509; ibid Angew. Chem. Int. Ed. 49 (2010) 2158).

For example, choline chloride is mixed with an H-bond donor (e.g., glycerol) at elevated temperature (e.g. at 80 °C) in the required molar ratio with agitation for half an hour. To remove traces of water drying with phosphor pentoxide can be performed afterwards (e.g. for two weeks at 45 °C).

Also for example, choline chloride and the hydrogen-bond donor (glycerol or ethylene glycol) are mixed at the intended ratio (i.e. 1:2 or 1:3, respectively) and mixed, using e.g. a magnetic stirrer, at 80 °C for 1–2 hours. After a colorless clear

liquid had formed the mixture is cooled down to room temperature and dried over phosphor pentoxide in a desiccator (e.g. at room temperature, 2 weeks)

Another exemplary method is reported in US 8,247,198: nitrogen salt (0.05 mol) and hydrogen bond donor (0.1 mol for choline chloride mixtures, 0.075 mol for ethyl ammonium chloride mixtures) were added to a 20-ml vial and heated at 80 °C, until a clear, homogenous liquid formed, typically one hour.

According to US 8,247,198 the components within the DES's are surprisingly and significantly 20 to more than 600-fold less reactive for enzymes than expected. It is reported therein a method comprising enzymatic catalysis of a chemical reaction in a solution comprising a deep eutectic solvent. The reaction may be, e.g., transesterification, aminolysis, hydrolysis, perhydrolysis, and/or alcohol dehydrogenase activity. The reaction may be a polymerization reaction. The polymerization reaction may be catalyzed by an enzyme that is, e.g., a member of the group consisting of enzymes that catalyze transesterification, aminolysis, hydrolysis, perhydrolysis, alcohol dehydrogenation, oxidation-reduction, or dehydrogenation. The reaction may produce, e.g., an addition product or a condensation product. The polymerization reaction may produce, e.g., a polyester or a polyamide. An enzyme that catalyzes the enzyme-catalyzed chemical reaction may be, e.g., a member of the group consisting of transesterase, hydrolase, lipase, amidase, and dehydrogenase (column 2, lines 4 to 19).

Deep eutectic solvents can comprise as first component a strong hydrogen bond donor and as second component a metal salt or a nitrogen salt. Organic or aqueous components as co-solvents can be present too. For example, in one embodiment, the deep eutectic solvent comprises between 10% and 99% (v/v) of deep eutectic solvent and the remainder is a non-deep eutectic solvent. In one embodiment the second component is a halide-containing salt of an amine or metals. In one embodiment the metal is a transition metal. In one embodiment the first component is selected from the group consisting of hydroxyl, amide, amine, aldehyde, and carboxylic acid. In one embodiment the first component is selected from the group consisting of organic acids, urea, thiourea, amide, hydroxyl groups, diols, glycerol, choline chloride, and combinations thereof. In one embodiment the first component is selected from the group comprising choline chloride, ethyl ammonium chloride, choline bromide, glycerol, tetra butyl ammonium chloride, triethyl benzyl ammonium chloride, zinc chloride, and acetylcholine chloride. In one embodiment the second component is selected from the group consisting of acetamide ethylene

glycol, glycerol, urea, malonic acid, formamide, arabinose, glucose, and xylose. In one preferred embodiment the deep eutectic solvent comprises as first component choline chloride and as second component glycerol.

The terms "deep eutectic solvent" or "DES", which can be used interchangeably herein, denote mixtures comprising a nitrogen or metal salt and a hydrogen bond donor forming a eutectic point at the specific applied ratio.

In one embodiment the nitrogen salt is selected from compounds with a positively charged nitrogen atom. In one embodiment the nitrogen salt is selected from the group consisting of primary, secondary, tertiary, and quaternary nitrogen components. In one embodiment the nitrogen salt is an ammonium salt or a substituted ammonium compound. In one embodiment the nitrogen salt is a halide-nitrogen salt. In one embodiment the metal salt is a transition metal salt. In one embodiment the metal salt is a halide-metal salt. In one embodiment the hydrogen bond donor is selected from the group consisting of hydroxyl, amide, amine, aldehyde, and carboxylic acid. In one embodiment the hydrogen bond donor is selected from the group consisting of hydrogen halides, organic acids, urea, choline chloride, thiourea, glycerol, diols, and propane diols. Transition metals are the elements from the period table of the chemical elements that have a number of 21 to 30, 39 to 48, 71 to 80, and 103 to 112. Halides are fluoride (F-), chloride (Cl-), bromide (Br-), iodide (I-) and astatide (At-). In one embodiment the components of the DES have a specific ratio based on their used volume. In one embodiment the ratio is from about 1:3 parts to about 3:1 parts of nitrogen or metal salt to strong hydrogen bond donor. When used in an enzymatic reaction also other components will be present in the DES, such as e.g. an enzyme and its substrates. If the enzyme is not used a solid also the buffer in which the enzyme is dissolved will be present in the DES of a reaction mixture. In one embodiment the reaction mixture of the enzymatic reaction comprises at least 95 % DES, at least 90 % DES, at least 85 % DES, at least 80 % DES.

Due to the fact that some enzyme are not available in solid form (either conjugated to a solid support or as powder) a DES used in an enzymatic reaction will comprises to a certain percentage a co-solvent, in most cases water or an aqueous buffer. Thus, the DES accounts for at least 95 %, at least 90 %, at least 85 %, or at least 80 % of the volume of the enzymatic reaction mixture and the rest is made up of water or an aqueous buffer. In one preferred embodiment the enzymatic reaction mixture comprises at least 90 % of DES, more preferably about 95 % of DES.

Deep eutectic solvents useful for enzymatic reactions are, e.g., the mixtures of choline chloride and acetamide (ChCl:Acet), choline chloride and ethylene glycol (ChCl:EG), choline chloride and glycerol (ChCl:G), choline chloride and urea (ChCl:U), choline chloride and malonic acid (ChCl:MA), ethyl ammonium chloride and acetamide (EAC:Acet), ethyl ammonium chloride and ethylene glycol (EAC:EG), ethyl ammonium chloride and glycerol (EAC:Gly), choline bromide and glycerol (ChBr:Gly), tetra butyl ammonium chloride and glycerol (Bu4NCl:Gly), triethyl benzyl ammonium chloride and glycerol (Et3BzNCl:Gly), acetylcholine chloride and glycerol (AcChCl:Gly), choline chloride and formamide (ChCl:F), choline chloride and arabinose (ChCl:Ara), choline chloride and glucose (ChCl:Glc) and choline chloride and xylose (ChCl:Xyl).

Many reactions can be performed in DES, e.g. comprising an enzyme and a DES and optionally a co-solvent. The enzyme and its substrates are present in effective concentrations and at ratios useful to make the intended product. An effective concentration does not include concentrations that are not suitable to make the intended product in meaningful quantities and in a reasonable time. In one embodiment the enzyme is used/is present in an amount/concentration of at least about 0.1, at least about 1, at least about 5, at least about 10, or at least about 20 mg enzyme per ml of solvent, i.e. from about 0.1 to about 10 mg/ml. Similarly, the enzymatic substrates are in one embodiment present in a concentration of at least about 1%, at least about 5%, at least about 10%, at least about 20%, or at least about 30% wt substrate/wt of solvent, i.e. from about 1% to about 25%.

III. ENZYMATIC CONJUGATION USING SORTASE A

A covalent conjugate (i.e. a fusion polypeptide) comprising two in vivo not covalently associated entities can be obtained in vitro by using the enzyme sortase, especially a Sortase A.

Transamidases in general catalyze the formation of a peptide bond (amide bond) between an acyl donor and a nucleophilic acyl acceptor. In order to form a peptide bond the acyl acceptor has to contain a NH2-CH2-moiety. Gram-positive bacteria include the following genera: Actinomyces, Bacillus, Bifidobacterium, Cellulomonas, Clostridium, Corynebacterium, Micrococcus, Mycobacterium, Nocardia, Staphylococcus, Streptococcus and Streptomyces.

Sortases have been classified into 4 classes, designated A, B, C, and D, based on sequence alignment and phylogenetic analysis of 61 sortases from gram-positive

bacterial genomes (Dramsi, S., et al., Res. Microbiol. 156 (2005) 289-297). These classes correspond to the following subfamilies, into which sortases have also been classified by Comfort and Clubb (Comfort, D. and Clubb, R.T., Infect. Immun. 72 (2004) 2710-2722): Class A (Subfamily 1), Class B (Subfamily 2), Class C (Subfamily 3), Class D (Subfamilies 4 and 5). The aforementioned references disclose numerous sortases and recognition motifs (see also Pallen, M. J., et al., Trends Microbiol. 9 (2001) 97-101). With this information a person skilled in the art can assign a sortase to the correct class based on its sequence and/or other characteristics such as those described in Dramsi (supra).

Sortase A (SrtA) is a membrane bound enzyme has transamidase activity. It has been identified and isolated from gram-positive bacteria. In vivo Sortase A attaches proteins covalently to the bacterial cell wall. The specific recognition motif on the SrtA substrate is LPXTG, whereby the enzyme cleaves between the residues threonine and glycine. The recognition motif on the peptidoglycan is a pentaglycine motif. It has been shown that a triglycine and even a diglycine motif on the N-terminus is sufficient to support the SrtA reaction (Clancy, K.W., et al., Peptide Science 94 (2010) 385-396). The reaction proceeds through a thioester acyl-enzyme intermediate, which is resolved by the attack of an amine nucleophile from the oligoglycine, covalently linking peptidoglycan to a protein substrate and regenerating SrtA. SrtA can be used to covalently conjugate chemically synthetized peptides to recombinantly expressed proteins.

Many gram-positive bacteria use sortase to covalently anchor a variety of surface proteins including virulence factors to their cell wall (peptidoglycan). Sortases are membrane associated enzymes. The wild-type Staphylococcus aureus sortase A (SrtA) is a polypeptide of 206 amino acids with an N-terminal membrane-spanning region. In a first step, sortase A recognizes substrate proteins that contain a LPXTG (SEQ ID NO: 01) amino acid sequence motif and cleaves the amide bond between the Thr and Gly by means of an active-site Cys. This peptide cleaving reaction results in a sortase A-substrate thioester intermediate. In a second step the thioester acyl-enzyme intermediate is resolved by nucleophilic attack of an amino group of an oligoglycine containing second substrate polypeptide (corresponding to the pentaglycine unit of peptidoglycan in S. aureus) leading to a covalently linked cell wall protein and the regeneration of sortase A. In the absence of oligoglycine nucleophiles, the acyl-enzyme intermediate can be hydrolyzed by a water molecule.

Sortase-mediated ligation/conjugation has begun to be applied for a variety of protein engineering and bioconjugation purposes. This technique enables the introduction of natural and synthetic functionalities into LPXTG-tagged recombinant or chemically synthesized polypeptides. Examples include the covalent attachment of oligoglycine derivatized polymers (e.g. PEG), fluorophores, vitamins (e.g. biotin and folate), lipids, carbohydrates, nucleic acids, synthetic peptides and proteins (e.g. GFP) (see e.g. Tsukiji, S. and Nagamune, T., ChemBioChem 10 (2009) 787-798; Popp, M.W.L. and Ploegh, H.L., Angew. Chem. Int. Ed. Engl. 50 (2011) 5024-5032).

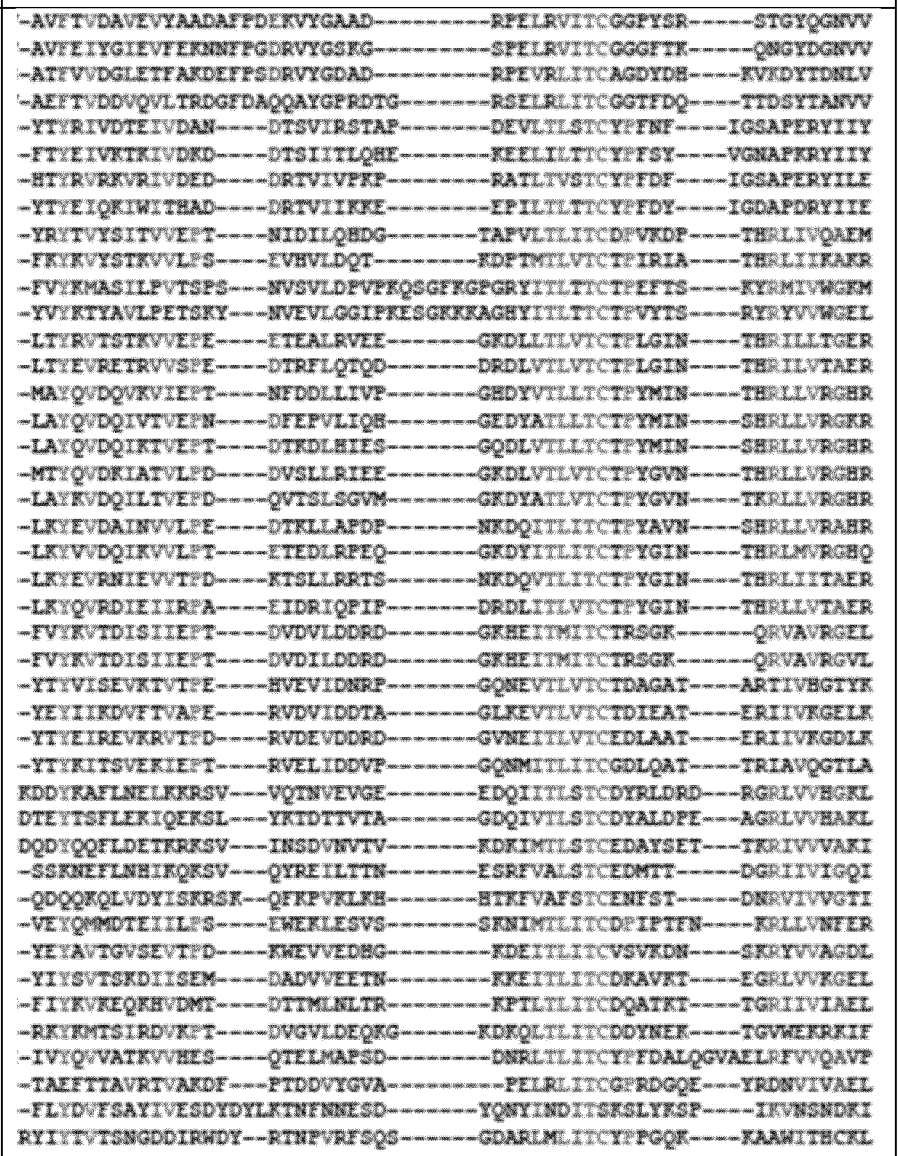
year / citation	**content**
1990 Fischetti et al. Mol. Microbiol. 4 (1990) 1603-1605	- LPETG sortase motif
1999 Mazmanian et al. Science 285 (1999) 760-763	- Staphylococcus aureus strain OS2 - conserved Leu-Pro-X-Thr-Gly (LPXTG) motif - srtA gene specifies a protein of 206 amino acids with a potential NH2-terminal signal peptide/membrane anchor sequence and a presumed active-site cysteine at position 184 - srtA homologs are present in Actinomyces naeslundii, Bacillus subtilis, Enterococcus faecalis, Staphylococcus aureus, Streptococcus mutans, Streptococcus pneumoniae, and Streptococcus pyogenes - AF162687 discloses coding sequence (frame 1 reading) MKKWTNRLMT IAGVVLILVA AYLFAKPHID NYLHDKDKDE KIEQYDKNVK EQASKDKKQQ AKPQIPKDKS KVAGYIEIPD ADIKEPVYPG PATPEQLNRG VSFAEENESL DDQNISIAGH TFIDRPNYQF TNLKAAKKGS MVYFKVGNET RKYKMTSIRD VKPTDVGVLD EQKGKDKQLT LITCDDYNEK TGVWEKRKIF VATEVK
1999 Ton-That et al. Proc. Natl. Acad. Sci. USA 96 (1999) 12424-12429	- Staphylococcus aureus strain OS2 - Staphylococcus aureus Sortase A - residues 2-25 deleted (N-terminal signal sequence) - MW: 22139 Da - mutation C184S abolishes catalytic activity - sortase homologs from Streptococcus pyogenes (Spyo), Enterococcus faecalis (Efea), Actinomyces naeslundii (Anei), Streptococcus mutans (Smut), Bacillus subtilis (Bsub), and Streptococcus pneumoniae (SpnA, SpnB, and SpnC) - water can resolve acyl-enzyme intermediate
2000 Dhar et al. Biochem. 39	- Listeria monocytogenes has a peptidoglycan cross-bridge (m-Dpm) which is chemically distinct and much shorter than the pentaglycine cross-bridge of staphylococci

year / citation	content
(2000) 3725-3733	- Listeria monocytogenes Sortase A has the same sortase motif as Staphylococcus aureus Sortase A: LPXTG (LPTTG)
2000 Ton-That et al. J. Biol. Chem. 275 (2000) 9876-9881	- purified recombinant Staphylococcus aureus Sortase A - hydrolyzes peptides with LPXTG motif - triple G-motif - in presence of H2N-GGG exclusive transpeptidation - in the presence of amino-donors sortase mediated LPXTG motif cleavage rate was increased - Sortase is a 206-amino acid polypeptide with an N-terminal signal sequence/stop transfer domain, is anchored in the cytoplasmic membrane of staphylococci - residues 1-25 correspond to N-terminal signal sequence - reaction conditions: - 5 mM amino group nucleophile - 4.71 µM SrtADN - 150 mM NaCl, 5 mM CaCl2, 50 mM Tris-HCl, pH 7.5) - volume of 520 µl - reaction conditions: - 10 µM fluorescent peptide - 5 mM amino group nucleophile H2NGGG - 15 µM SrtADN - 150 mM NaCl, 5 mM CaCl2, 50 mM Tris-HCl, pH 7.5) - volume of 520 µl - 37 °C, 16 h

year / citation	content
	TABLE III *Kinetic analysis of $SrtA_{\Delta N}$* Kinetic constants K_m, V_{max}, and k_{cat} were calculated from the curve fit for the Michaelis-Menten equation using the Lineweaver-Burk plot. Reaction conditions are described in the legend to Fig. 4. Nucleophile / K_m / V_{max} / K_{cat} / K_m/K_{cat} μM / $\mu M/s$ / l/s / $l/\mu M \cdot s$ H_2O / 10.88 / 5.08×10^{-5} / 1.06×10^{-5} / 9.77×10^{-7} NH_2-Gly_3 / 16.48 / 1.08×10^{-4} / 2.27×10^{-5} / 1.38×10^{-6} TABLE IV *The effect of different nucleophiles on the rate of LPXTG peptide cleavage by sortase ($SrtA_{\Delta N}$)* Nucleophile / rate (s^{-1})[a] H_2O / 1.84 (±0.11) NH_2OH / 1.91 (±0.07) NH_2-Gly / 1.95 (±0.05) NH_2-Gly_2 / 2.03 (±0.13) NH_2-Gly_3 / 2.91 (±0.03) [a] Slope of the kinetic curves as shown in Fig. 4. The substrate peptide *d*-QALPETGEE-e was incubated with $SrtA_{\Delta N}$ and various nucleophiles. Substrate cleavage between the threonine and the glycine was measured as an increase in fluorescence. With the exception of water, all nucleophiles were added at a concentration of 5 mM. Averages were calculated from three independent experiments and standard deviations are reported (parentheses).
2001 Bolken et al. Infect. Immun. 69 (2001) 75-80	- sortase A from Streptococcus gordonii - 252 amino acid residues with N-terminal signal sequence - cysteine at position 210 - 12-amino-acid extension at the carboxy-terminus of the S. gordonii protein compared to S. aureus
2001 Glaser et al. Science 294 (2001) 849-852	- Listeria monocytogenes genome contains 41 proteins containing an LPXTG - Lm genome contains more LPXTG proteins than any other gram-positive bacterium (Strep. pyogenes: 13; St.au. aureus: 18)
2001 Ilangovan et al. Proc. Natl. Acad. Sci. USA 98 (2001) 6056-6061	- NMR structure of St.au. aureus Sortase A - unique b-barrel structure including two short helices and several loops - the active-site sulfhydryl of cysteine-184 is poised for ionization by histidine-120, presumably enabling the resultant thiolate to attack the LPXTG peptide - conservation of H120 and C184 - calcium binding near the active site stimulates catalysis - stimulation by about 2 mM calcium ions - magnesium and manganese ions can substitute for calcium ions - sortase with residues 1-29 deleted - sortase with residues 1-59 deleted - average mass of 16,595.12 Da observed

year / citation	content
	- reaction conditions: - 2 mM H2N-GGG - 5 µM sortase N-terminal deletion variant - 150 mM NaCl, 5 mM CaCl2, 50 mM Tris*HCl pH 7.5 - reaction volume 520 µL - S. aureus SrtA residues 26–59 display no amino acid conservation; core SrtA residues 60–206 present in all sortase homologs examined
2001 Mazmanian et al. Mol. Microbiol. 40 (2001) 1049-1057	- St. aureus sorting motifs: LPETG, LPDTG, LPKTG, LPNTG, PLAAG, LPKAG, LPQTG - Actinomyces naeslundii, Bacillus anthracis, Bacillus subtilis, Clostridium acetabutylicum, Corynebacterium diphtheria, Enterococcus faecalis, Listeria monocytogenes, Streptococcus mutans, Streptococcus pneumoniae, Streptococcus pyogenes
2001 Pallen et al. TRENDS Microbiol. 9 (2001) 97-101	- LPXTG like motif, followed shortly by a membrane-spanning hydrophobic domain and a charged carboxy-terminal tail

year / citation	content
	SC7A8.19 VAGHVDNA--------EGPAVFYRLGALEKGSAIEIDRRDGGV- SCM10.23 VVGHVDNQ--------QGPAVFYGLGALKKGNKVEVHRQDGKT- SC5C11.07 IAGHVDTK--------TSAAVFARLDQLDKGDKFQVRRADGRS- SCE20.15c MVGHVDTE--------TRPAVFYQLSTLEPGQTIRVARDDDEV- BH3596_bachd LSGHRDT----------VFRDMGKLEIGDDLTVHMPYGS-- yhcs_bacsu LSGHRDT----------VFRRTGELEKGDQLRLLLSYGE-- BH4010_bachd LSGHRDT----------VFRELGEVGVGDLLIVETATGT-- slp2_bacan LSGHRDT----------VFTDLGQLKEKDTLVLEYDNKT-- BH2127_bachd IAAHRSR---------TYGRQFNRLDEVEVGDVITVTTNNHM-- slp_cloab LAGHRSY---------TFGEYFNRLGEIGSGDEIDVETVNGT-- SCH69.20c LAGHRN----------THGEPFRYINKLEPGDPIVVETQDKY-- SCH69.19c LAAHRD----------GHGARFHNIDKIEKGDPIVFETKDTW-- fap2_actna ITGHRGL---------AEATMFTNLDKVKTGDSLIVEVFGEV-- slp4_cordi ITAHRGL---------AEATMFTNLNKVGVGDRFTIEVMGEV-- slp2_strpn ITAHTGL---------PTAEMFTDLTKLKVGDKFYVHNIKEV-- slp4_strpn ITAHRGL---------PTAELFSQLDKMKKGDIFYLHVLDQV-- slp2_enfae ISGHRGL---------PQAKLFTDLPELKKGDEFYIEVNGKT-- slp_streq ISGHRGL---------PSAKLFTNIDKLRINDTFTITSLNRT-- slp2_strpy ISAHRGL---------PSAEMFTNLMLVKKGDTFYFRVLNKV-- slp1_cordi ITGHSGL---------ANATLFDNLEDVKEHDPIYITVQGET-- slp2_cordi ITGHTGL---------ANSTMFDHLNKAEKGDTFYVQVSGEK-- slp3_cordi LSAHTGL---------QNATLWDNLIQIKKGDPVYVAAAGEK-- slp5_cordi LTAHSGI---------QKSTFFDNLEKVKKGDAIYVRNIGET-- BH2015_bachd IAGHRGYRGNR---------HFSRLPDVTIGDEVFLHTKEET-- BH0362_bachd IAGHRGYRGNR---------HFSRLPDVTIGDEVFLHTKEET-- slp_strmu LASHHVFGMTG-----SSQMLFSPLERAKEGMEIYLTDKNKV-- slp1_strpy LASHHIFGITG-----SSQMLFSPLERAQNGMSIYLTDKEKI-- slp1_strpn LASHHIFGVDN-----ANKMLFSPLDNAKNGMKIYLTDKNKV-- slp1_enfae LASHRTEDGVS---------LFSPLERTKKDELIYITDLSTV-- BH3294_bachd VDHHEGFYYDT-LYNRYDVEVFSAYVTTT--DFYYIETEFPS-I slp3_bacan FMSHRKLYYDT-LFEGYDLEVFSVYTTTT--DFYYIETDFSS-I slp_staau YEKHKIIEFDN-KYGKYQLQVFSAYKTTT--KDNYIRTDFEN-I slp3_strpy FNKHKEFSIETKTKQKLKINIFACIQTDAFDSLLFNPIDVDI-- slp4_strpy FNKHNKAIIETKERKKLTVTIFACLKTDAFDQLVFNPNAITN-- slp3_strpn IAGHRAE---------PSHVFFRHLDQLKVGDALYYDNGQEI-- slp1_bacan LAGHNMS---------KKGVLFSDIASLKKGDKIYLYDNENE-- ywpe_bacsu LAGHHLK---------QKNLLFGPLENIKTGAQIVITDFKKD-- slp3_enfae LASHNAG---------YEGLLFTSLNKVSVGDLVKLNDREGHS- sortase_staau IAGHTFID--------RFNYQFTNLKAAKKGSMVYFKVGNET-- slp_shepu IAGHRDT----------HFAILKGMTVGRRLALQTAAGKE- SCH22A.15c VLGHVTVG--------RYDGVFRHLAGRR-GERIEARENGT--- slp_clodi IYGHNMKN-----KTMFNNLNKFKDADFFKKNNKIKITLNGKE-- MTH1829_metth ILGHRTT---------YSGPFRKIGALRKGDRVIIEDASSSII

year / citation	content
	 St.au. aureus: LPKTG St.au. pyogenes: LPITG Bac. anthracis: LPKTG Bac. subtilis: LPDTA Clos. difficile: SPKTG Clos. acetabutylicum: LPKTG S. coelicolor: LAETG, LAATG, LAHTG, LASTG Arthrobacter. sp.: LASTG A. naeslundii: LPLTG A. viscosus: LPLTG, LSRTG S. pneumoniae: LPETG, VPDTG, IPQTG, YPRTG C. diphtheria: LPMTG, LALTG, LPKTG, LGNTG, LPLTG, LAFTG S. putrefasciens: LPQTS
2002	- in silico identification of sortase from L. monocytogenes

year / citation	content
Garandeau et al. Infect. Immun. 70 (2002) 1382-1390	- 222 amino acid residues - TLXTC consensus motif for sortase active site MLKKTIAAAA LAAGLLLIFS PFIKNGIVKY MSGHETIEQY KASDIKKNNE KDATFDFESV QLPSMTSVIK GAANYDKDAV VGSIAVPSVD VNLLVFKGTN TANLLAGATT MRSDQVMGKG NYPLAGHHMR DESMLFGPIM KVKKGDKIYL TDLENLYEYT VTETKTIDET EVSVIDNTKD ARITLITCDK PTETTKRFVA VGELEKTEKL TKELENKYFP SK
2002 Bierne et al. Mol. Microbiol. 43 (2002) 869-881	- Listeria monocytogenes - BLAST analysis: one sequence that is 28% identical to S. aureus SrtA and encodes a protein of 222 amino acids with an expected molecular weight of 24.7 kDa - contains a putative signal peptide/transmembrane region, the expected TLXTC sequence and two stretches of 13 and 31 amino acids that are not present in SrtA
2004 Mao et al. J. Am. Chem. Soc. 126 (2004) 2670-2671	- peptide with single aminoglycoside as nucleophile - attack 50 times faster than water - yield not affected by number of glycine residues - 30 % yield after 30 min, 50 % after 6 h, 90 % after 24 h - conjugates D-amino acid containing peptide; rate half that for L-amino acid peptides - conjugation of small molecules coupled to triglycine (folate) - branched peptide as efficiently coupled as linear peptide
2004 Kruger et al. Biochem. 43 (2004) 1541-1551	- SrtA in Staphylococcus aureus, Streptococcus gordonii, Listeria monocytogenes, Streptomyces coelicolor, Streptococcus pyogenes, Streptococcus suis - S. aureus two sortase isoforms: SrtA, SrtB (MPQTN motif) - LPXTG motif highly conserved among all gram-positive bacteria - NPQTN motif appears to be conserved only among at least three bacteria containing the heme iron acquisition isd gene locus (Bacillus anthracis, Bacillus halodurans, and S. aureus) - Staphylococcus aureus Sortase A - residues 2-24 deleted - transpeptidase reaction conditions: - 100 µL reaction volume - 150 mM NaCl, 300 mM Tris, 5 mM CaCl2 (pH 7.5), pentaglycine (2 mM), SrtAΔN24 (840 nM), and 0 to 10 mM Abz-LPETG-Dap(Dnp)-NH2 - 37 °C, 30 min - transpeptidase reaction conditions: - 100 µL reaction volume - 150 mM NaCl, 300 mM Tris, 5 mM CaCl2 (pH 7.5) - SrtAΔN24 (15 µM), Gly5 (2 mM), peptide (300 µM) - 37 °C for 30 min - quenched by 1 N HCl (50 µL) - transpeptidase reaction conditions: - 100 µL reaction volume

year / citation	content
	- 150 mM NaCl, 300 mM Tris, 5 mM CaCl2 (pH 7.5) - SrtAΔN24 (60 μM), Gly5 (2 mM), peptide (300 μM) - 37 °C for 360 min - no reaction with inverted sortase motif or SrtB motif - initial velocity motif: LPXTG (X=any except P,C; X=M fastest) - end point motif: L/M-P-X-A/L/S/T/V-G (X=any except P,C; L better than M; T and A comparable, then S and V and L comparable) - Table 3 sorting signals: LM: IPKTG, IPALG, LAASS, LPATG, LPKAG, LPISS, IPALG, LPKTS
2007 Parthasarathy et al. Bioconjug. Chem. 18 (2007) 469-476	- conjugation to surface of non-protein species (polystyrene beads, PEG) - alkylamine as nucleophile - making of cyclic peptides - use of 159 amino acid sortase - transpeptidase reaction conditions: - 50 mM Tris, 150 mM NaCl, pH 8, 0.1% Tween-20, 6 mM CaCl2, 3 mM beta-mercaptoethanol - 37°C, 3 h - 15 μM eGFP, 10 μM sortase - add 10 mM EDTA to stop the reaction - transpeptidase reaction conditions: - 50 mM Tris, 150 mM NaCl, pH 8 - 37°C, 3 h - 12 μM eGFP-LPETG, 40 μM sortase, 36 μM GGG-eGFP - with beads higher concentration of sortase has to be used for comparable yield - use of a cleavable upstream tag to protect nucleophile (GGG) - enzyme displays a strong preference for glycine in the first position; alanine and valine can apparently substitute for glycine in the second position, although the reaction is not as efficient
2007 Chan et al. PlosOne 11 (2007) e1164	- conjugation to solid supports - tetraglycine beads react faster than diglycine beads which react faster than monoglycine beads - reaction conditions: - 85 μM eGFP-LPETGG-His6 - 40 nM His6-sortase - 50 mM Tris-HCl, 150 mM NaCl, 5 mM CaCl2, pH 7.5 - sortase amplified for S. aureus genome - N-terminal membrane targeting sequence removed - 30 kDa on SDS gel
2008 Samantaray et al. J. Am. Chem. Soc. 130	- peptide-sugar conjugation - 6-aminohexoses - peptide antibiotic ligation (aminoglycosides) - conjugates between antibiotics and peptides with yields of 35 to 70% for kanamycin class, about 18-30% for ribostamycin class

year / citation	content
(2008) 2132-2133	- YALPET-sugar adduct - YALPMTGK- sugar adduct - LPNTG motif with S. aureus sortase and peptide

For the enzymatic conjugation a soluble truncated sortase A lacking the membrane-spanning region (SrtA; amino acid residues 60-206 of Staphylococcus aureus SrtA) can be used (SEQ ID NO: 05; see also Ton-That, H., et al., Proc. Natl. Acad. Sci. USA 96 (1999) 12424-12429; Ilangovan, H., et al., Proc. Natl. Acad. Sci. USA 98 (2001) 6056-6061).

The sortase A-mediated reaction results in the ligation of species containing a sortase motif (sequence) with those bearing one or more N-terminal glycine residues. The sortase motif can be the amino acid sequence LPXTG, but can also different therefrom (see below). However, a drawback of using such sequences as acyl donors is that the transfer of the LPXT unit to a nucleophilic acyl acceptor liberates a stoichiometric amount of a corresponding fragment containing at least one N-terminal glycine residue. The liberated glycine-containing fragment competes with the intended acyl acceptor for the enzymatic intermediate and works against the progress of the enzymatic ligation reaction. Additionally the hydrolytic cleavage of the enzymatic intermediate as well as the LPXTG containing substrate, although a relatively slow process, compete with the reaction. In the beginning of the use of the sortase-mediated reaction useful levels of ligation could only be obtained using concentrations of at least 5 mM of the acyl donor comprising the sortase-motif.

The general sortase-motif has the amino acid sequence LPXT, wherein X can be any amino acid residue, i.e. a naturally occurring amino acid residue or a non-naturally occurring amino acid residue. In some embodiments, X is selected from the group of amino acid residues comprising or consisting of (in one letter code) D, E, A, N, Q, K, and R. In some embodiments, the sortase-motif is selected from the group comprising or consisting of the amino acid sequences LPXT, LPXA, SPXT, LAXT, LSXT, NPXT, VPXT, IPXT, LGXT, and YPXR. In some embodiments, the sortase motif is selected from the group of amino acid sequences consisting of LPST, LPKT, LPIT, LPDT, SPKT, LAET, LAAT, LAET, LAST, LAET, LPLT, LSRT, LPET, VPDT, IPQT, YPRR, LPMT, LPLT, LAFT, and LPQT. In certain embodiments in which sortase A is used, the sortase-motif comprises the amino acid sequence X1PX2X3, wherein i) X1 is selected from the group consisting of

the amino acid residues leucine, isoleucine, valine and methionine, ii) X2 is any amino acid, and iii) X3 is selected from the group consisting of threonine, serine and alanine. In specific embodiments, as noted above X1, is leucine and X3 is threonine. In certain embodiments X2 is selected from the group consisting of aspartate, glutamate, alanine, glutamine, lysine and methionine.

In some embodiments the sortase-motif is selected from the group of amino acid sequences comprising or consisting of LPKTG, LPITG, LPDTA, SPKTG, LAETG, LAATG, LAHTG, LASTG, LAETG, LPLTG, LSRTG, LPETG, VPDTG, IPQTG, YPRRG, LPMTG, LPLTG, LAFTG, and LPQTS. In some embodiments of the invention the sortase is a sortase A (SrtA). SrtA recognizes a sortase-motif with the amino acid sequence LPXTG. Common sortase-motif amino acid sequences are, e.g., LPKTG, LPATG, LPETG and LPNTG. In some embodiments LPETG is used. However, sortase-motifs not in line with this consensus sortase-motif amino acid sequence may also be recognized. For example, in some embodiments the sortase-motif comprises the amino acid residue A rather than the amino acid residue T at position 4, e.g. LPXAG or LPNAG. In some embodiments the sortase-motif comprises the amino acid residue A rather than the amino acid residue G at position 5, e.g. LPXTA or LPNTA. In some embodiments the sortase-motif comprises the amino acid residue G rather than the amino acid residue P at position 2, e.g. LGXTG or LGATG. In some embodiments the sortase-motif comprises the amino acid residue I rather than the amino acid residue L at position 1, e.g., IPXTG or IPNTG or IPETG.

In some embodiments, where the sortase-motif is LPXTG or LPXT, X is selected from the group consisting of D, E, A, N, Q, K, and R. In some embodiments X is selected from the group of amino acid residues consisting of K, E, N, Q, and A in an LPXTG or LPXT motif where the sortase is a sortase A. In one embodiment the sortase-motif is LPET or LPETG or LPETA.

In certain embodiments where sortase A from Staphylococcus aureus (St.au. SrtA) is used the sortase-motif has the amino acid sequence LPX1TX2, wherein i) X1 is selected from the group of amino acid residues consisting of D, E, A, N, Q, K, and R, and ii) X2 is selected from the group of amino acid residues consisting of alanine and glycine. In certain embodiments the sortase-motif of St.au. SrtA is LPX1TA. In other embodiments the sortase-motif of St.au. SrtA is LPX1TG. X1 has the meaning as outlined before.

Streptococcus pyogenes sortase A (St.py. SrtA) will accept di-alanine based nucleophiles. This sortase will efficiently cleave the sortase-motif amino acid sequence LPXTA between the threonine and the alanine residue and install modified alanine-based nucleophiles. St.py. SrtA will also recognize and cleave LPXTG motifs, albeit with reduced efficiency.

Staphylococcus aureus sortase A (St.au. SrtA) will not significantly cleave LPXTA motifs or accept alanine based nucleophiles.

In one embodiment, a polypeptide is contacted with Strep. SrtA and an alanine-containing nucleophile. The polypeptide comprises a sortase-motif amino acid sequence that can be recognized by Strep. SrtA at or near its C-terminus and the nucleophile comprises one or more amino acids capable of serving as nucleophile for a St.au. SrtA-mediated reaction at or near its N-terminus (e.g., (G)n, where n is between 1 and 10, e.g., between 1 and 5). This leads to the formation of an LPXTA sequence at the reactive site, a motif refractory to cleavage by St.au. SrtA. This allows for example St.au. SrtA to act on the N-terminus without affecting the C-terminal modification installed with Strep. SrtA.

Sortase fragments having sortase transamidation activity can be used in the methods as reported herein. Sortase fragments can be identified by producing fragments of sortase, for example, by recombinant techniques or proteolytic digestion of full length sortase, and determining the rate of peptide bond formation, i.e. ligation. The fragment can comprise about 80 % of amino acid sequence of full-length sortase, about 70 %, about 60 %, about 50 %, about 40 % or about 30 % of the amino acid sequence of full-length sortase such as that of S. aureus Sortase A (GenBank Accession number AAD48437). In some embodiments the fragment lacks an N-terminal portion of the full-length sortase amino acid sequence that is not essential to the catalytic activity of sortase, for example the fragment lacks the N-terminal portion extending to the end of the membrane anchor sequence. In some embodiments the fragment comprises the C-terminus of a full-length sortase amino acid sequence. In some embodiments, the fragment comprises the catalytic core region of a sortase. In one embodiment the core region is from about position 60 to about position 206 of SrtA, e.g., S. aureus SrtA, or about from position 82 to about position 249 of Strep. SrtA.

Sortases from other organisms also can be utilized in the processes as reported herein. Such sortases often are encoded by nucleotide sequences substantially

identical or similar to the nucleotide sequences that encode SrtA. A similar or substantially identical nucleotide sequence may include modifications to the native sequence, such as substitutions, deletions, or insertions of one or more nucleotides. Included are nucleotide sequences that are at least 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, or 85 % or more identical to a native nucleotide sequence, and often are 90 % or 95 % or more identical to the native nucleotide sequence (each identity percentage can include a 1 %, 2 %, 3 % or 4 % variance). One test for determining whether two nucleic acids are substantially identical is to determine the percentage of identical nucleotide positions shared between two nucleic acids.

SrtA nucleotide sequences may be used as "query sequences" to perform a search against public databases to identify related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (J. Mol. Biol. 215 (1990) 403-410). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, word-length = 12 to obtain homologous nucleotide sequences. To obtain gapped alignments for comparison purposes, gapped BLAST can be utilized as described in Altschul, et al. (Nuc. Acids Res. 25 (1997) 3389-3402). When utilizing BLAST and Gapped BLAST programs, default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used (see e.g. www.ncbi.nlm.nih.gov).

A variant amino acid sequence departs from a native amino acid sequence. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, helix-forming properties and/or amphipathic properties and the resulting variants are screened for enzymatic activity with a suitable assay, such as that reported in European patent application EP14198535. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar or non-polar head groups having similar hydrophilicity values include leucine, isoleucine, valine, glycine, alanine, asparagine, glutamine, serine, threonine, phenylalanine, and tyrosine. In certain embodiments, conservative substitutions may be made, according to the following Table. Amino acids in the same block in the second column and in the same line in the third column may be substituted for one another other in a conservative substitution. Certain conservative substitutions are substituting an amino acid in one row of the third column corresponding to a block in the second column with an amino acid from another row of the third column within the same block in the second column.

aliphatic amino acid residues	non-polar	G, A, P
		I, L, V
	polar, non-charged	C, S, T, M
		N, Q
	polar, charged	D, E
		K, R
aromatic		H, F, W, Y

In certain embodiments homologous substitution may occur, which is a substitution or replacement of like amino acids, such as basic for basic, acidic for acidic, polar for polar amino acids, and hydrophobic for hydrophobic, for example. Non-homologous substitutions can be introduced to a native sequence, such as from one class of residue to another (e. g. a non-hydrophobic to a hydrophobic amino acid), or substituting a naturally occurring amino acid with an unnatural amino acids or non-classical amino acid replacements.

In the methods as reported herein the sortase, the sortase-motif comprising polypeptide (i.e. the acyl donor), and the nucleophile (i.e. the acyl acceptor) are incubated together under conditions suitable to effect the formation of a peptide bond between the N-terminal part of the sortase-motif comprising polypeptide and the nucleophile. As used herein, the term "incubating" or grammatical equivalents thereof denotes that the components of the process are brought in close proximity to one another to allow contact between the molecules. Incubating can be done by adding them to one reaction vessel, for example. The components in the system may be mixed in a variety of manners, such as by oscillating a vessel, subjecting a vessel to a vortex generating apparatus, or repeated mixing with a pipette or pipettes, for example. The components may be added in any order to the system.

The sortase reaction may be performed in any convenient vessel (e.g., tubes such as microfuge tubes, flask, dish), microtiter plates (e.g., 96-well or 384-well plates), glass slides, silicon chips, filters, or any solid or semisolid support having surface (optionally coated) having molecules immobilized thereon and optionally oriented in an array (see, e.g., US 6,261,776 and Fodor, Nature 364 (1993) 555-556), and microfluidic devices (see, e.g., US 6,440,722; US 6,429,025; US 6,379,974; and US 6,316,781).

The reaction mixture is generally cell free and further does not include bacterial cell wall components or intact bacterial cell walls. In some embodiments, the sortase-motif comprising polypeptide and/or the nucleophile are expressed by one

or more recombinant nucleotide sequences in a cell, which nucleotide sequences are integrated into the cell genome or non-integrated (e.g., in a plasmid).

The reaction mixture is maintained at any convenient temperature at which the sortase reaction can be performed. In some embodiments, the sortase reaction is performed at a temperature between and including about 15 °C and about 50 °C. In some embodiments, the sortase reaction is performed at a temperature between and including about 23 °C and about 37 °C. In certain embodiments, the temperature is room temperature (i.e. about 20 °C to 25 °C). The temperature can be optimized by repetitively performing the same sortase reaction at different temperatures and determining ligation rates.

Any convenient volume and component ratio can be used.

In certain embodiments, a (molar) ratio of 1:1000 or greater of sortase enzyme to sortase-motif comprising polypeptide is utilized, or a (molar) ratio of 1:1000 or greater of sortase enzyme to nucleophile is utilized. In specific embodiments, ratios of sortase enzyme to sortase-motif comprising polypeptide or enzyme to nucleophile is about 1:1, including 1:2 or greater, 1:3 or greater, 1:4 or greater, 1:5 or greater, 1:6 or greater, 1:7 or greater, 1:8 or greater, and 1:9 or greater.

In some embodiments, the sortase-motif comprising polypeptide is present at a concentration ranging from about 10 μM to about 10 mM. In some embodiments, the sortase-motif comprising polypeptide is present at a concentration ranging from about 100 μM to about 1 mM. In some embodiments, the sortase-motif comprising polypeptide is present at a concentration ranging from about 100 μM to about 50 mM. In some embodiments, the sortase-motif comprising polypeptide is present at a concentration ranging from about 200 μM to about 1 mM. In some embodiments, the sortase-motif comprising polypeptide is present at a concentration ranging from about 200 μM to about 800 μM. In some embodiments, the sortase-motif comprising polypeptide is present at a concentration ranging from about 400 μM to about 600 μM.

In certain embodiments the nucleophile is present in excess with respect to the sortase-motif comprising polypeptide. In certain embodiments, the nucleophile is present in 10-fold excess with respect to the sortase-motif polypeptide. In certain embodiments, the nucleophile is present in 25-fold excess with respect to the sortase-motif polypeptide. In certain embodiments, the nucleophile is present in 50-fold excess with respect to the sortase-motif polypeptide. In certain embodiments,

the nucleophile is present in 100-fold excess with respect to the sortase-motif polypeptide. In certain embodiments, the nucleophile is present in 250-fold excess with respect to the sortase-motif polypeptide.

In certain embodiments, the nucleophile is present at a concentration ranging from about 1 μM to about 50 mM. In certain embodiments, the nucleophile is present at a concentration ranging from about 15 μM to about 1500 μM. In certain embodiments, the nucleophile is present at a concentration ranging from about 25 μM to about 1000 μM. In certain embodiments, the nucleophile is present at a concentration ranging from about 40 μM to about 250 μM.

In certain embodiments, the sortase is present at a concentration ranging from about 1 μM to about 500 μM. In certain embodiments, the sortase is present at a concentration ranging from about 15 μM to about 150 μM. In certain embodiments, the sortase is present at a concentration ranging from about 25 μM to about 100 μM. In certain embodiments, the sortase is present at a concentration ranging from about 40 μM to about 60 μM.

In certain embodiments, the method is performed in a reaction mixture comprising an aqueous environment. Water with an appropriate buffer and/or salt content often may be utilized. An alcohol or organic solvent may be included in certain embodiments. The amount of an organic solvent often does not appreciably esterify a protein or peptide in the ligation process (e.g., esterified protein or peptide often increase only by 5 % or less upon addition of an alcohol or organic solvent). Alcohol and/or organic solvent contents sometimes are 20 % or less, 15 % or less, 10 % or less or 5 % or less, and in embodiments where a greater amount of an alcohol or organic solvent is utilized, 30 % or less, 40 % or less, 50 % or less, 60 % or less, 70 % or less, or 80 % or less alcohol or organic solvent is present. In certain embodiments, the reaction mixture includes only an alcohol or an organic solvent, with only limited amounts of water if it is present.

In some embodiments, the reaction mixture comprises a buffer. A person skilled in the art will be familiar with a variety of buffers that could be used in accordance with the methods as reported herein. In some embodiments, the buffer solution comprises calcium ions. In certain embodiments, the buffer solution does not contain substances that precipitate calcium ions. In some embodiments, the buffer solution does not include phosphate ions. In some embodiments, the buffer solution does not contain chelating agents.

In some embodiments, the method is performed at a pH value in the range of from 6 to 8.5. In some embodiments, the method is performed at a pH value in the range of from 6 to 8. In some embodiments, the method is performed at a pH value in the range of from 6 to 7.5. In some embodiments, the method is performed at a pH value in the range of from 6.5 to 8.5. In some embodiments, the method is performed at a pH value in the range of from 7 to 8.5. In some embodiments, the method is performed at a pH value in the range of from 7.5 to 8.5. In some embodiments, the method is performed at a pH value in the range of from 7.0 to 8.5. In some embodiments, the method is performed at a pH value in the range of from 7.3 to 7.8.

One or more components of the reaction mixture or the product may be immobilized to a solid support. The attachment between the reaction mixture component and the solid support may be covalent or non-covalent (see, e.g., US 6,022,688 for non-covalent attachments). The solid support may be one or more surfaces of the system, such as one or more surfaces in each well of a microtiter plate, a surface of a glass slide or silicon wafer, BIAcore chip, a surface of a particle, e.g., a bead (see e.g., Lam, Nature 354 (1991) 82-84) that is optionally linked to another solid support, or a channel in a microfluidic device, for example. Types of solid supports, linker molecules for covalent and non-covalent attachments to solid supports, and methods for immobilizing molecules to solid supports are known (see, e.g., US 6,261,776; US 5,900,481; US 6,133,436; US 6,022,688; WO 2001/18234). Any material may be used, e.g., plastic (e.g., polystyrene), metal, glass, cellulose, gels (e.g., formed at least in part from organic polymers such as PDMS), etc. In some embodiments the solid support is semi-solid and/or gel-like, deformable, flexible, or the like.

Any polypeptide, eventually after introduction of a sortase-motif or an oligoglycine or -alanine, may be used as sortase-motif comprising polypeptide or nucleophile in the methods as reported herein.

Summarizing the above, the first substrate, also denoted as donor, comprises the sortase recognition motif. It is cleaved by the sortase after the threonine residue in the recognition motif. Thereby a C-terminal activated carboxyl group (acyl intermediate) is generated. The second substrate, also denoted as acceptor or nucleophile, provides a free N-terminal amino group. Between the free amino group and the activated carboxyl group a peptide bond is formed in the sortase catalyzed transpeptidation reaction.

Thus, for the enzymatic sortase mediated transpeptidation reaction it is only required that a donor comprising a sortase recognition motif and an acceptor comprising an N-terminal free glycine, alanine, cysteine or an equivalent functional group is incubated with a polypeptide having sortase A catalytic activity. The remainder of the donor as well as of the acceptor does not interfere with the reaction.

Thus, a sortase mediated transpeptidation reaction can be performed with virtually any protein or small molecule independently of each other as donor or acceptor as long as these comprise a pair of sortase recognition sequence and nucleophile.

This is confirmed by the art.

For example, Marraffini et al. (Microbiol. Mol. Biol. Rev. 70 (2006) 192-221) reported that sortase A can be used to incorporate chemicals containing glycine residues with a free amino group to the LPXTG motif of recombinant proteins, i.e. without limitation of the protein. Presented examples are the conjugation of triglycyl-lysine-folate with (GFP or Cre or p27)-LPETG-His6 with high efficiency, the incorporation of the branched peptide AT-P-022 into polypeptides, and the self-cleavage of chimeras of His6-sortase-LPETG-target protein (the fusion cleaves itself once the enzyme has been activated by the addition of calcium and triglycine).

Further, Antos et al. (J. Am. Chem. Soc. 131 (2009) 10800-10801) reported that the transpeptidation reaction catalyzed by sortase A allows site-specific derivatization of proteins with virtually any type of functional material. Target proteins are engineered to contain the recognition site (LPXTG) near their C terminus, thus allowing a trans-acylation reaction in which the residues C-terminal to threonine are exchanged for a synthetic oligoglycine peptide. It is reported that the terminal G residue of the sortase recognition motif can be replaced by a methyl ester without imparting the reaction. In this document nucleophiles comprising either a fluorescent label or a protein were used for the conjugation to cholera toxin B subunit.

Further, Popp et al. (Proc. Natl. Acad. Sci. USA 108 (2011) 3169-3174) reported the use of Sortase for polypeptide cyclization and PEGylation. The method is general and applicable to a wide variety of proteins. The sortase transpeptidase reaction allows facile site-specific PEGylation of multiple distinct proteins, as exemplified using interferon a2, GCSF, and erythropoietin. In all cases tested, the site-specific C-terminal PEGylation proceeded efficiently.

In EP 2 990 423 a self-cleaving sortase construct is reported. In this construct the sortase recognition sequence LPETG and the catalytic sortase domain have been combined in the same molecule. As protein comprising the sortase recognition sequence any protein, such as e.g. a protein selected from the group comprising polymer proteins, glycoproteins, cytokines, growth factor, blood preparations, vaccines, hormones, enzymes, antibodies and parts or fragments thereof (isolated light or heavy chains).

IV. Sortases

Full length Streptococcus pyogenes Sortase A (Uniprot Q1J6K9; catalytic core underlined; conserved histidine underlined) has the following amino acid sequence:

MVKKQKRRKI KSMSWARKLL IAVLLILGLA LLFNKPIRNT LIARNSNKYQ
VTKVSKKQIK KNKEAKSTFD FQAVEPVSTE SVLQAQMAAQ QLPVIGGIAI
PELGINLPIF KGLGNTELIY GAGTMKEEQV MGGENNYSLA SHHIFGITGS
SQMLFSPLER AQNGMSIYLT DKEKIYEYII KDVFTVAPER VDVIDDTAGL
KEVTLVTCTD IEATERIIVK GELKTEYDFD KAPADVLKAF NHSYNQVST
(SEQ ID NO: 33).

The amino acid sequence of the mature soluble sortase A derived from Streptococcus pyogenes is

VLQAQMAAQQ LPVIGGIAIP ELGINLPIFK GLGNTELIYG AGTMKEEQVM
GGENNYSLAS HHIFGITGSS QMLFSPLERA QNGMSIYLTD KEKIYEYIIK
DVFTVAPERV DVIDDTAGLK EVTLVTCTDI EATERIIVKG ELKTEYDFDK
APADVLKAFN HSYNQVST (SEQ ID NO: 06).

Full length Staphylococcus aureus Sortase A (see Mazmanian et al.; catalytic core underlined; conserved histidine underlined) has the following amino acid sequence:

MKKWTNRLMT IAGVVLILVA AYLFAKPHID NYLHDKDKDE KIEQYDKNVK
EQASKDKKQQ AKPQIPKDKS KVAGYIEIPD ADIKEPVYPG PATPEQLNRG
VSFAEENESL DDQNISIAGH TFIDRPNYQF TNLKAAKKGS MVYFKVGNET
RKYKMTSIRD VKPTDVGVLD EQKGKDKQLT LITCDDYNEK TGVWEKRKIF
VATEVK (SEQ ID NO: 26).

Staphylococcus aureus Sortase A without the N-terminal 28 amino acid residues (N(2-29) transmembrane domain) has the following amino acid sequence:

MDNYLHDKDK DEKIEQYDKN VKEQASKDKK QQAKPQIPKD KSKVAGYIEI
PDADIKEPVY PGPATPEQLN RGVSFAEENE SLDDQNISIA GHTFIDRPNY
QFTNLKAAKK GSMVYFKVGN ETRKYKMTSI RDVKPTDVGV LDEQKGKDKQ
LTLITCDDYN EKTGVWEKRK IFVATEVK (SEQ ID NO: 34).

Staphylococcus aureus Sortase A without the N-terminal 59 amino acid residues (transmembrane domain) has the following amino acid sequence:

QAKPQIPKDK SKVAGYIEIP DADIKEPVYP GPATPEQLNR GVSFAEENES
LDDQNISIAG HTFIDRPNYQ FTNLKAAKKG SMVYFKVGNE TRKYKMTSIR
DVKPTDVGVL DEQKGKDKQL TLITCDDYNE KTGVWEKRKI FVATEVK
(SEQ ID NO: 05).

Full length Sortase A from Listeria monocytogenes has the following amino acid sequence (the catalytic center is underlined; the conserved histidine is underlined):

MLKKTIAAAA LAAGLLLIFS PFIKNGIVKY MSGHETIEQY KASDIKKNNE
KDATFDFESV QLPSMTSVIK GAANYDKDAV VGSIAVPSVD VNLLVFKGTN
TANLLAGATT MRSDQVMGKG NYPLAGHHMR DESMLFGPIM KVKKGDKIYL
TDLENLYEYT VTETKTIDET EVSVIDDTKD ARITLITCDK PTETTKRFVA
VGELEKTEKL TKELENKYFP SK (SEQ ID NO: 35).

The amino acid sequence of the mature soluble sortase A derived from Listeria monocytogenes Sortase A has the following amino acid sequence:

ANYDKDAVVG SIAVPSVDVN LLVFKGTNTA NLLAGATTMR SDQVMGKGNY
PLAGHHMRDE SMLFGPIMKV KKGDKIYLTD LENLYEYTVT ETKTIDETEV
SVIDDTKDAR ITLITCDKPT ETTKRFVAVG ELEKTEKLTK ELENKYFPSK
(SEQ ID NO: 38).

With the assay as reported in Example 6 it is easily possible to determine which polypeptide or fragment of a sortase A (still) has sortase A enzymatic activity.

V. Sortases in different solvent systems

Sortase A from Staphylococcus aureus was mixed with hexane, triethylamine, acetone, isopropanol, acetonitrile, dimethyl sulfoxide, DES (choline chloride:glycerol 1:2, choline chloride:ethylene glycol, choline chloride:glucose/fructose 1:1 and choline chloride:2,3-butanediol 2,3-butanediol 1:4) or water and analyzed for solubility. The Sa-SrtA is highly soluble in the DES and in water but not in the other solvents.

The activity of the Sa-SrtA in all mentioned solvents was tested. The water-soluble or water-based solvents were supplemented with 15 % buffer to ensure the optimal pH and Ca^{2+} concentration. The reaction mixture (0.1 mM Sa-SrtA and 1 mM of substrates (ULPETGGRR and GGGG-PEG-Biotin)) was incubated (8 h, 37 °C), the reaction mixtures were analyzed by HPLC and the peaks were identified by Mass spectrometry.

In the following Table the solubility and activity of Sa-SrtA in different solvents is shown (+: completely soluble/significant amounts of ligation product were formation). The DES from fructose or glucose was nearly solid at RT and therefore not suitable for this approach. In the DES from ethylene glycol ligation product could be observed but Sa-SrtA is not stable over time.

Solvent	Sa-SrtA solubility	Sa-SrtA activity
Hexane	-	-
Triethylamine	-	-
Acetone	-	-
Isopropanol	-	-
Acetonitrile	-	-
Dimethyl sulfoxide	-	-
ChCl:Glycerin 1:2	+	+
ChCl:Ethylene glycol 1:3	+	+/- (instable)
ChCl:Fructose 1:1	DES is to viscous at mesophilic temperatures	
ChCl:Glucose/ 1:1		
ChCl:2,3-Butanediol 2,3-Butanediol 1:4	+	-
Water	+	+

Thus, Sa-SrtA is active in ChCl:Glycerol based DES containing aqueous co-solvent. To evaluate the influence of the amount of co-solvent on the Sortase reaction, varied ratios of water and DES were investigated. The reactions were analyzed for ligation product formation over time. The results are shown in Figure 7.

To exclude stability issues of Sa-SrtA towards the DES, 0.1 mM Sortase was incubated for 24 h at room temperature in 90 % DES and no loss of activity could be detected.

VI. The method and use as reported herein

One aspect as reported herein is a method for the enzymatic production of a polypeptide comprising the following step

- incubating

  i) a first polypeptide (optionally comprising within the 100 C-terminal amino acid residues) the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue),

ii) a second polypeptide that has i) a glycinyl, an alaninyl, or a cysteinyl compound at its N-terminus, or ii) an oligoglycine, or oligoalanine, or a cysteine amino acid residue followed by one to three glycine or alanine amino acid residues at its N-terminus, or iii) a lysine amino acid residue within its 5 N-terminal amino acid residues, and

iii) a third polypeptide with sortase A activity,

in a deep eutectic solvent

and thereby producing a polypeptide.

In one embodiment the cysteinyl compound is a compound that comprises at position 1 a cysteine amino acid residue with free alpha amino group (in one embodiment a NH_2 or NH_3^+), and a carboxy group, which is part of a peptide bond.

In one embodiment the alanyl compound is a compound that comprises at position 1 an alanine amino acid residue with free alpha amino group (in one embodiment a NH_2 or NH_3^+), and a carboxy group, which is part of a peptide bond.

In one embodiment the glycinyl compound is a compound that comprises at position 1 a glycine amino acid residue with free alpha amino group (in one embodiment a NH_2 or NH_3^+), and a carboxy group, which is part of a peptide bond.

In one embodiment the third polypeptide with sortase A activity is derived from Staphylococcus aureus sortase A, or from Streptococcus pyogenes Sortase A, or from Listeria monocytogenes Sortase A.

In one embodiment the third polypeptide is a sortase A or a sortase A fragment that has sortase A catalytical activity. In one embodiment sortase A catalytical activity is determined using a bond forming assay. In one embodiment the bond forming assay is the assay according to current example 6.

In one embodiment the third polypeptide comprises the amino acid sequence of SEQ ID NO: 05, SEQ ID NO: 06 or SEQ ID NO: 38. In one preferred embodiment the third polypeptide comprises the amino acid sequence of SEQ ID NO: 38.

In one embodiment the third polypeptide comprises additionally a tag at its N- or C-terminus either conjugated directly or via an intervening linker. In one embodiment the third polypeptide is consisting of the amino acid sequence of SEQ

ID NO: 38 and the C-terminal tag of SEQ ID NO: 32. In one embodiment the third polypeptide is consisting of the amino acid sequence of SEQ ID NO: 38.

In one embodiment the method is for the enzymatic conjugation of two polypeptides.

In one embodiment the deep eutectic solvent comprises choline chloride. In one embodiment the deep eutectic solvent is a mixture of choline chloride with glycerol at a molar ratio of 1:2 (v/v). In one embodiment the deep eutectic solvent comprises an aqueous co-solvent. In one embodiment the deep eutectic solvent comprises up to 10 % (v/v) co-solvent. In one embodiment the deep eutectic solvent comprises up to 5 % (v/v) co-solvent. In one preferred embodiment the deep eutectic solvent is a mixture of choline chloride with glycerol at a molar ratio of 1:2 (v/v) comprising up to 5 % (v/v) aqueous co-solvent.

In one embodiment the second polypeptide has at its N-terminus the amino acid sequence GGG, AAA, CGG, CAA, KGG, or KAA.

In one embodiment the first polypeptide comprises within the 250 C-terminal amino acid residues the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue). In one embodiment the first polypeptide comprises within the 100 C-terminal amino acid residues the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue). In one embodiment the first polypeptide comprises within the 25 C-terminal amino acid residues the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue). In one embodiment the first polypeptide comprises within the 10 C-terminal amino acid residues the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue).

In one embodiment the first polypeptide comprises at its C-terminus the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue). In one embodiment the first polypeptide comprises at its C-terminus the amino acid sequence LPETG (SEQ ID NO: 04) or LPETA (SEQ ID NO: 42).

In one embodiment the first polypeptide and the second polypeptide are independently of each other selected from an antibody variable domain, an antibody heavy chain Fab-fragment, an antibody Fc-region, a tag, a peptide comprising the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue), a linker, and a non-sortase motif moiety.

One aspect as reported herein is the use of a deep eutectic solvent comprising choline chloride and glycerol at a molar ratio of 1:2 (v/v) and further comprising up to 5 % (v/v) aqueous co-solvent as solvent in an enzymatic transamidation reaction catalyzed by Sortase A.

One aspect as reported herein is a method for the Sortase A catalyzed production of a polypeptide comprising the following step

- incubating in a deep eutectic solvent comprising choline chloride and glycerol at a molar ratio of 1:2 (v/v) and further comprising up to 5 % (v/v) aqueous co-solvent

    i) a first polypeptide comprising (optionally within the 100 C-terminal amino acid residues) the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue),

    ii) a second polypeptide that has i) a glycinyl, an alaninyl, or a cysteinyl compound at its N-terminus, or ii) an oligoglycine, or oligoalanine, or a cysteine amino acid residue followed by one to three glycine or alanine amino acid residues at its N-terminus, or iii) a lysine amino acid residue within its 5 N-terminal amino acid, and

    iii) a (soluble) sortase that has the amino acid sequence of SEQ ID NO: 05, or SEQ ID NO: 06, or SEQ ID NO: 38 and optionally comprising the C-terminal tag of SEQ ID NO: 32,

and thereby producing a polypeptide.

<u>The first or second polypeptide</u>

The sortase-motif (amino acid sequence) may be conjugated to or incorporated in, if it is not directly comprised in one of these molecules, a therapeutic agent (drug),

a cytotoxic agent (e.g. a toxin such as doxorubicin or pertussis toxin), a fluorophore such as a fluorescent dye like fluorescein or rhodamine, a chelating agent for an imaging or radiotherapeutic metal, a peptidyl or non-peptidyl label, a tag, or a clearance-modifying agent such as various isomers of polyethylene glycol, a peptide that binds to a third component, another carbohydrate or lipophilic agent, or a small molecule, such as e.g. a synthetic small molecule (e.g. acetyl salicylic acid). If the motif is incorporated via conjugation the conjugation can be either directly or via an intervening linker. Furthermore the first and/or second polypeptide can either be recombinantly produced or can be synthetic or semi-synthetic, i.e. recombinantly produced and thereafter chemically modified.

a) Therapeutic agents

The therapeutic agent can be any compound, moiety or group which has a therapeutic effect, such as e.g. an antibody, a cytotoxic or cytostatic compound. The antibody can be a full length or complete antibody or an antigen-binding fragment thereof.

A number of therapeutic antibodies directed against cell surface molecules and their ligands are known, such as Rituxan/MabThera/Rituximab, 2H7/Ocrelizumab, Zevalin/Ibrizumomab, Arzerra/Ofatumumab (CD20), HLL2/Epratuzumab, Inotuzomab (CD22), Zenapax/Daclizumab, Simulect/Basiliximab (CD25), Herceptin/Trastuzumab, Pertuzumab (Her2/ERBB2), Mylotarg/Gemtuzumab (CD33), Raptiva/Efalizumab (Cd11a), Erbitux/Cetuximab (EGFR, epidermal growth factor receptor), IMC-1121B (VEGF receptor 2), Tysabri/Natalizumab (α4-subunit of α4ß1 and α4ß7 integrins), ReoPro/Abciximab (gpIIb-gpIIa and αvß3-integrin), Orthoclone OKT3/Muromonab-CD3 (CD3), Benlysta/Belimumab (BAFF), Tolerx/Oteliximab (CD3), Soliris/Eculizumab (C5 complement protein), Actemra/Tocilizumab (IL-6R), Panorex/Edrecolomab (EpCAM, epithelial cell adhesion molecule), CEA-CAM5/Labetuzumab (CD66/CEA, carcinoembryonic antigen), CT-11 (PD-1, programmed death-1 T-cell inhibitory receptor, CD-d279), H224G11 (c-Met receptor), SAR3419 (CD19), IMC-A12/Cixutumumab (IGF-1R, insulin-like growth factor 1 receptor), MEDI-575 (PDGF-R, platelet-derived growth factor receptor), CP-675, 206/Tremelimumab (cytotoxic T lymphocyte antigen 4), RO5323441 (placenta growth factor or PGF), HGS1012/Mapatumumab (TRAIL-R1), SGN-70 (CD70), Vedotin(SGN-35)/Brentuximab (CD30), and ARH460-16-2 (CD44).

The conjugates obtained with the method as reported herein can be used in the preparation of medicaments for the treatment of e.g. an oncologic disease, a cardiovascular disease, an infectious disease, an inflammatory disease, an autoimmune disease, a metabolic (e.g., endocrine) disease, or a neurological (e.g. neurodegenerative) disease. Exemplary non-limiting examples of these diseases are Alzheimer's disease, non-Hodgkin's lymphomas, B-cell acute and chronic lymphoid leukemias, Burkitt lymphoma, Hodgkin's lymphoma, hairy cell leukemia, acute and chronic myeloid leukemias, T-cell lymphomas and leukemias, multiple myeloma, glioma, Waldenstrom's macroglobulinemia, carcinomas (such as carcinomas of the oral cavity, gastrointestinal tract, colon, stomach, pulmonary tract, lung, breast, ovary, prostate, uterus, endometrium, cervix, urinary bladder, pancreas, bone, liver, gall bladder, kidney, skin, and testes), melanomas, sarcomas, gliomas, and skin cancers, acute idiopathic thrombocytopenic purpura, chronic idiopathic thrombocytopenic purpura, dermatomyositis, Sydenham's chorea, myasthenia gravis, systemic lupus erythematosus, lupus nephritis, rheumatic fever, polyglandular syndromes, bullous pemphigoid, diabetes mellitus, Henoch-Schonlein purpura, post-streptococcal nephritis, erythema nodosum, Takayasu's arteritis, Addison's disease, rheumatoid arthritis, multiple sclerosis, sarcoidosis, ulcerative colitis, erythema multiforme, IgA nephropathy, polyarteritis nodosa, ankylosing spondylitis, Goodpasture's syndrome, thromboangitis obliterans, Sjogren's syndrome, primary biliary cirrhosis, Hashimoto's thyroiditis, thyrotoxicosis, scleroderma, chronic active hepatitis, polymyositis/dermatomyositis, polychondritis, pemphigus vulgaris, Wegener's granulomatosis, membranous nephropathy, amyotrophic lateral sclerosis, tabes dorsalis, giant cell arteritis/polymyalgia, pernicious anemia, rapidly progressive glomerulonephritis, psoriasis, or fibrosing alveolitis.

A number of cell surface markers and their ligands are known. For example cancer cells have been reported to express at least one of the following cell surface markers and or ligands, including but not limited to, carbonic anhydrase IX, alpha fetoprotein, alpha-actinin-4, A3 (antigen specific for A33 antibody), ART-4, B7, Ba-733, BAGE, BrE3-antigen, CA125, CAMEL, CAP-1, CASP-8/m, CCCL19, CCCL21, CD1, CD1a, CD2, CD3, CD4, CDS, CD8, CD1-1A, CD14, CD15, CD16, CD18, CD19, CD20, CD21, CD22, CD23, CD25, CD29, CD30, CD32b, CD33, CD37, CD38, CD40, CD40L, CD45, CD46, CD54, CD55, CD59, CD64, CD66a-e, CD67, CD70, CD74, CD79a, CD80, CD83, CD95, CD126, CD133, CD138, CD147, CD154, CDC27, CDK-4/m, CDKN2A, CXCR4, CXCR7, CXCL12, HIF-

l-alpha, colon-specific antigen-p (CSAp), CEA (CEACAM5), CEACAM6, c-met, DAM, EGFR, EGFRvIII, EGP-1, EGP-2, ELF2-M, Ep-CAM, Flt-1, Flt-3, folate receptor, G250 antigen, GAGE, GROB, HLA-DR, HM1.24, human chorionic gonadotropin (HCG) and its subunits, HER2/neu, HMGB-1, hypoxia inducible factor (HIF-1), HSP70-2M, HST-2or la, IGF-1R, IFN-gamma, IFN-alpha, IFN-beta, IL-2, IL-4R, IL-6R, IL-13R, IL-15R, IL-17R, IL-18R, IL-6, IL-8, IL-12, IL-15, IL-17, IL-18, IL- 25, insulin-like growth factor-1 (IGF-1), KC4-antigen, KS-1-antigen, KS1-4, Le-Y, LDR/FUT, macrophage migration inhibitory factor (MIF), MAGE, MAGE-3, MART-1, MART-2, NY-ESO-1, TRAG-3, mCRP, MCP-1, MIP-1A, MIP-1B, MIF, MUC1, MUC2, MUC3, MUC4, MUC5, MUM-1/2, MUM-3, NCA66, NCA95, NCA90, pancreatic cancer mucin, placental growth factor, p53, PLAGL2, prostatic acid phosphatase, PSA, PRAME, PSMA, P1GF, ILGF, ILGF-1R, IL-6, IL-25, RS5, RANTES, T101, SAGE, S100, survivin, survivin-2B, TAC, TAG-72, tenascin, TRAIL receptors, TNF-alpha, Tn-antigen, Thomson-Friedenreich antigens, tumor necrosis antigens, VEGFR, ED-B fibronectin, WT-1, 17-1A-antigen, complement factors C3, C3a, C3b, C5a, C5, an angiogenesis marker, bcl-2, bcl-6, Kras, cMET, an oncogene marker and an oncogene product (see, e.g., Sensi, et al., Clin. Cancer Res. 12 (2006) 5023-5032; Parmiani, et al, J. Immunol. 178 (2007) 1975-1979; Novellino, et al., Cancer Immunol. Immunother. 54 (2005) 187-207).

Thus, antibodies recognizing specific cell surface receptors including their ligands can be used for specific and selective targeting and binding to a number/multitude of cell surface markers that are associated with a disease. A cell surface marker is a polypeptide located on the surface of a cell (e.g. a disease-related cell) that is e.g. associated with signaling event or ligand binding.

In one embodiment, for the treatment of cancer/tumors multispecific binding molecules/bispecific antibodies are produced that target tumor-associated antigens, such as those reported in Herberman, "Immunodiagnosis of Cancer", in Fleisher (ed.), "The Clinical Biochemistry of Cancer", page 347 (American Association of Clinical Chemists (1979)) and in US 4,150,149; US 4,361,544; and US 4,444,744.

Reports on tumor associated antigens (TAAs) include Mizukami, et al., (Nature Med. 11 (2005) 992-997); Hatfield, et al., (Curr. Cancer Drug Targets 5 (2005) 229-248); Vallbohmer, et al., (J Clin. Oncol. 23 (2005) 3536-3544); and Ren, et al., (Ann. Surg. 242 (2005) 55-63), each incorporated herein by reference with respect to the TAAs identified.

Where the disease involves a lymphoma, leukemia or autoimmune disorder, targeted antigens may be selected from the group consisting of CD4, CD5, CD8, CD14, CD15, CD19, CD20, CD21, CD22, CD23, CD25, CD33, CD37, CD38, CD40, CD40L, CD46, CD54, CD67, CD74, CD79a, CD80, CD126, CD138, CD154, CXCR4, B7, MUC1 or la, HM1.24, HLA-DR, tenascin, VEGF, P1GF, ED-B fibronectin, an oncogene, an oncogene product (e.g., c-met or PLAGL2), CD66a-d, necrosis antigens, IL-2, T101, TAG, IL-6, MIF, TRAIL-R1 (DR4) and TRAIL-R2 (DR5).

A number of bispecific antibodies are known directed against two different targets , such as BCMA/CD3, different antigens of the HER family in combination (EGFR, HER2, HER3), CD19/CD3, IL17RA/IL7R, IL-6/IL-23, IL-1-beta/IL-8, IL-6 or IL 6R/ IL-21 or IL-21R, first specificity directed to a glycoepitope of an antigen selected from the group consisting of Lewis x-, Lewis b- and Lewis y-structures, Globo H-structures, KH1, Tn-antigen, TF-antigen and carbohydrate structures of Mucins, CD44, glycolipids and glycosphingolipids, such as Gg3, Gb3, GD3, GD2, Gb5, Gm1, Gm2, sialyltetraosylceramide and a second specificity directed to an ErbB receptor tyrosine kinase selected from the group consisting of EGFR, HER2, HER3 and HER4, GD2 in combination with a second antigen binding site is associated with an immunological cell chosen from the group consisting of T lymphocytes NK cell, B-lymphocytes, dendritic cells, monocytes, macrophages, neutrophils, mesenchymal stem cells, neural stem cells, ANG2/VEGF, VEGF/PDGFR-beta, Vascular Endothelial Growth Factor (VEGF) acceptor 2/CD3, PSMA/CD3, EPCAM/CD3, combinations of an antigen is selected from a group consisting of VEGFR-1, VEGFR-2, VEGFR-3, FLT3, c FMS/CSF1R, RET, c-Met, EGFR, Her2/neu, HER3, HER4, IGFR, PDGFR, c-KIT, BCR, integrin and MMPs with a water-soluble ligand is selected from the group consisting of VEGF, EGF, PIGF, PDGF, HGF, and angiopoietin, ERBB-3/C-MET, ERBB-2/C-MET, EGF receptor 1/CD3, EGFR/HER3, PSCA/CD3, C-MET/CD3, ENDOSIALIN/CD3, EPCAM/CD3, IGF-1R/CD3, FAPALPHA/CD3, EGFR/IGF-1R, IL 17A/F, EGF receptor 1/CD3, and CD19/CD16.

Toxic drug moieties include: (i) chemotherapeutic agents, which may function as microtubule inhibitors, mitosis inhibitors, topoisomerase inhibitors, or DNA intercalators; (ii) protein toxins, which may function enzymatically; and (iii) radioisotopes.

Exemplary toxic drug moieties include, but are not limited to, a maytansinoid, an auristatin, a dolastatin, a trichothecene, CC1065, a calicheamicin and other enediyne antibiotics, a taxane, an anthracycline, and stereoisomers, isosters, analogs or derivatives thereof.

Protein toxins include diphtheria-A chain, non-binding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain (Vitetta et al (1987) Science, 238:1098), abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP -5), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes (WO 93/21232).

Therapeutic radioisotopes include 32P, 33P, 90Y, 125I, 131I, 131In, 153Sm, 186Re, 188Re, 211At, 212B, 212Pb, and radioactive isotopes of Lu.

The radioisotope or other labels may be incorporated in known ways (Fraker et al (1978) Biochem. Biophys. Res. Commun. 80: 49-57; "Monoclonal Antibodies in Immunoscintigraphy" Chatal, CRC Press 1989). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triamine pentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of a radionuclide to the complex (WO 94/11026).

b) Labels

The non-sortase motif moiety can be a label. Any label moiety which can be covalently attached to the sortase amino acid sequence can be used (see e.g. Singh et al (2002) Anal. Biochem. 304:147-15; Harlow E. and Lane, D. (1999) Using Antibodies: A Laboratory Manual, Cold Springs Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Lundblad R. L. (1991) Chemical Reagents for Protein Modification, 2nd ed. CRC Press, Boca Raton, Fla.). The label may function to: (i) provide a detectable signal; (ii) interact with a second label to modify the detectable signal provided by the first or second label, e.g. to give FRET (fluorescence resonance energy transfer); (iii) affect mobility, e.g. electrophoretic mobility or cell-permeability, by charge, hydrophobicity, shape, or other physical parameters, or (iv) provide a capture moiety, e.g. to modulate ionic complexation.

Conjugates comprising a haptenylated label as reported herein may be useful in diagnostic assays, e.g., for detecting expression of an antigen of interest in specific

cells, tissues, or serum. For diagnostic applications, a bispecific antibody will be used wherein the first binding specificity binds to a target and the second binding specificity binds to a haptenylated label. The hapten will typically be labeled with a detectable moiety. Numerous labels are available which can be generally grouped into the following categories:

(a) Radioisotopes (radionuclides), such as 3H, 11C, 14C, 18F, 32P, 35S, 64Cu, 68Gn, 86Y, 89Zr, 99TC, 111In, 123I, 124I, 125I, 131I, 133Xe, 177Lu, 211At, or 131Bi. Radioisotope labeled conjugates are useful in receptor targeted imaging experiments. The antigen (hapten) can be labeled with ligand reagents that bind, chelate or otherwise complex a radioisotope metal using the techniques described in Current Protocols in Immunology, (1991) Volumes 1 and 2, Coligen et al, Ed. Wiley-Interscience, New York, N.Y., Pubs. Chelating ligands which may complex a metal ion include DOTA, DOTP, DOTMA, DTPA and TETA (Macrocyclics, Dallas, Tex.). Radionuclides can be targeted via complexation with the complex as reported herein (Wu et al, Nature Biotechnology 23(9) (2005) 1137-1146). Receptor target imaging with radionuclide labeled complexes can provide a marker of pathway activation by detection and quantification of progressive accumulation of complexes or corresponding therapeutic antibodies in tumor tissue (Albert et al (1998) Bioorg. Med. Chem. Lett. 8:1207-1210).

Metal-chelate complexes suitable as labels for imaging experiments (US 2010/0111856; US 5,342,606; US 5,428,155; US 5,316,757; US 5,480,990; US 5,462,725; US 5,428,139; US 5,385,893; US 5,739,294; US 5,750,660; US 5,834,456; Hnatowich et al, J. Immunol. Methods 65 (1983) 147-157; Meares et al, Anal. Biochem. 142 (1984) 68-78; Mirzadeh et al, Bioconjugate Chem. 1 (1990) 59-65; Meares et al, J. Cancer (1990), Suppl. 10:21-26; Izard et al, Bioconjugate Chem. 3 (1992) 346-350; Nikula et al, Nucl. Med. Biol. 22 (1995) 387-90; Camera et al, Nucl. Med. Biol. 20 (1993) 955-62; Kukis et al, J. Nucl. Med. 39 (1998) 2105-2110; Verel et al., J. Nucl. Med. 44 (2003) 1663-1670; Camera et al., J. Nucl. Med. 21 (1994) 640-646; Ruegg et al., Cancer Res. 50 (1990) 4221-4226; Verel et al., J. Nucl. Med. 44 (2003) 1663-1670; Lee et al., Cancer Res. 61 (2001) 4474-4482; Mitchell, et al., J. Nucl. Med. 44 (2003) 1105-1112; Kobayashi et al., Bioconjugate Chem. 10 (1999) 103-111; Miederer et al., J. Nucl. Med. 45 (2004) 129-137; DeNardo et al, Clinical Cancer Research 4 (1998) 2483-90; Blend et al, Cancer Biotherapy & Radiopharmaceuticals 18 (2003) 355-363; Nikula et al J. Nucl. Med. 40 (1999) 166-76; Kobayashi et al., J. Nucl. Med. 39 (1998) 829-36;

Mardirossian et al., Nucl. Med. Biol. 20 (1993) 65-74; Roselli et al., Cancer Biotherapy & Radiopharmaceuticals, 14 (1999) 209-20).

(b) Fluorescent labels such as rare earth chelates (europium chelates), fluorescein types including FITC, 5-carboxyfluorescein, 6-carboxy fluorescein; rhodamine types including TAMRA; dansyl; Lissamine; cyanines; phycoerythrins; Texas Red; and analogs thereof. The fluorescent labels can be conjugated to the antigen (hapten) using the techniques disclosed in Current Protocols in Immunology, supra, for example. Fluorescent dyes and fluorescent label reagents include those which are commercially available from Invitrogen/Molecular Probes (Eugene, Oregon, USA) and Pierce Biotechnology, Inc. (Rockford, Ill.).

Detection labels such as fluorescent dyes and chemiluminescent dyes (Briggs et al "Synthesis of Functionalised Fluorescent Dyes and Their Coupling to Amines and Amino Acids," J. Chem. Soc., Perkin-Trans. 1 (1997) 1051-1058) provide a detectable signal and are generally applicable for labeling, especially with the following properties: (i) the labeled conjugate should produce a very high signal with low background so that small quantities of conjugate can be sensitively detected in both cell-free and cell-based assays; and (ii) the labeled conjugate should be photostable so that the fluorescent signal may be observed, monitored and recorded without significant photo bleaching. For applications involving cell surface binding of labeled conjugates to membranes or cell surfaces, especially live cells, the labels should (iii) have good water-solubility to achieve effective conjugate concentration and detection sensitivity and (iv) are non-toxic to living cells so as not to disrupt the normal metabolic processes of the cells or cause premature cell death.

(c) Various enzyme-substrate labels are available or disclosed (see e.g. US 4,275,149). The enzyme generally catalyzes a chemical alteration of a chromogenic substrate that can be measured using various techniques. For example, the enzyme may catalyze a color change in a substrate, which can be measured spectrophotometrically. Alternatively, the enzyme may alter the fluorescence or chemiluminescence of the substrate. The chemiluminescent substrate becomes electronically excited by a chemical reaction and may then emit light which can be measured (using a chemiluminometer, for example) or donates energy to a fluorescent acceptor. Examples of enzymatic labels include luciferases (e.g., firefly luciferase and bacterial luciferase; US 4,737,456), luciferin, 2,3-dihydrophthalazinediones, malate dehydrogenase, urease, peroxidase such as

horseradish peroxidase (HRP), alkaline phosphatase (AP), (3-galactosidase, glucoamylase, lysozyme, saccharide oxidases (e.g., glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase), heterocyclic oxidases (such as uricase and xanthine oxidase), lactoperoxidase, microperoxidase, and the like. Techniques for conjugating enzymes to polypeptides are described in O'Sullivan et al "Methods for the Preparation of Enzyme-Antibody Conjugates for use in Enzyme Immunoassay", in Methods in Enzym. (ed. by J. Langone & IT Van Vunakis), Academic Press, New York, 73 (1981) 147-166.

Examples of enzyme-substrate combinations (US 4,275,149; US 4,318,980) include, for example:

(i) Horseradish peroxidase (HRP) with hydrogen peroxidase as a substrate, wherein the hydrogen peroxidase oxidizes a dye precursor (e.g., orthophenylene diamine (OPD) or 3,3',5,5'-tetramethylbenzidine hydrochloride (TMB));

(ii) alkaline phosphatase (AP) with para-nitrophenyl phosphate as chromogenic substrate; and

(iii) (3-D-galactosidase ((3-D-Gal) with a chromogenic substrate (e.g., p-nitro phenyl-(3-D-galactosidase) or fluorogenic substrate 4-methylumbelliferyl-(3-D-galactosidase.

The labeled conjugate as reported herein may be employed in any known assay method, such as ELISA, competitive binding assays, direct and indirect sandwich assays, and immunoprecipitation assays (Zola, Monoclonal Antibodies: A Manual of Techniques (1987) pp. 147-158, CRC Press, Inc.).

Labeled conjugates as reported herein are useful as imaging biomarkers and probes by the various methods and techniques of biomedical and molecular imaging such as: (i) MRI (magnetic resonance imaging); (ii) MicroCT (computerized tomography); (iii) SPECT (single photon emission computed tomography); (iv) PET (positron emission tomography) Tinianow, J. et al, Nuclear Medicine and Biology, 37(3) (2010) 289-297; Chen et al, Bioconjugate Chem. 15 (2004) 41-49; US 2010/0111856 (v) bioluminescence; (vi) fluorescence; and (vii) ultrasound. Immunoscintigraphy is an imaging procedure in which conjugates labeled with radioactive substances are administered to an animal or human patient and a picture is taken of sites in the body where the conjugate localizes (US 6,528,624). Imaging biomarkers may be objectively measured and evaluated as an indicator of normal

biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention. Biomarkers may be of several types: Type 0 markers are natural history markers of a disease and correlate longitudinally with known clinical indices, e.g. MRI assessment of synovial inflammation in rheumatoid arthritis; Type I markers capture the effect of an intervention in accordance with a mechanism-of-action, even though the mechanism may not be associated with clinical outcome; Type II markers function as surrogate endpoints where the change in, or signal from, the biomarker predicts a clinical benefit to "validate" the targeted response, such as measured bone erosion in rheumatoid arthritis by CT. Imaging biomarkers thus can provide pharmacodynamic (PD) therapeutic information about: (i) expression of a target protein, (ii) binding of a therapeutic to the target protein, i.e. selectivity, and (iii) clearance and half-life pharmacokinetic data. Advantages of in vivo imaging biomarkers relative to lab-based biomarkers include: non-invasive treatment, quantifiable, whole body assessment, repetitive dosing and assessment, i.e. multiple time points, and potentially transferable effects from preclinical (small animal) to clinical (human) results. For some applications, bioimaging supplants or minimizes the number of animal experiments in preclinical studies.

Peptide labeling methods are well known. See Haugland (2003) Molecular Probes Handbook of Fluorescent Probes and Research Chemicals, Molecular Probes, Inc.; Brinkley (1992) Bioconjugate Chem. 3:2; Garman, (1997) Non-Radioactive Labeling: A Practical Approach, Academic Press, London; Means (1990) Bioconjugate Chem. 1:2; Glazer et al Chemical Modification of Proteins. Laboratory Techniques in Biochemistry and Molecular Biology (T. S. Work and E. Work, Eds.) American Elsevier Publishing Co., New York; Lundblad, R. L. and Noyes, C. M. (1984) Chemical Reagents for Protein Modification, Vols. I and II, CRC Press, New York; Pfleiderer, G. (1985) "Chemical Modification of Proteins", Modern Methods in Protein Chemistry, H. Tschesche, Ed., Walter DeGruyter, Berlin and New York; and Wong (1991) Chemistry of Protein Conjugation and Cross-linking, CRC Press, Boca Raton, Fla.); DeLeon-Rodriguez et al, Chem. Eur. J. 10 (2004) 1149-1155; Lewis et al, Bioconjugate Chem. 12 (2001) 320-324; Li et al, Bioconjugate Chem. 13 (2002) 110-115; Mier et al Bioconjugate Chem. 16 (2005) 240-237.

c) Linker

The term "linker" denotes a bifunctional or multifunctional moiety which can be used to conjugate (link) a first moiety with a second moiety. Linked conjugates can be conveniently prepared using a linker having two reactive functionalities.

In one embodiment, a linker has a reactive site which has an electrophilic group that is reactive to a nucleophilic group present in the sortase amino acid sequence. Useful electrophilic groups include, but are not limited to, another thiol, maleimide and haloacetamide groups (see e.g. conjugation method at page 766 of Klussman et al, Bioconjugate Chemistry 15(4) (2004) 765-773).

Examples of thiol-reaction functional groups include, but are not limited to, thiol, maleimide, and alpha-haloacetyl.

The linker may comprise amino acid residues which link the sortase amino acid sequence to the non-sortase motif moiety. The amino acid residues may form a dipeptide, tripeptide, tetrapeptide, pentapeptide, hexapeptide, heptapeptide, octapeptide, nonapeptide, decapeptide, undecapeptide or dodecapeptide unit. Amino acid residues include those occurring naturally, as well as non-naturally occurring amino acid analogs, such as e.g. citrulline or β–amino acids, such as e.g. β-alanine, or ω-amino acids such as 4-amino-butyric acid.

In another embodiment, the linker has a reactive functional group which has a nucleophilic group that is reactive to an electrophilic group present in the sortase amino acid sequence. Useful electrophilic groups include, but are not limited to, aldehyde and ketone carbonyl groups. The heteroatom of a nucleophilic group of a linker can react with an electrophilic group in the sortase amino acid sequence and form a covalent bond to the sortase amino acid sequence. Useful nucleophilic groups on a linker include, but are not limited to, hydrazide, oxime, amino, hydrazine, hydrazine carboxylate, and arylhydrazide. The electrophilic group on an antigen (hapten) provides a convenient site for attachment to a linker.

Typically, peptide-type linkers can be prepared by forming a peptide bond between two or more amino acids and/or peptide fragments. Such peptide bonds can be prepared, for example, according to the liquid phase synthesis method (E. Schroder and K. Lubke "The Peptides", volume 1 (1965) 76-136, Academic Press) which is well known in the field of peptide chemistry.

In another embodiment, the linker may be substituted with groups which modulated solubility or reactivity. For example, a charged substituent such as sulfonate (SO3-) or ammonium or a polymer such as PEG, may increase water solubility of the reagent and facilitate the coupling reaction of the linker reagent with the antigen (hapten) or the drug moiety, or facilitate the coupling reaction depending on the synthetic route employed.

The conjugates comprising a non-sortase motif moiety as reported herein expressly contemplate, but are not limited to, complexes prepared with linker reagents: BMPEO, BMPS, EMCS, GMBS, HBVS, LC-SMCC, MBS, MPBH, SBAP, SIA, SIAB, SMCC, SMPB, SMPH, sulfo-EMCS, sulfo-GMBS, sulfo-KMUS, sulfo-MBS, sulfo-SIAB, sulfo-SMCC, and sulfo-SMPB, and SVSB (succinimidyl-(4-vinylsulfone) benzoate), and including bis-maleimide reagents: DTME, BMB, BMDB, BMH, BMOE, BM(PEO)3, and BM(PEO)4, which are commercially available from Pierce Biotechnology, Inc. Bis-maleimide reagents allow the attachment of e.g. a thiol group to a thiol-containing drug moiety, label, or linker intermediate, in a sequential or concurrent fashion. Other functional groups besides maleimide, which are reactive with e.g. a thiol group, include iodoacetamide, bromoacetamide, vinyl pyridine, disulfide, pyridyl disulfide, isocyanate, and isothiocyanate.

Exemplary linker include a valine-citrulline (val-cit or vc) dipeptide linker reagent having a maleimide stretcher and a para-aminobenzylcarbamoyl (PAB) self-immolative spacer, and a phe-lys(Mtr) dipeptide linker reagent having a maleimide Stretcher unit and a p-amino benzyl self-immolative spacer.

Cysteine thiol groups are nucleophilic and capable of reacting to form covalent bonds with electrophilic groups on linker reagents and the non-sortase motif moiety or the sortase amino acid sequence including: (i) active esters such as NHS esters, HOBt esters, haloformates, and acid halides; (ii) alkyl and benzyl halides, such as haloacetamides; (iii) aldehydes, ketones, carboxyl, and maleimide groups; and (iv) disulfides, including pyridyl disulfides, via sulfide exchange. Nucleophilic groups on a haptenylated compound include, but are not limited to: amine, thiol, hydroxyl, hydrazide, oxime, hydrazine, hydrazine carboxylate, and arylhydrazide groups capable of reacting to form covalent bonds with electrophilic groups on linker moieties and linker reagents.

Coupling of two lipophilic substrates

A DES comprising ChCl:Glycerol and 25 % (v/v) of aqueous co-solvent was used for the conjugation of two both highly lipophilic compounds. Equimolar (0.15 mM) amounts of the substrates were employed and the reaction was analyzed over time. The time course is shown in Figure 8. A convenrsion of 39 % could be obtained in the DES comprising 25 % aqueous co-solvent whereas in the 100 % aqueous system no conversion could be detected.

Glycerol can act as nucleophile in the deep eutectic solvents

A DES comprising ChCl:Glycerol and 25 % (v/v) of aqueous co-solvent was used for the conjugation of two both highly lipophilic compounds.

It has been found that a second unexpected ligation product is formed. Mass spectroscopic analysis revealed the presence of a reaction product from the sortases intermediate (LCRed640-LPET-sortase) with glycerol from the solvent. This high degree of promiscuity was unpredicted due to the fact that the substrate spectrum of the nucleophile is very narrow (Kruger, R. G., et al., Anal. Biochem. 326 (2004) 42-48; Baer, S., et al., Org. Biomol. Chem. 12 (2014) 2675-2685) and that even hydrolysis is rarely observed with wild type Sa-SrtA (Pritz, S. (2008) “Enzymatische Ligation von Peptiden, Peptidnucleinsäuren und Proteinen”; Heck, T., et al., Bioconj. Chem. 25 (2014) 1492-1500).

Thus, one aspect as reported herein is a method for the enzymatic conjugation of a polypeptide with glycerol comprising the following step

- incubating

  i) a polypeptide (optionally comprising within the 100 C-terminal amino acid residues) the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue),

  with

  iii) a second polypeptide with sortase A activity,

  in a deep eutectic solvent comprising choline chloride and glycerol

and thereby producing a polypeptide conjugated to glycerol.

In one embodiment the deep eutectic solvent is a mixture of choline chloride with glycerol at a molar ratio of 1:2 comprising about 25 % (v/v) aqueous co-solvent.

VII. RECOMBINANT METHODS

Any polypeptide domain (e.g. a single chain antigen binding polypeptide such as a scFv, a scFab, or a darpin, or a multi chain antigen binding polypeptide such as a dsFv or a Fab) comprising an nucleophilic amino acid sequence at its N-terminus, such as e.g. an oligoglycine motif (GG (SEQ ID NO: 28), GGG (SEQ ID NO: 29), GGGG (SEQ ID NO: 30), GGGGG (SEQ ID NO: 31)) can be expressed und purified from the supernatant of eukaryotic cells (e.g. HEK293 cells, CHO cells). It does not matter if the polypeptide is an isolated polypeptide or comprised in a multimeric or heteromeric entity.

Suitable host cells for cloning or expression/secretion of polypeptide-encoding vectors include prokaryotic or eukaryotic cells described herein. For example, polypeptides may be produced in bacteria, in particular when glycosylation is not needed (see, e.g., US 5,648,237, US 5,789,199 and US 5,840,523, Charlton, Methods in Molecular Biology 248 (2003) 245-254 (B.K.C. Lo, (ed.), Humana Press, Totowa, NJ), describing expression of antibody fragments in *E. coli*.). After expression, the polypeptide may be isolated from the bacterial cell paste in a soluble fraction or may be isolated from the insoluble fraction so called inclusion bodies which can be solubilized and refolded to bioactive forms. Thereafter the polypeptide can be further purified.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeasts are suitable cloning or expression hosts for polypeptide-encoding vectors, including fungi and yeast strains whose glycosylation pathways have been "humanized", resulting in the production of a polypeptide with a partially or fully human glycosylation pattern (see e.g. Gerngross, Nat. Biotech. 22 (2004) 1409-1414, and Li, et al., Nat. Biotech. 24 (2006) 210-215).

Suitable host cells for the expression of glycosylated polypeptides are also derived from multicellular organisms (invertebrates and vertebrates). Examples of invertebrate cells include plant and insect cells. Numerous baculoviral strains have been identified which may be used in conjunction with insect cells, particularly for transfection of Spodoptera frugiperda cells.

Plant cell cultures can also be utilized as hosts (see, e.g., US 5,959,177, US 6,040,498, US 6,420,548, US 7,125,978 and US 6,417,429 (describing PLANTIBODIES™ technology for producing antibodies in transgenic plants)).

Vertebrate cells may also be used as hosts. For example, mammalian cell lines that are adapted to grow in suspension may be useful. Other examples of useful mammalian host cell lines are the COS-7 cell line (monkey kidney CV1 cell transformed by SV40); the HEK293 cell line (human embryonic kidney); the BHK cell line (baby hamster kidney); the TM4 mouse sertoli cell line (TM4 cells as described, e.g., in Mather, Biol. Reprod. 23 (1980) 243-251); the CV1 cell line (monkey kidney cell); the VERO-76 cell line (African green monkey kidney cell); the HELA cell line (human cervical carcinoma cell); the MDCK cell line (canine kidney cell); the BRL-3A cell line (buffalo rat liver cell); the W138 cell line (human lung cell); the HepG2 cell line (human liver cell); the MMT 060562 cell line (mouse mammary tumor cell); the TRI cell line (e.g. described in Mather, et al., Anal. N.Y. Acad. Sci. 383 (1982) 44-68); the MRC5 cell line; and the FS4 cells-line. Other useful mammalian host cell lines include the CHO cell line (Chinese hamster ovary cell), including DHFR negative CHO cell lines (see e.g. Urlaub, et al., Proc. Natl. Acad. Sci. USA 77 (1980) 4216), and myeloma cell lines such as Y0, NS0 and Sp2/0 cell line. For a review of certain mammalian host cell lines suitable for polypeptide production, see, e.g., Yazaki, and Wu, Methods in Molecular Biology, Antibody Engineering 248 (2004) 255-268 (B.K.C. Lo, (ed.), Humana Press, Totowa, NJ).

Description of the Figures

Figure 1 Chromatogram of the reaction product of Example 3; 1: sortase; 2: GGGWW-BHQ2; 3: LCR640-ULPETGGRRC + B: GGGWW-BHQ2; 4: LCR640-ULPETGGGWW-BHQ2

Figure 2 Mass-spectrum of the peak of the chromatogram of Figure 1; A:1, B:2, C:3, D:4.

Figure 3 Chromatogram of the reaction product of Example 4.

Figure 4 Mass-spectrum of the peak of the chromatogram of Figure 3; A:1, B:2, C:3, D:4, E:5, F:6, G:7, H:8.

Figure 5 Enzymatic activity of soluble St.au. SrtA using a REIA in different solvents; diamond: water; square: DES-1 (ChCl:G); triangle: DES-2 (ChCl: PE).

Figure 6 Enzymatic activity of soluble Listeria monocytogenes SrtA using a REIA in different solvents; diamond: water; square: DES-1 (ChCl:G); triangle: DES-2 (ChCl: PE).

Figure 7 Absolute reaction yields (area under peak) of sortase reaction in DES containing different amounts of aqueous co-solvent. Indicated mixtures of ChCl:Glycerol and water were prepared with 200 mM NaCl, 10 mM $CaCl_2$ and 50 mM Tris/HCl pH 7.5 and 0.05 mM Sa-SrtA, 0.5 mM LCRed640-LPETGGRRC and 5 mM GGGG-Peg-Biotin. Samples were taken at indicated time points and analyzed by HPLC. Error bars represent the standard deviation of three independent measurements.

Figure 8 Absolute reaction yields (area under peak) of sortase reactions in ChCl:Glycerol containing 25 % and 100 % water. Both solvents were prepared with 200 mM NaCl, 10 mM $CaCl_2$ and 50 mM Tris/HCl pH 7.5, 15 µM Sa-SrtA, 0.15 mM LCRed640-LPETGGRRC and 0.15 mM GGWWK-BHQ2. The reaction was analyzed at indicated time points for the generation of ligation product. Error bars represent the standard deviation of three independent measurements.

The following examples, figures and sequences are provided to aid the understanding of the present invention, the true scope of which is set forth in the appended claims. It is understood that modifications can be made in the procedures set forth without departing from the spirit of the invention.

Examples

Recombinant DNA techniques

Standard methods were used to manipulate DNA as described in Sambrook, J. et al., Molecular cloning: A laboratory manual; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989. The molecular biological reagents were used according to the manufacturer's instructions.

Gene and oligonucleotide synthesis

Desired gene segments were prepared by chemical synthesis at Geneart GmbH (Regensburg, Germany). The synthesized gene fragments were cloned into an E. coli plasmid for propagation/amplification. The DNA sequences of subcloned gene

fragments were verified by DNA sequencing. Alternatively, short synthetic DNA fragments were assembled by annealing chemically synthesized oligonucleotides or via PCR. The respective oligonucleotides were prepared by metabion GmbH (Planegg-Martinsried, Germany).

Description of the basic/standard mammalian expression plasmid

For the expression of a desired gene/protein (e.g. full length antibody heavy chain, full length antibody light chain, or an Fc-chain containing an oligoglycine at its N-terminus) a transcription unit comprising the following functional elements is used:

- the immediate early enhancer and promoter from the human cytomegalovirus (P-CMV) including intron A,
- a human heavy chain immunoglobulin 5'-untranslated region (5'UTR),
- a murine immunoglobulin heavy chain signal sequence,
- a gene/protein to be expressed (e.g. shortened Sortase A of Staphylococcus aureus), and
- the bovine growth hormone polyadenylation sequence (BGH pA).

Beside the expression unit/cassette including the desired gene to be expressed the basic/standard mammalian expression plasmid contains

- an origin of replication from the vector pUC18 which allows replication of this plasmid in E. coli, and
- a beta-lactamase gene which confers ampicillin resistance in E. coli.

Protein determination

The protein concentration of purified polypeptides was determined by determining the optical density (OD) at 280 nm, using the molar extinction coefficient calculated on the basis of the amino acid sequence of the polypeptide.

Example 1

Generation of an expression plasmid for sortase A

Staphylococcus aureus derived Sortase A

The sortase gene encodes an N-terminally truncated Staphylococcus aureus sortase A (60-206) molecule (amino acid sequence of SEQ ID NO: 05).

The expression plasmid for the expression of sortase in E.coli cells comprised besides the sortase expression cassette an origin of replication from the vector pUC18, which allows replication of this plasmid in E. coli, and the URA3 gene as selectable marker, and the LacI gene to allow induction of transcription using IPTG.

The transcription unit of the sortase comprised the following functional elements:

- a T5 promoter,
- a purification tag,
- an N-terminally truncated S. aureus sortase A encoding nucleic acid, and
- the To and fd termination sequences.

The expression plasmid for the transient expression of sortase in HEK293 cells comprised besides the sortase expression cassette an origin of replication from the vector pUC18, which allows replication of this plasmid in E. coli, and a beta-lactamase gene which confers ampicillin resistance in E. coli. ^

The transcription unit of the sortase comprised the following functional elements:

- the immediate early enhancer and promoter from the human cytomegalovirus (P-CMV) including intron A,
- a human heavy chain immunoglobulin 5'-untranslated region (5'UTR),
- a murine immunoglobulin heavy chain signal sequence,
- a purification tag encoding nucleic acid,
- an N-terminally truncated S. aureus sortase A encoding nucleic acid, and
- the bovine growth hormone polyadenylation sequence (BGH pA).

The amino acid sequence of the mature sortase is

QAKPQIPKDKSKVAGYIEIPDADIKEPVYPGPATPEQLNRGVSFAEENESLD
DQNISIAGHTFIDRPNYQFTNLKAAKKGSMVYFKVGNETRKYKMTSIRDV
KPTDVGVLDEQKGKDKQLTLITCDDYNEKTGVWEKRKIFVATEVK
(SEQ ID NO: 05).

The purification tag has the amino acid sequence MRGSHHHHHHGS (SEQ ID NO: 32).

<u>Streptococcus pyogenes derived Sortase A</u>

The sortase gene encodes an N-terminally truncated Streptococcus pyogenes sortase A molecule (amino acid sequence of SEQ ID NO: 06).

The expression plasmid for the expression of sortase in E.coli cells comprised besides the sortase expression cassette an origin of replication from the vector pUC18, which allows replication of this plasmid in E. coli, and the URA3 gene as selectable marker, and the LacI gene to allow induction of transcription using IPTG.

The transcription unit of the sortase comprised the following functional elements:

- a T5 promoter,
- a purification tag,
- an N-terminally truncated S. pyogenes sortase A encoding nucleic acid, and
- the To and fd termination sequences.

The expression plasmid for the transient expression of sortase in HEK293 cells comprised besides the sortase expression cassette an origin of replication from the vector pUC18, which allows replication of this plasmid in E. coli, and a beta-lactamase gene which confers ampicillin resistance in E. coli.

The transcription unit of the sortase comprised the following functional elements:

- the immediate early enhancer and promoter from the human cytomegalovirus (P-CMV) including intron A,
- a human heavy chain immunoglobulin 5'-untranslated region (5'UTR),
- a murine immunoglobulin heavy chain signal sequence,
- a purification tag encoding nucleic acid,
- an N-terminally truncated S. pyogenes sortase A encoding nucleic acid, and
- the bovine growth hormone polyadenylation sequence (BGH pA).

The amino acid sequence of the mature sortase is

VLQAQMAAQQLPVIGGIAIPELGINLPIFKGLGNTELIYGAGTMKEEQVMG GENNYSLASHHIFGITGSSQMLFSPLERAQNGMSIYLTDKEKIYEYIIKDVFT VAPERVDVIDDTAGLKEVTLVTCTDIEATERIIVKGELKTEYDFDKAPADV LKAFNHSYNQVST (SEQ ID NO: 06).

The purification tag has the amino acid sequence MRGSHHHHHHGS (SEQ ID NO: 32).

Listeria monocytogenes derived Sortase A

The sortase gene encodes an N-terminally truncated Listeria monocytogenes sortase A (73-222) molecule (amino acid sequence of SEQ ID NO: 38).

The expression plasmid for the expression of the truncated Listeria monocytogenes sortases in E.coli cells comprised besides the sortase expression cassette an origin of replication from the vector pUC18, which allows replication of this plasmid in E. coli, and the URA3 gene as selectable marker, and the LacI gene to allow induction of transcription using IPTG.

The transcription unit of the sortase comprised the following functional elements:

- a T5 promoter,
- a purification tag,
- the Listeria monocytogenes sortase A variant encoding nucleic acid, and
- the To and fd termination sequences.

The expression plasmid for the transient expression of truncated sortases in HEK293 cells comprised besides the sortase expression cassette an origin of replication from the vector pUC18, which allows replication of this plasmid in E. coli, and a beta-lactamase gene which confers ampicillin resistance in E. coli.

The expression plasmid for the transient expression of sortase in HEK293 cells comprised besides the sortase expression cassette an origin of replication from the vector pUC18, which allows replication of this plasmid in E. coli, and a beta-lactamase gene which confers ampicillin resistance in E. coli.

The transcription unit of the sortase comprised the following functional elements:

- the immediate early enhancer and promoter from the human cytomegalovirus (P-CMV) including intron A,
- a human heavy chain immunoglobulin 5'-untranslated region (5'UTR),
- a murine immunoglobulin heavy chain signal sequence,
- a purification tag encoding nucleic acid,
- an N-terminally truncated L. monocytogenes sortase A encoding nucleic acid, and
- the bovine growth hormone polyadenylation sequence (BGH pA).

The purification tag has the amino acid sequence MRGSHHHHHHGS (SEQ ID NO: 32).

Example 2

Transient expression and analytical characterization

E.coli:

The recombinant production of Sortase was performed by growing E.coli cells transformed with the respective Sortase expression plasmids to an OD578 of approx. 0.9 at 37 °C (pre-culture). At this OD578 of approx. 0.9 protein expression was induced by adding 2 mM IPTG and growing the cells for an additional 24 hours at 28 °C. Thereafter, cells were harvested by centrifugation and lysed via high pressure using a homogenizer. Cell lysates were centrifuged to remove cell debris and subsequently the cell lysates were stored at reduced temperature (e.g. -80°C) until purification. Soluble Sortase was purified using Ni-NTA chromatography followed by size exclusion chromatography. For depletion of endotoxins an anion exchange chromatography was performed in flow through mode. The protein concentration of sortase preparations was determined by measuring the optical density (OD) at 280 nm, using the molar extinction coefficient calculated on the basis of the amino acid sequence. Purity and integrity of sortase was determined by SDS-PAGE in the presence and absence of a reducing agent (5 mM 1,4-dithiotreitol) and staining with Coomassie brilliant blue.

The protein concentration was determined by measuring the optical density (OD) at 280 nm, using the molar extinction coefficient calculated on the basis of the amino acid sequence. Purity was analyzed by SDS-PAGE in the presence and absence of a reducing agent (5 mM 1,4-dithiotreitol) and staining with Coomassie brilliant blue.

HEK:

The recombinant production was performed by transient transfection of HEK293 cells (human embryonic kidney cell line 293-derived) cultivated in F17 Medium (Invitrogen Corp.). For transfection "293-Fectin" Transfection Reagent (Invitrogen) was used. Transfection was performed as specified in the manufacturer's instructions. Cell culture supernatants were harvested three to seven (3-7) days after transfection. Supernatants were stored at reduced temperature (e.g. -80°C).

General information regarding the recombinant expression of human immunoglobulins in e.g. HEK293 cells is given in: Meissner, P. et al., Biotechnol. Bioeng. 75 (2001) 197-203.

The protein concentration was determined by measuring the optical density (OD) at 280 nm, using the molar extinction coefficient calculated on the basis of the amino acid sequence. Purity was analyzed by SDS-PAGE in the presence and absence of a reducing agent (5 mM 1,4-dithiotreitol) and staining with Coomassie brilliant blue.

Example 3

Sortase mediated transamidation in deep eutectic solvents

DES-1 in this example is choline chloride mixed with glycerol at a molar ratio of 1:2.

As substrates the following components were used:

sortase-motif containing compound: LCR640-ULPETGGRRC (SEQ ID NO: 45)

nucleophile: GGGWW-BHQ_2 (SEQ ID NO: 46)

The two substrates which have a low solubility in water but good in DES were incubated with two different sortases and the activity determined.

Sortase 1: soluble Staphylococcus aureus Sortase A (SEQ ID NO: 05)

Sortase 2: soluble Listeria monocytogenes Sortase A (SEQ ID NO: 38)

The educts were dissolved in DES-1 (glycerol and choline chloride 2:1) to a final concentration of 0.5 mM. 1 mM Sortase 1 was stored in 50 mM Tris*HCl pH 7.5, 150 mM NaCl and 10 mM CaCl. This two solutions where mixed 19:1 (v/v) and incubated for 18 hours at 37°C.

Reaction mixture (10 µl) was injected on a Vydac C18 column of an LC-Ms system and separated with a 30 min. linear gradient to 100 % buffer B (buffer A (v/v): 95% water, 5% acetonitrile, 0.1% trifluoro acetic acid (TFA); buffer B (v/v): 5% water, 95% Acetonitrile, 0.1 % TFA). The respective chromatogram is shown in Figure 1.

The Analysis of the reaction mixture with LC-MS shows in peak 4 the product of the sortase reaction.

Example 4

Sortase mediated transamidation in aqueous solution

A reaction mixture comprising 0.5 mM of the polypeptide LCR640-ULPETGGGRRC (LCR640 fluorophore conjugated to beta alanine (U); SEQ ID NO: 45) Fc-region fragment comprising a LPETG sortase motif (SEQ ID NO: 04), 1.5 mM of an N-terminal biotinylated N-terminal cysteine comprising peptide with the C-terminally biotinylated amino acid sequence CAAA (SEQ ID NO: 03) and 50 µM Staphylococcus aureus Sortase A in 50 mM Tris pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$ was incubated at 37 °C for 18 h hours.

The reaction was analyzed without stopping.

The samples (10 µl) were injected on a Vydac C18 column of an LC-Ms system and separated with a 30 min. linear gradient to 100 % buffer B (buffer A (v/v): 95% water, 5% acetonitrile, 0.1% trifluoro acetic acid (TFA); buffer B (v/v): 5% water, 95% Acetonitrile, 0.1 % TFA). The respective chromatogram is shown in Figure 3.

The Analysis of the reaction mixture with LC-ESI-TOF MS in positive ion mode shows in peak 4 the product of the native chemical ligation reaction with the mass of 2155 Da.

The respective fragment pattern and masses are shown in the following Table.

LCR-----U-------L------P-----E----T-------G----G----G----R-----R---C-NH												
1840												
726	1114											
1237 (1254)						603						
726	512					603						
1	Educt: LPETG	1840										
2	LCR-GGRR	1329										
3	Educt: LPETG dimer	3676										
4	Educt: G-Bio	830										
5	Educt: C-Bio	919										
6	LCR-LPETG-Bio	2067										
7	LCR-LPETC-Bio	2156										

Example 5

Kinetic assays

Two deep eutectic solvents have been tested in this example. DES-1 is choline chloride mixed with glycerol at a molar ratio of 1:2. DES-2 is choline chloride mixed with ethylene glycol at a molar ratio of 1:3.

The mixture comprising the 2 compounds was heated slowly over a flame and shaken until a clear, uniform solution was formed.

The liquid was then allowed to cool to room temperature. The yield was quantitative and the product had a melting point lower than room temperature.

Reaction 1:

Ratio	Components
1	1 mM St.au. SrtA (SEQ ID NO: 05), 50 mM Tris*HCl, pH 7.5, 200 mM NaCl, 10 mM $CaCl_2$ in water
33	60 µM C-terminally biotinylated oligoglycine (GGGG, SEQ ID NO: 30) in DES
16.5	60 µM glucose dehydrogenase (containing one of the substrates of the sortase reaction (LPXTG; SEQ ID NO: 04) in 50 mM Tris*HCl, pH 7.5, 200 mM NaCl, 10 mM $CaCl_2$ in water.
49.5	DES

Reaction 2:

Ratio	Components
10	1 mM Listeria monocytogenes SrtA (SEQ ID NO: 38), 50 mM Tris*HCl pH 7.5, 200 mM NaCl, 10 mM $CaCl_2$ in water
33	60 µM C-terminally biotinylated oligoglycine (GGGG, SEQ ID NO: 30) in DES
16.5	60 µM glucose dehydrogenase (containing one of the substrates of the sortase reaction (LPXTG; SEQ ID NO: 04)) in 50 mM Tris*HCl, pH 7.5, 200 mM NaCl, 10 mM $CaCl_2$ in water.
40.5	DES

The reaction mixture was incubated at 37 °C for 1, 2, 3 and up to 4 hours. The reaction was stopped by addition of a 20-fold excess of inhibition buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 10 mM $CaCl_2$, 5 mM iodoacetamide). The stopped reaction mixture was centrifuged for 10 min at 5000 x g. The supernatant (50 µL) was added to 100 µL of 50 mM Tris buffer (pH 7.5) comprising 200 mM NaCl, 10 mM $CaCl_2$ and Streptavidin coated magnetic beads. The mixture was incubated for

30 min at 30 °C with shaking at 200 rpm. Thereafter the magnetic beads were washed five times with 300 µL washing buffer each (50 mM Tris, pH 7.5, 200 mM NaCl, 10 mM $CaCl_2$, 5 mg/mL BSA, 0.1 % Triton X-100) in V-bottom micro-titer-plates using a magnet and a vacuum pump. Afterwards the beads were resuspended in 100 µL citrate buffer (200 mM, pH 5.8) and 50 µL thereof are transferred to a new well. Thereto 150 µL test buffer (0.2 M sodium citrate, pH 5.8, 0.3 g/L 4-nitrosoanilin, 1 mM $CaCl_2$, 30 mM glucose) were added. The kinetic of the reporter enzyme was measured over a time period of 5 min at 620 nm.

In comparison the same reaction has been performed in an aqueous buffer system.

The results are shown in Figures 5 and 6.

Example 6

Reporter immobilization assay

Determination of Sortase A enzymatic activity can be done using a reporter immobilization assay (REIA) as reported in European Patent application EP14198535 and as outlined below.

Reaction mixture:

20 µM polypeptide in question
100 µM nucleophile (GGGG/AAAA/CAAA)
20 µM glucose dehydrogenase with C-terminal sortase motif (LPXTG)
250 mM MESNA
0.5 mM TCEP.

The glucose dehydrogenase is expressed and purified as described in WO 2007/118647.

The reaction mixture is prepared in 50 mM Tris-HCl buffer pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$.

The reaction mixture is incubated at 37 °C for up to 60 hours. The reaction is stopped by addition of a 60-fold excess of inhibition buffer (50 mM Tris, pH 7.5, 200 mM NaCl, 10 mM $CaCl_2$, 5 mM iodoacetamide). The stopped reaction mixture is centrifuged for 10 min at 5000 x g. The supernatant (50 µL) is added to 100 µL of 50 mM Tris buffer (pH 7.5) comprising 200 mM NaCl, 10 mM $CaCl_2$ and streptavidin coated magnetic beads. The mixture is incubated for 30 min at 30 °C with shaking at 200 rpm. Thereafter the magnetic beads are washed five times with

300 µL washing buffer each (50 mM Tris, pH 7.5, 200 mM NaCl, 10 mM $CaCl_2$, 5 mg/mL BSA, 0.1% Triton X-100) in V-bottom micro-titer-plates using a magnet and a vacuum pump. Afterwards the beads are resuspended in 100 µL citrate buffer and 80 µL thereof is transferred to a new well. Thereto 150 µL test buffer (0.2 M sodium citrate, pH 5.8, 0.3 g/L 4-nitrosoaniline, 1 mM $CaCl_2$, 30 mM glucose) are added. The kinetic of the reporter enzyme is measured over a time period of 5 min at 620 nm.

The gerenation of a signal is showing that the tested polypeptide has sortase A enzymatic activity.

Example 7

Reaction in different Solvents

Trimethylamine, Acetone, Isopropanol, Acetonitrile, Dimethyl sulfoxide, DES (ChCl:Glycerol 1:2, ChCl:Ethylene glycol 1:3 and ChCl:2,3-Butanediol) and water were used as solvents. Beside Hexane the solvents supplemented with (15 % water, 200 mM NaCl, 10 mM $CaCl_2$ and 50 mM Tris/HCl pH 7.5). 0.1 mM Sa-SrtA, 2 mM ULPETGGRR and 4 mM GGGG-PEG-Biotin were used. For Hexane the Sa-SrtA was used lyophilized. The reactions were incubated for 6 h at 37 °C at 800 rpm and stopped with the addition of 0.2 M HCl 1:1 (v:v) to the reaction mix. 20 µl was injected on a Aeris C18 Column of an HPLC-system and separated with a 31 min. linear gradient from 2 % to 100 % buffer B (buffer A (v/v): 95 % water, 5 % acetonitrile, 0.1% trifluoroacetic acid (TFA); buffer B (v/v): 5 % water, 95 % Acetonitrile, 0.1 % TFA). Peaks were identified using a LC-MS-ESI-system under the same condition.

Example 8

Influence of water Content

Mixtures of (ChCl:Glycerol) DES and water (10 %, 20 %, 30 %, 40 % and 50 %) were prepared and additives were added to a final concentration of (200 mM NaCl, 10 mM CaCl2 and 50 mM Tris/HCl pH 7.5). 0.05 mM Sa-SrtA, 0.5 mM LCRed640-LPETGGRRC and 5 mM GGGG-PEG-Biotin was used. The reaction was stopped after 0.5 h, 1.5 h, 2.5 h, 6 h and 20 h with the addition of 0.2 M HCl 1:1 (v:v) with the reaction mix. 20 µl was injected on a Aeris C18 Column of an HPLC-system and separated with a 31 min. linear gradient from 2 % to 100 % buffer B (buffer A (v/v): 95 % water, 5 % acetonitrile, 0.1 % trifluoroacetic acid (TFA); buffer B (v/v): 5 % water, 95 % acetonitrile, 0.1 % TFA). Peaks were

identified using a LC-MS-ESI-system. Retention times: 8.1 min Sort-tag (ULPETGGRR), 8.6 min Nucleophile (GGGG-Peg-Biotin), 10.2 ligation product (ULPETGGGG-Peg-Biotin) and 13.5 min Sa-SrtA.

Example 9

Coupling of two lipophilic Substrates

A buffer containing 25 % and 100 % water in DES (ChCl:Glycerol) with (200 mM NaCl, 10 mM CaCl2 and 50 mM Tris/HCl pH 7,5) was used and 0.15 mM of both substrates (LCRed640-LPETGGRRC and GGWWK-BHQ2) were weighed in. The mixture was shaken for 1 h at 40 °C and centrifuged. To start the reaction 30 µM Sa-SrtA were added to the supernatant and incubated at 37°C. At indicated time points (20 µl) was injected on a Aeris C18 Column of an HPLC-system and separated with a 8 min. linear gradient from 5 % to 48 % and in 15 min from 48 % to 95 % buffer B (buffer A (v/v): 95 % water, 5 % acetonitrile, 0.1 % trifluoroacetic acid (TFA); buffer B (v/v): 5 % water, 95 %, 0.1 % TFA). Peaks were identified using a LC-MS-ESI-system.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, the descriptions and examples should not be construed as limiting the scope of the invention.

Patent Claims

1. A method for the enzymatic production of a polypeptide comprising the following step

   - incubating

      i) a first polypeptide comprising the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue),

      ii) a second polypeptide that has an oligoglycine, or oligoalanine, or a cysteine amino acid residue followed by one to three glycine or alanine amino acid residues at its N-terminus, and

      iii) a third polypeptide with sortase A activity wherein the third polypeptide

         a) has the amino acid sequence of Staphylococcus aureus sortase A or Listeria monocytogenes Sortase A with 0 to 10 amino acid sequence changes, or

         b) is a shortened variant of the amino acid sequence of Staphylococcus aureus sortase A or Listeria monocytogenes Sortase A,

   in a deep eutectic solvent which is a mixture of choline chloride with glycerol at a molar ratio of 1:2 comprising up to 5 % (v/v) aqueous co-solvent

   and thereby producing a polypeptide.

2. The method according to claim 1, wherein the method is for the enzymatic conjugation of two polypeptides.

3. The method according to any one of claims 1 to 2, wherein the second polypeptide has at its N-terminus the amino acid sequence GGG, AAA, CGG, or CAA.

4. The method according to any one of claims 1 to 3, wherein the first polypeptide comprises at its C-terminus the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue).

5. The method according to any one of claims 1 to 4, wherein the first polypeptide comprises at its C-terminus the amino acid sequence LPETG (SEQ ID NO: 04) or LPETA (SEQ ID NO: 42).

6. The method according to any one of claims 1 to 5, wherein the first polypeptide and the second polypeptide are independently of each other selected from an antibody variable domain, an antibody heavy chain Fab-fragment, an antibody Fc-region, a tag, and a peptide comprising the amino acid sequence LPXTG (SEQ ID NO: 01, wherein X can be any amino acid residue) or LPXTA (SEQ ID NO: 41, wherein X can be any amino acid residue), a linker and a non-sortase motif moiety.