



(51) International Patent Classification:

C07D 207/16 (2006.01) C07K 5/06 (2006.01)
C07K 1/00 (2006.01)

Published:

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

(21) International Application Number:

PCT/EP2018/058488

(22) International Filing Date:

03 April 2018 (03.04.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

17163878.6 30 March 2017 (30.03.2017) EP

(71) Applicant: RIJKSUNIVERSITEIT GRONINGEN
[NL/NL]; Broerstraat 5, 9712 CP Groningen (NL).

(72) Inventors: BARTA, Katalin; c/o Faculty of Science and Engineering, Stratingh Institute of Chemistry, Nijenborgh 4, 9747 AG Groningen (NL). YAN, Tao; c/o Faculty of Science and Engineering, Stratingh Institute of Chemistry, Nijenborgh 4, 9747 AG Groningen (NL). FERINGA, Bernard Lucas; c/o Faculty of Science and Engineering, Stratingh Institute of Chemistry, Nijenborgh 4, 9747 AG Groningen (NL).

(74) Agent: JANSEN, C.M. et al.; V.O., Carnegieplein 5, 2517 KJ Den Haag (NL).

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

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

(54) Title: N-ALKYLATED AMINO ACIDS AND OLIGOPEPTIDES, USES THEREOF AND METHODS FOR PROVIDING THEM.

(57) Abstract: The invention relates to the synthesis of amphiphilic amino acid derivatives, in particular to a method for the N-alkylation of an unprotected amino acid or the N-terminus of an oligopeptide substrate, comprising reacting said unprotected amino acid or oligopeptide substrate with an alcohol, e.g. a fatty alcohol, in the presence of a homogeneous transition metal catalyst.



WO 2018/178397 A1

Title: *N*-alkylated amino acids and oligopeptides, uses thereof and methods for providing them.

5

The invention relates to the synthesis of amphiphilic amino acid derivatives, in particular to *N*-alkylation of amino acids and oligopeptides by the direct alkylation of amines with alcohols. It also relates to the use of the *N*-alkylated products as surfactant.

10 Long-chain alkyl amino acid-based surfactants are known in the art, including *N*-alkyl amino acids, *N*-acyl amino acids, *N*-alkyl amino amides and alkyl amino acid esters (Infante et al., C. R. Chimie, 2004, 7, 583–592). However, the synthesis of *N*-alkyl amino acids is relatively demanding (Bordes and Holmberg, Advances in Colloid and Interface
15 Science, 2015, 222, 79–91). The traditional pathways including *N*-alkylation of amino acids with alkyl halides (Bordes et al., Journal of Colloid and Interface Science 2013, 411, 47–52) and reductive *N*-alkylation of amino acids with aldehyde (Meng et al., Catal. Lett., 2016, 146, 1249–1255). These suffer from various drawbacks, including the production of stoichiometric
20 amount of toxic halide salts as wastes and unstable substrates leading to complex side products, respectively. Furthermore, alkyl halides and aldehydes are mostly synthesized from alkanes and alkenes which are derived from fossil fuels.

It is also known that alcohols are abundant chemical reagents
25 that can be derived from renewable resources through fermentation, depolymerization of lignocellulose (Barta and Ford, Acc. Chem. Res., 2014, 47, 1503–1512) as well as reduction of fatty acids contained in plants oil (Kreutzer, J. Am. Oil. Chem. Soc. 1984, 61, 343 – 348). Whereas alcohols have been already used as reagent to alkylate amines in industrial scale,
30 these known alkylation processes require harsh reaction conditions. Hence,

the present inventors set out to provide an improved method for the synthesis of N-alkylated amino acids which method is environmentally benign and preferably uses building blocks from renewable sources. In particular, they aimed at providing fully sustainable production of (novel) surfactants based on renewable resources i.e. amino acids and long chain alcohols that can be obtained from fatty acids.

It was surprisingly found that this is advantageously achieved by the direct N-alkylation of amino acids with alcohols using a variety of transition metal complexes as catalysts under mild conditions. The process is also highly selective, only produces water as byproduct and allows for much easier purification than other methods.

Accordingly, the invention provides a method for the N-alkylation of an unprotected amino acid or the N-terminus of an oligopeptide substrate, comprising reacting said unprotected amino acid or oligopeptide substrate with an alcohol in the presence of a homogeneous transition metal catalyst.

Direct alkylation of amines with alcohols has been achieved mainly with Ru- or Ir- based catalytic system. However, none of them were successfully applied to zwitterionic amines like unprotected amino acids (Yu et al., Chem. Soc. Rev., 2015, 44, 2305-2329).

Recently, the inventors developed the first Fe- based catalytic system for direct alkylation of amines with alcohols (Yan, Feringa, Barta, Nature comm. 2014, 5, 5602; Yan, Feringa, Barta, ACS Catal. 2016, 6, 381). Notably however, zwitterionic amines like unprotected amino acids were not tested. The present finding that unprotected amino acids and oligopeptides can be directly N-alkylated is highly surprising for at least the following reasons. First, it were to be expected that the alcohol would react with the available carboxyl-moieties such that ester formation would occur instead of or in addition to N-alkylation. Second, the reported Fe- catalyst system used low

or medium polarity solvents such as toluene and cyclopentyl methyl ether (CPME) in which unprotected amino acids are not soluble. Hence, selective *N*-alkylation of unprotected amino acids or oligopeptides according to the present invention could not have predicted.

5 Leonard *et al.* (Org. Process Res. Dev. 2015, 19, 1400-1410) describe the use of the "borrowing hydrogen strategy" in the synthesis of a number of pharmaceutically relevant intermediates, such as the alkylation of a protected amino acid ((*S*)-amino acid ester compound 29) using a diol. In contrast to the present invention Hence, Leonard *et al.* does not relate to the
10 *N*-alkylation of an unprotected amino acid.

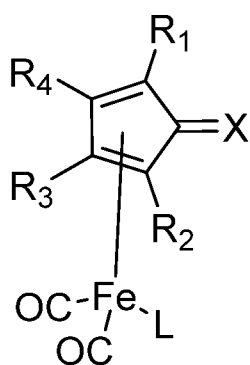
As is demonstrated in the Examples herein below, the present inventors are the first to show direct *N*-alkylation of free α -amino acids and simple peptides with a variety of alcohols using low loadings of a
15 homogeneous transition metal catalyst. The presented atom-economic transformation only results in water as byproduct, thereby significantly simplifying the purification procedure. The reaction is highly selective, and most products were obtained in quantitative yield. Reaction temperature as low as 60°C were used with several substrates. The method is modular and
20 allows for the use of a range of alcohol reaction partners leading to functionalized amino acids with modular properties such as low or high hydrophilicity. Particularly, the use of long chain alcohols and amino acids as only reaction partners to obtain mono-*N*-alkyl amino acids, especially with a molecular iron catalyst, holds great potential for the fully sustainable
25 production of completely bio-based surfactants.

Chiral amines, *N*-alkylated amino acids and peptides moieties are frequently used as substrates or key intermediates to produce high valued chemicals (e.g. drugs), A method of the invention will dramatically reduce the cost and potential pollution (e.g. halide salts are toxic compounds
30 produced in the current methods). Importantly, it was found that in a

method of the invention using a Ru- or Fe-based catalyst, no racemization occurred. Natural amino acids are chiral, and this chirality was maintained in the obtained corresponding *N*-alkylated products. The mild reaction temperature and selectivity of the method as provided herein thus offer a distinct advantage when chirality of the final product is important, e.g. for further application.

Herewith, the invention provides not only a significant addition to sustainable homogeneous catalysis related to the atom economic modification of challenging substrates such as amino acids, but also opens up new possibilities for material science for the production of surfactant as well as for the easy and selective chemical modification of proteins or peptides in biochemistry.

Preferably, the alkylation is performed in the presence of a Fe- or Ru-based catalyst. Very suitable Fe catalysts for use in the present invention include those according to one of the following Formula's A, B or C herein below.



wherein **R1** and **R2** are independently selected from the group consisting of: alkyl, aryl, -CH₂Ph and silyl moieties (e.g. TMS (trimethylsilyl), TBDMS

(tertbutyldimethylsilyl), TIPS (triisopropylsilyl) or TBDPS (tertbutyldiphenylsilyl))

R3 and **R4** are independently selected from the group consisting of:

hydrogen, optionally substituted alkyl and optionally substituted, aryl (with

5 broad range of substitution e.g. -OCH₃, CN, NO₂, COOR, Alkyl)

L is selected from the group consisting of CO, acetonitrile phosphine,

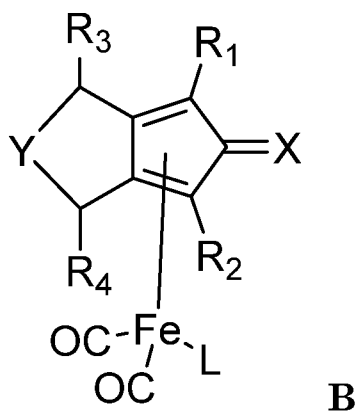
phosphite, phosphoramidite, primary or secondary amine, primary,

secondary or tertiary alcohol,

X is O, NH or N-R where R is an alkyl or aryl, preferably X is O.

10

B



15

wherein R1 and R2 are independently selected from the group consisting of:

alkyl, aryl, -CH₂Ph and,

silyl moieties ((e.g. TMS (trimethylsilyl), TBDMS (tertbutyldimethylsilyl),

TIPS (triisopropylsilyl) or TBDPS (tertbutyldiphenylsilyl));

20 **R3** and **R4** are independently selected from the group consisting of:

hydrogen, optionally substituted alkyl (with broad range of substitution),

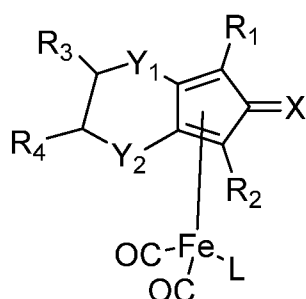
and optionally substituted aryl (with broad range of substitution e.g. H, -

OCH₃, CN, NO₂, COOR, Alkyl), silyl, or OH, O-R; NH₂, NH-R (where R can

be alkyl, aryl, sulphonyl, tosyl, mesyl); **Y** = CH₂; CH₂-CH₂; CH₂-O-; CH₂-NH-; O; NH or NHR₅. wherein R₅=alkyl or aryl, sulphonyl, tosyl, mesyl;

L is selected from the group consisting of CO, acetonitrile phosphine, phosphite, phosphoramidite, primary or secondary amine,
5 primary, secondary or tertiary alcohol,

X is O, NH or N-R₆ where R₆ is an alkyl or aryl, preferably X is O.



C

wherein **R1** and **R2** are independently selected from the group consisting of:
alkyl, aryl, -CH₂Ph and silyl moieties (e.g. TMS [trimethylsilyl], TBDMS
15 [tertbutyldimethylsilyl], TIPS [triisopropylsilyl] or TBDPS
[tertbutyldiphenylsilyl])

where **R3** and **R4** are independently selected from the group consisting of H,
Alkyl (with broad range of substitution), silyl, Aryl (with broad range of
20 substitution on the aromatic ring: e.g. H, -OCH₃, CN, NO₂, COOR, Alkyl); or
OH, O-R₅ (where R₅=alkyl, aryl); NH₂, NH-R₆ (where R₆=alkyl, aryl,
sulphonyl), or wherein R₃ and R₄ are connected to form a cyclic structure,
Y1 and **Y2** = CH₂; O; NH; NHR₇ (where R₇=alkyl, aryl, sulphonyl, tosyl,
mesyl); CH₂-CH₂; CH₂-O-; or CH₂-NH-

L is selected from the group consisting of CO, acetonitrile phosphine, phosphite, phosphoramidite, primary or secondary amine, primary, secondary or tertiary alcohol,

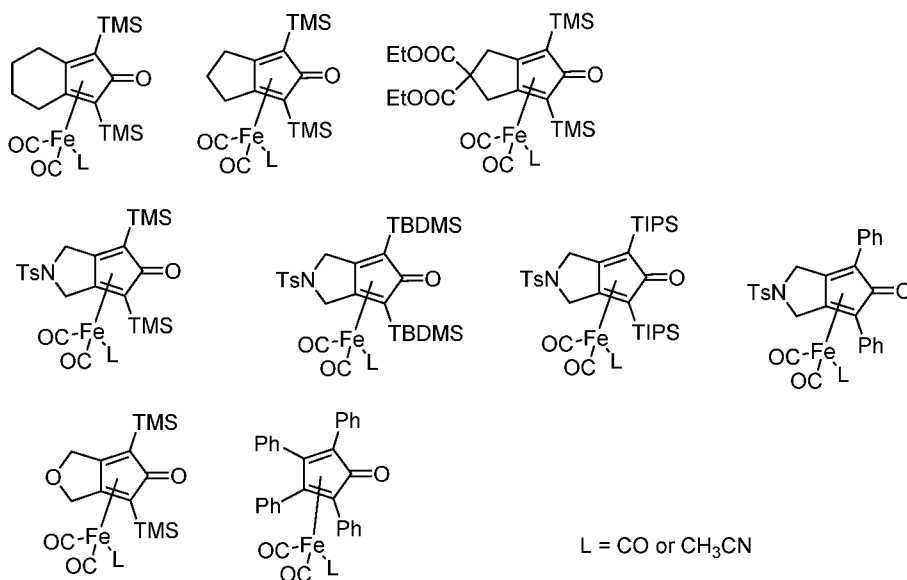
Preferably, in any of the structures A, B or C above, L is CO or acetonitrile.

- 5 In one embodiment, L=CO and catalyst activation is with Me_3NO , base or UV light. In another embodiment, L=acetonitrile and catalyst activation is with heat

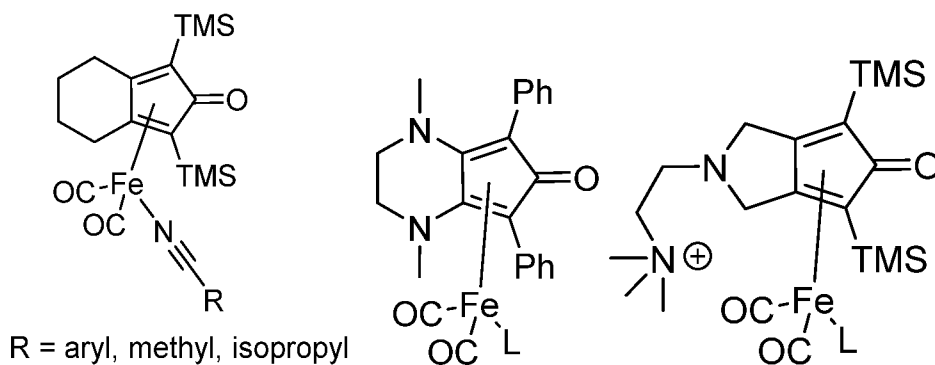
Preferably, in any of the structures A, B or C above, X is oxygen.

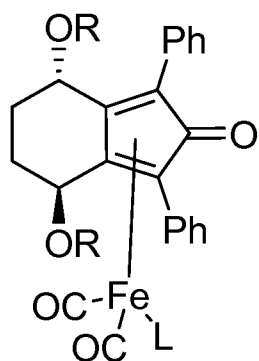
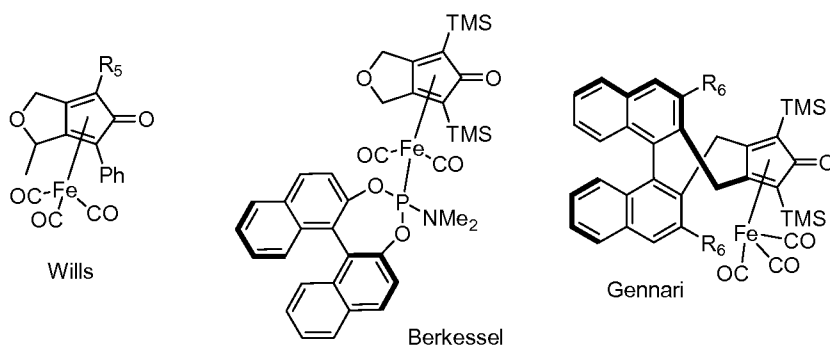
10

Exemplary Fe-based catalysts for use in the present invention include



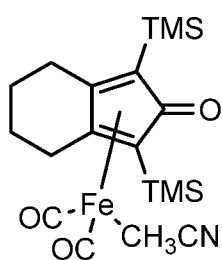
15





R = H, CH₂Ph, TBDMS, TIPS or TBDPS

5 In a specific aspect, the catalyst is of the formula:



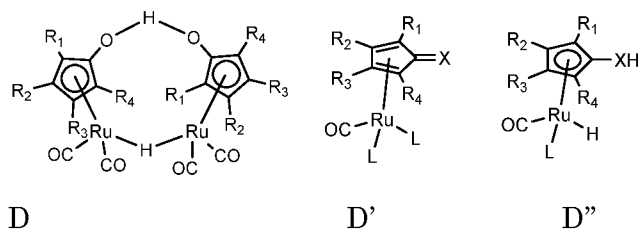
Cat 1b

10

In another embodiment, a method of the invention involves direct *N*-alkylation catalyzed by a Ru catalyst.

Very suitable Ru catalysts for use in the present invention include those

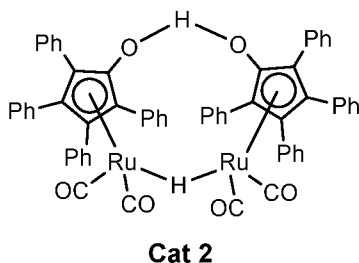
15 according to one of the following Formula's D:



wherein R₁, R₂, R₃ and R₄ are independently selected from the group
 5 consisting of alkyl, aryl, -CH₂Ph, optionally with one or more substitution
 on the aromatic ring, preferably wherein R₁, R₂, R₃ and R₄ are (optionally
 substituted) aryl or -CH₂Ph.

In a preferred embodiment, the Ru-based catalyst is of the formula D.

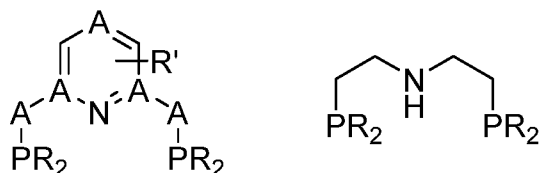
For example, in a specific aspect, the catalyst is 1-
 hydroxytetraphenylcyclopentadienyl-(tetraphenyl-2,4-cyclopentadien-1-one)-
 μ-hydrotetracarbonyldiruthenium(II) (Shvo catalyst) of the formula



Furthermore, a Ru-based catalyst may be prepared *in situ* from a Ru
 precursor consisting of RuCl₃, Ru₃(CO)₁₂, [Ru(*p*-cymene)Cl₂]₂, Ru(acac)₃
 (acetylacetonate) and a ligand which is a monodentate (e.g. triaryl or
 20 trialkyl phosphine), bidentate (dcpe, dppf, DPEphos, Duphos.. etc) or
 tridentate phosphine (triphos and its derivatives).

Furthermore, an Fe catalyst may be prepared in situ from a Fe precursor
 consisting of Fe(CO)₅, Fe₃(CO)₁₂, Fe(BF₄)₂*6H₂O [iron (II) tetrafluoroborate,

hexahydrate]; Fe(OTf)₂ [Iron (II) trifluoromethanesulfonate]; Fe(OAc)₂ [iron (II) acetate]; FeCl₂; FeBr₂ or iron salts with organic acids as anions and a ligand, consisting of monodentate, bidentate, tridentate, or tetradentate phosphines, N-heterocyclic carbenes or the combination thereof. In addition,
 5 the ligand is that of a PNP pincer type E and E' shown below.



R = aryl, alkyl;

R' = aryl, alkyl, amino, alkoxyl.

E

E'

Further, an additional base such as KOH or KOtBu (potassium tert-butoxide) may need to be added.

10

It was found that, when primary amino acids and secondary alcohols were used as reactants, both catalysts gave mono-*N*-alkylated products.

Surprisingly, when primary amino acids and primary alcohols were used, a Ru-based catalyst yielded di-*N*-alkylated products. In contrast, a Fe-based
 15 catalyst could be used to specifically produce mono-*N*-alkylated products, e.g. by control of reaction conditions including reaction time and/or substrate ratio. Herewith, by choosing the catalyst system, a method of the invention therefore also allows to selectively produce mono – or bis-alkylated products.

20

As will be appreciated by a person skilled in the art, a method of the invention is broadly applicable to a diverse set of unprotected amino acids and small unprotected peptides. Amino acids are biologically important organic compounds containing amine (-NH₂) and carboxyl (-COOH)
 25 functional groups, along with a side-chain (R group) specific to each amino

acid. The key elements of an amino acid are carbon, hydrogen, oxygen, and nitrogen, though other elements are found in the side-chains of certain amino acids. About 500 amino acids are known (though only 20 are based on the genetic code) and can be classified in many ways. They can be classified according to the core structural functional groups' locations as alpha- (α -), beta- (β -), gamma- (γ -) or delta- (δ -) amino acids; other categories relate to polarity, pH level, and side-chain group type (aliphatic, acyclic, aromatic, containing hydroxyl or sulfur, etc.). In one aspect, the invention involves the *N*-alkylation of a naturally occurring amino acid. As used herein, a naturally occurring amino acid is one of the 20 common α -amino acids (Gly, Ala, Val, Leu, Ile, Ser, Thr, Asp, Asn, Lys, Glu, Gln, Arg, His, Phe, Cys, Trp, Tyr, Met, and Pro), and other amino acids that are natural products, such as norleucine, ethylglycine, ornithine, methylbutenyl-methylthreonine, and phenylglycine. Typically, the amino acid or oligopeptide substrate used in the present invention comprises or consists of α - amino acids. However, β - and γ - amino acids were also found to be suitable substrates for *N*-alkylation.

As to the side-chain, the amino acid or oligopeptide substrate preferably consists of or comprises an amino acid residue having an uncharged or hydrophobic side chain. In one embodiment, the oligopeptide substrate consists of from two to eight amino acids, preferably two to five amino acids, more preferably wherein the oligopeptide is a dipeptide or a tripeptide.

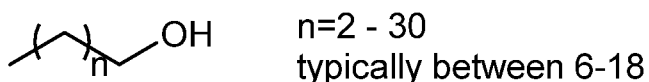
However, the *N*-alkylation of Asp, Glu, His, Lys and Arg may be achieved through technical adaptation of the catalytic system, such as pH tuning e.g. involving the addition of an organic acid or base such as AcOH or Et₃N). Good results were obtained with all naturally occurring amino acids excluding acidic amino acids (Asp, Glu) and basic amino acids (His, Lys, Arg). Accordingly, the unprotected amino acid/oligopeptide is or comprises a natural amino acid other than aspartic acid, glutamic acid, histidine and/or

arginine residue(s). In one embodiment, the amino acid residue is selected from the group consisting of Ser, Thr, Cys, Met, Phe, Try, Ala, Gly, Pro, Val, Leu, Ile, Gln and Asn. Preferred oligopeptides are those comprising or consisting of two or more residues selected from Ser, Thr, Cys, Met, Phe, Try, Ala, Gly, Pro, Val, Leu, Ile, Gln, Asn. For example, it is Gly-Ala, Ala-Gly, Gly-Gly, Leu-Gly, Ala-Leu, Leu-Gly-Gly, Ala-Gly-Ala, and the like.

According to a method of the invention, any alcohol can be directly coupled to the amino acid or oligopeptide substrate. Whereas alcohols containing a single reactive hydroxyl group are typically preferred, e.g. in the synthesis of surfactants, diols may also be used.

In one preferred embodiment, the alcohol is a fatty alcohol. Fatty alcohols (or long-chain alcohols) are usually high-molecular-weight, straight-chain primary alcohols, but can also range from as few as 4–6 carbons to a typical medium chain length of C10-C20 to as many as 26-34, derived from natural fats and oils. The precise chain length varies with the source. Fatty alcohols usually have an even number of carbon atoms and a single alcohol group (–OH) attached to the terminal carbon. Some are unsaturated and some are branched.

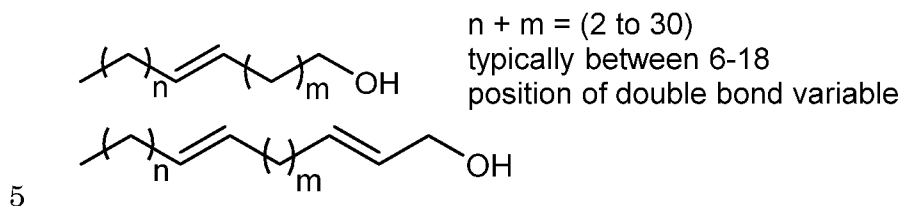
In one embodiment, the alcohol is of the formula



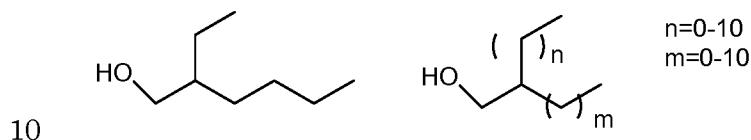
Preferred alcohols for use in the present invention include C8-C18 fatty alcohols, more preferably saturated fatty C8-C18 alcohols. Synthetic fatty alcohols from industrial processes that use syngas and an appropriate catalyst for 'higher alcohol synthesis' can also be used.

In a further aspect, in particular for synthesis of an amino acid based surfactant, the alcohol is an unsaturated long chain, primary alcohol. Many of them can be derived from natural triglycerides (such as sunflower,

palm, cashew, rapeseed, coconut, almond, soy oil) or triglycerides or lipids derived from algae. The chain can have 1, 2 or 3 double bonds.

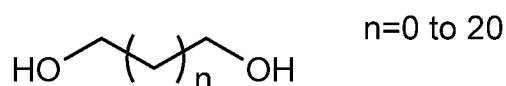


Alternatively, the alcohol is a branched aliphatic alcohol, for instance of the formula



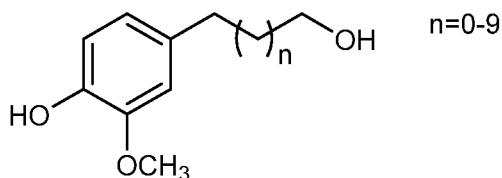
Still further, the alcohol is a long chain aliphatic diol without branching on the chain or a long chain aliphatic diol with branching of alkyl- or aryl substituents on the chain, or a long chain diol with unsaturation in the chain.

In one embodiment, the diol is of the formula



Other embodiments of the invention involve the use of alcohols derived from monolignols (such as coniferyl, coumaryl and synapyl alcohol) or phenol-alcohols derived from lignin depolymerization.

For example, the diol is of the formula



In a specific embodiment, the alcohol is an optionally substituted linear, branched or aromatic C1-C6 alcohol. Suitable substituents include one or more of halogen, oxo, hydroxyl, -CN, alkoxy, amino, amido, phenyl, benzyl hetero-aryl, cyclo-alkyl, heterocycloalkyl, -C(=O)O-R, wherein R is H, C1-C6 alkyl or CH₂-aryl. For example, the alcohol is selected from the group consisting of ethanol, isopropanol, 1-butanol, 2-chloroethanol, 2-butanol, cyclopropylmethanol, benzylalcohol and 4-chlorobenzyl alcohol.

Preferably, the alcohol is derived from biomass or other renewable source. Methods for obtaining (fatty) alcohols are known in the art. For example, US2012/0115195 discloses a method for the production of fatty alcohols from biomass, including cellulose, xylan, hemicellulose, lignin, mannan, and other materials commonly found in biomass. Non-limiting examples of renewable sources include grasses (e.g., switchgrass, Miscanthus), rice hulls, bagasse, cotton, jute, hemp, flax, bamboo, sisal, abaca, straw, leaves, grass clippings, corn stover, corn cobs, distillers grains, legume plants, sorghum, sugar cane, sugar beet pulp, wood chips, sawdust, and biomass. It is also known that alcohols are abundant chemical reagents that can be derived from renewable resources through fermentation, depolymerization of lignocellulose (Barta and Ford, *Acc. Chem. Res.*, 2014, 47, 1503–1512). Preferably, linear long chain alcohols are derived from the reduction of natural fatty acids contained in plant oil (Kreutzer, *J. Am. Oil. Chem. Soc.* 1984, 61, 343 – 348)."

The reaction conditions for performing the *N*-alkylation method of the invention can be optimized using routine skills. Typically, the conditions are

relatively mild. Suitable reaction temperatures are those up to about 110°C, preferably up to about 100°C, like 95°C, 90°C, 85°C or lower. For example, good yields could be obtained at a temperature of only up to 80°C, like about 70°C or even about 60°C. Reaction times can vary e.g. depending on reaction temperature, solvent, catalyst and/or desired yields. Typical incubation times are at least 12 hours, preferably at least 16 hours, like 18 or 20 hours, more preferably at least 24 hours. Preferably, the reaction is performed under an argon atmosphere. In one embodiment, about 0.1 to about 1 mmol amino acid or oligopeptide substrate is reacted with about 2-6 ml alcohol. In a preferred embodiment, CF₃CH₂OH (typically 1 – 4 ml) is added to increase the solubility of the amino acid or oligopeptide substrate. Additional toluene (or an other low polar solvent, e.g. THF, CPME) can be added to tune the polarity of the mixture to obtain a higher catalytic reactivity.

The catalyst is preferably used at about 0.5 to 5 mol%. Preferably, the Fe-based catalyst is used at about 4-5 mol%. The Ru-based catalyst is preferably used at about 0.5-1 mol%.

A further embodiment of the invention relates to an *N*-alkylated amino acid or oligopeptide obtainable by a method according to the invention. For example, there is provided an *N*-alkylated amino acid or oligopeptide selected from the group consisting of the reaction products of Tables 1 to 6 shown herein below. In one embodiment, the invention provides an *N*-alkylated amino acid or oligopeptide selected from the group consisting of the compounds shown in Scheme 2 or 3. Preferred *N*-alkylated amino acid or oligopeptide compounds, e.g. for use as surfactant, include *N,N*-di-(5-hydroxypentyl)-phenylalanine (3dk), *N,N*-di-(*n*-pentyl)-glycine (3gm), *N,N*-(di-*n*-nonyl)-phenylalanine (3dj), *N,N*-di-*n*-nonyl-glycine (3gj), *N,N*-di-dodecylglycine (3gn), *N,N*-di-(5-hydroxy-pentyl)-glycyl-alanine (5ak), *N,N*-di-(*n*-dodecyl)-glycyl-alanine (5an), *N*-nonyl-glycine (3gj'), *N*-decyl-glycine

(3go'), *N*-dodecyl-alanine (3gn'), *N*-tetradecylglycine (3gp'), *N*-hexadecylglycine (3gq'), *N*-octadecylglycine (3gr'), *N*-dodecyl-alanine (3fn') and *N*-nonyl-proline (3aj).

- 5 A method of the invention has many interesting applications. However, it is especially advantageous in the synthesis of an amino acid-based surfactant, preferably in the synthesis of an amino acid-based surfactant from natural building blocks/renewable sources. Hence, in one embodiment the invention provides a method for the synthesis of an amino acid based surfactant,
- 10 preferably a long-chain *N*-alkyl amino acid, comprising the *N*-alkylation of an unprotected amino acid or the N-terminus of an oligopeptide substrate by reacting said unprotected amino acid or oligopeptide substrate with a long chain (fatty) alcohol in the presence of a homogeneous transition metal catalyst. In a preferred aspect, the method comprises mono-*N*-alkylation
- 15 using a Fe-based catalyst as described herein above or in the Examples.

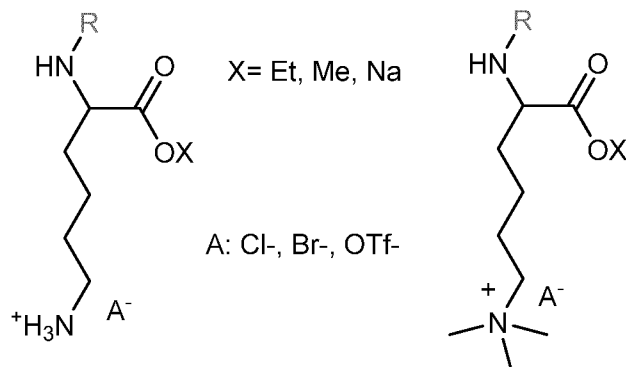
N-alkyl amino acid-based surfactants carry a positive charge at low pH, are zwitterionic at intermediate pH and negatively charged at high pH. This not only affects their physicochemical behavior, for instance their ability to

20 adsorb at charged surfaces, but it also brings about a special feature with respect to self-assembly in bulk and at surfaces. Therefore, the invention also provides the use of an *N*-alkylated amino acid or oligopeptide obtainable by a method of the invention as a surfactant. Preferred surfactants include compounds 3dk, 3gm, 3dj, 3gj, 3gn, 5ak, 5an and those

25 shown in Scheme 3.

Other preferred compounds (obtainable by a method) of the invention include those of the formula

R=alcohol as specified above



Also provided is the use of these compounds as surfactant or as antimicrobial agent.

5

EXPERIMENTAL SECTION

10 Materials and Methods

General methods

Chromatography: Merck silica gel type 9385 230-400 mesh or Merck Al₂O₃ 90 active neutral, TLC: Merck silica gel 60, 0.25 mm. Components were visualized by UV, Ninhydrin or I₂ staining. Progress of the reactions was determined by NMR. Mass spectra were recorded on an AEI-MS-902 mass spectrometer (EI⁺) or a LTQ Orbitrap XL (ESI⁺). ¹H- and ¹³C NMR spectra were recorded on a Varian AMX400 (400 and 100.59 MHz, respectively) using CDCl₃, CD₃OD, D₂O or DMSO-d₆ as solvent. Chemical shift values are reported in ppm with the solvent resonance as the internal standard (CDCl₃: 7.26 for ¹H, 77.00 for ¹³C; CD₃OD: 3.31 for ¹H, 49.00 for ¹³C; D₂O: 4.79 for ¹H; DMSO-d₆: 2.50 for ¹H, 39.52 for ¹³C). Data are reported as follows: chemical shifts, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, br. = broad, m = multiplet), coupling constants (Hz), and integration. All reactions were carried out under an Argon atmosphere

25

using oven (110 °C) dried glassware and using standard Schlenk techniques. Toluene were collected from a MBRAUN solvent purification system (MB SPS-800). CF₃CH₂OH (>99.0%) was purchased from TCI without further purification. Complex **Cat 1a** was synthesized according to T. Yan, B. L. Feringa, K. Barta ACS Catal., 2016, 6, 381–388. **Cat 2** was purchased from Strem. All other reagents were purchased from Sigma, TCI or Acros in reagent or higher grade and were used without further purification.

Representative procedures

General procedure: An oven-dried 20 ml Schlenk tube, equipped with stirring bar, was charged with amino acid (or peptide, given amount), Cat 1 or Cat 2 (given amount) and alcohol (given amount), solvent (or neat). Amino acid (or peptide) and catalyst were added into the Schlenk tube under air, the Schlenk tube was subsequently connected to an argon line and a vacuum-argon exchange was performed three times. Alcohol and solvent was charged under an argon stream. The Schlenk tube was capped. Then the Schlenk was fast sequentially connected to vacuo and argon line, repeat 3 times, for degassing the mixture. And the mixture was rapidly stirred at room temperature for 1 min, then was placed into a pre-heated oil bath at the appropriate temperature and stirred for a given time. The reaction mixture was cooled down to room temperature and concentrated *in vacuo*. The residue was characterized by crude ¹H NMR for determining conversion. Further purification was conducted through flash column chromatography or crystallization to provide the pure *N*-alkyl amino acid (or peptide) product.

Esterification procedure (for preparation of methyl ester of alky amino acids 3gp', 3gq' and 3gr'): Continuing the General procedure, until the reaction mixture was cooled down to room temperature, then 3 ml benzene was added, and under stirring TMSCHN₂ (2M in toluene) was added. The

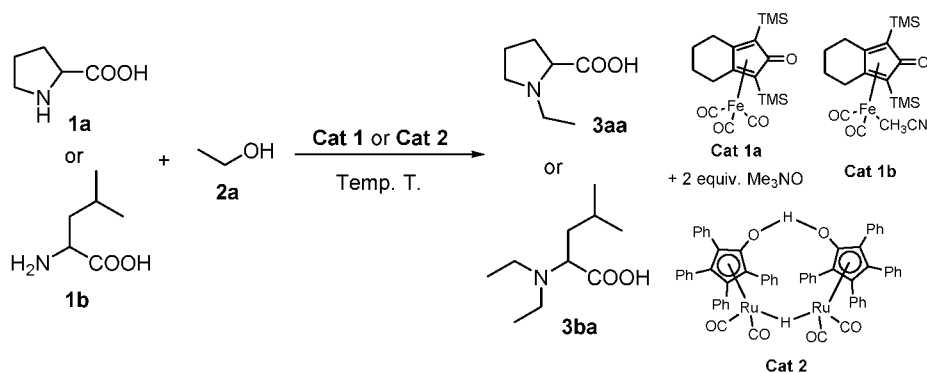
reaction can be monitored by TLC (SiO₂, mono-*N*-alkyl amino acid, R_f = 0.3 in ethylacetate/MeOH = 1/1; methyl mono-*n*-alkyl amino acid ester, R_f = 0.3 in Et₂O). Then the corresponding methyl ester was purified by flash chromatography (tol/Et₂O 50/50 – 0/100).

- 5 Preparation of Cat 1b: Oven-dried 250 ml Schlenk was charged with 100 ml dry acetone and 2 ml dry CH₃CN, degas with N₂ for 20 min. Then 1 g Cat 1a (2.38 mmol) was added under N₂, stirring for 1 min until it fully solubilized. 216 mg Me₃NO (1.2eq) was added under N₂. The direct color changing from yellow to orange was observed in 5 s. The conversion of Cat 1a can be
- 10 monitored by TLC (pentane/ethyl acetate = 1/1, on silica gel, R_f_{Cat1a} = 0.95, R_f_{Cat1b} = 0.35). After 1h, the solvent was removed by vacuum; Cat 1b was purified through flash chromatography, and obtained as brown solid (0.91g, 88% yield). ¹H NMR (400 MHz, CDCl₃): δ 2.05 – 2.48 (m, 4H), 2.21 (s, 3H), 1.38 – 1.73 (m, 4H), 0.22 (s, 18 H). ¹³C NMR (100 MHz, CDCl₃): δ 212.80,
- 15 180.12, 126.00, 106.58, 69.91, 24.83, 22.31, 4.43, -0.12. The physical data were identical in all respects to those previously reported. Cat 1a and Cat 1b are slightly light sensitive but air stable.

20 **EXAMPLE 1 : Direct N-alkylation of amino acids.**

This example exemplifies direct N-alkylation of unprotected amino acids using proline (1a) as exemplary substrate, ethanol (2a) as the alkylation reagent and Cat 1a as the catalyst (Table 1).

- 25 **Table 1:** Optimization of reaction conditions for direct *N*-ethylation of proline (1a) or leucine (1b) with ethanol (2a).



Entry	1	Cat. [mol %]	2a	T.	Temp.	Conv.	Sele. 3 [%] ^a
	[mmol]		[ml]	[h]	[°C]	1	
						[%] ^a	
1	1a / 0.5	Cat 1a / 4, Me ₃ NO / 8	2	18	110	>99	>99
2	1a / 0.5	Cat 1a / 4, Me ₃ NO / 8	2	16	90	>99	>99 (99)
3	1a / 0.5	Cat 1a / 2, Me ₃ NO / 4	2	16	80	14	14
4	1b / 0.5	Cat 1a / 4, Me ₃ NO / 8	4	15	90	15	n.d.
5	1b / 0.2	Cat 1a / 10, Me ₃ NO / 20	5	24	90	60	n.d.
6	1b / 0.2	Cat 1a / 10, Me ₃ NO / 20	5	18	110	70	n.d.
7 ^b	1b / 0.5	Cat 1a / 4, Me ₃ NO / 8	4	24	100	30	n.d.
8 ^b	1b / 0.5	Cat 1b / 4	4	24	100	70	n.d.
9 ^c	1b / 0.5	Cat 1b / 4	4	24	110	18	n.d.
10 ^b	1b / 0.5	Cat 1b / 4	4	24	110	90	n.d.
11 ^b	1b / 0.5	Cat 1b / 4	4	42	110	>95	(49)
12	1a / 0.5	Cat 2 / 0.5	2	18	110	>99	>99
13	1a / 0.5	Cat 2 / 0.5	5	18	90	>99	>99 (99)
14	1a / 0.2	Cat 2 / 1	5	18	60	53	53
15	1b / 0.5	Cat 2 / 1	4	14	90	>99	86
16	1b / 0.2	Cat 2 / 1	5	18	90	>99	>99 (99)

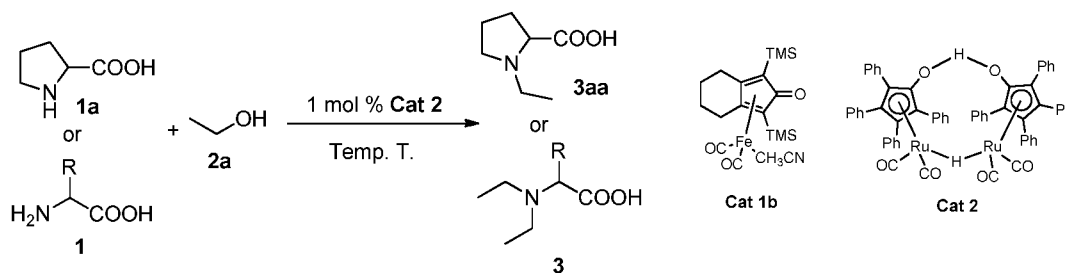
General reaction conditions: General procedure, 0.2 or 0.5 mmol 1a (or 1b), 2 – 5 ml 2a, Cat 1a + Me₃NO, Cat 1b, or Cat 2 (given amount), neat, 14 - 24 h, 60 - 110 °C, isolated yields in parenthesis. ^aConversion and selectivities are based on crude H NMR. ^b1 ml

5 CF₃CH₂OH was added. ^c1 ml H₂O was added.

Quantitative yield of *N*-ethyl proline (3aa) was obtained with 4 mol % of Cat 1a, under 90 °C for 16 h (Table 1, entry 2). However, when leucine (1b) was tried as the substrate, the results of producing *N,N*-diethyl leucine (3ba) turned out to be less satisfactory (Table 1, entry 4 - 7). Then Cat 1a's
5 analogue, Cat 1b was employed as the catalyst, full conversion of 1b and 49% yield of 3aa were obtained, with using CF₃CH₂OH as the solvent, under 110 °C, for 42 h (Table 1, entry 8 - 11). The ruthenium catalyst (Cat 2) was also selected as the catalyst. Surprisingly, quantitative yields of both 3aa and 3ba from 1a and 1b were observed (Table 1, entry 13 and 16). The Shvo
10 catalyst (Cat 2) has been well studied in a variety of reactions including transfer hydrogenation. However, its ability to promote *N*-alkylation of amines with alcohols has not been suggested.

General reaction conditions: General procedure 0.2 mmol **1**, 4 - 5 ml **2**, 1 mol % **Cat 2**, neat when using ethanol (**2a**), 1 ml CF₃CH₂OH (**2d**) when using *i*PrOH (**2b**), 18 - 28 h, 90 °C, yields are based on crude H NMR and mass balance, ee was measured through corresponding amide, unless otherwise specified; ^aee was measured through corresponding methyl ester; ^b47h; ^cneat; ^dH₂O was used instead of **2d**; ^eMeOH (**2c**) was used instead of **2d**; ^f2 mol % **Cat 2** was used; ^g100 °C; ^hisolated yield. For details see Table 2 and 3.

Table 2: Direct *N*-ethylation of free amino acid (**1**) with ethanol (**2a**) and the ee retention.



Entry	1 [mmol]	2a [ml]	T. [h]	Temp. [°C]	Conv. 1 [%] ^a	Sele. 3 [%] ^a	ee [%] ^b
1	1a L-Proline / 0.2	5	18	90	>99	3aa >99 (99)	93.2
2 ^e	1a L-Proline / 0.5	5	18	90	>99	3aa >99 (99)	99.2
3	1b L-Leucine / 0.2	5	18	90	>99	3ba >99 (99)	98.5
4 ^{de}	1b L-Leucine / 0.5	4	42	110	>95	3ba (49)	79.7
5	1c L-Valine / 0.2	5	18	90	>99	3ca 55	n.d.
6	1c L-Valine / 0.2	5	24	90	>99	3ca >99	99.5
7	1d L-Phenylalanine / 0.2	5	18	90	>99	3da >99	97.2 ^c
8 ^{de}	1d L-Phenylalanine / 0.5	4	42	110	>90	3da (55)	72.0
9	1e L-Serine / 0.2	5	18	90	>99	3ea >99	86.1
10	1f L-Alanine / 0.2	5	18	90	>99	3fa >99	84.1
11	1f L-Alanine / 0.2	5	47	60	>99	3fa >99	85.6

12	1g	Glycine / 0.2	5	18	90	>99	3ga	>99	/
13	1h	Lysine	5	18	90	<1%	3ha	/	/
14 ^d	1h	Lysine	4	18	100	<1%	3ha	/	/
15	1i	N ⁶ -Ac-lysine	5	18	90	>99	3ia	<5%	/
16 ^d	1i	N ⁶ -Ac-lysine	4	18	90	>99	3ia	(74)	/

General reaction conditions: General procedure (see page S2), 0.2 mmol **1**, 5 ml **2a**, 1 mol % **Cat 2**, neat, 18 - 47 h, 60 or 90 °C, unless otherwise specified, isolated yields in parenthesis; ^aConversion and selectivities are based on crude H NMR; ^bee was measured through corresponding amide unless otherwise specified, see page S55 for details of ee determination; ^cee was measured through corresponding methyl ester; ^d1 ml CF₃CH₂OH was added; ^e5 mol % **Cat 1b** was used instead of **Cat 2**.

As can be seen in Table 2, Cat 1b gave 3aa quantitative yield with 99.2 % ee retention, 3ba 45 % yield with 79.7 % ee retention (Table 2, entry 2 and 4). At the meantime, Cat 2 gave both 3aa and 3ba quantitative yields as described, with retention of ee of 93.2 % and 98.5 %, respectively (Table 2, entry 1 and 3). Phenylalanine (1d) was selected to react with 2a, catalyzed by both Cat 1b and Cat 2 for further comparison. Cat 1b gave 3da 55 % yield with 72.0 % ee retention, and Cat 2 gave 3da quantitative yield with 97.2 % ee retention (Table 2, entry 7 and 8). Thus Cat 2 was chosen for the further investigation.

Then, other α-amino acids including valine (1c), serine (1e) and alanine (1f) were evaluated. Quantitative yields of corresponding di-ethylation products 3ca, 3ea and 3fa were obtained, with the retention of ee of 99.5, 86.1 and 84.1 %, respectively (Table 2, entry 6, 9 and 10). The racemization in the latter cases was probably due to the activity of the Shvo catalyst in the dehydrogenation/rehydrogenation of amines. A lower steric hindrance on the chiral carbons of 3ea and 3fa compared to 3ca presumably allows for easier dehydrogenation/rehydrogenation of amines. In order to improve the ee

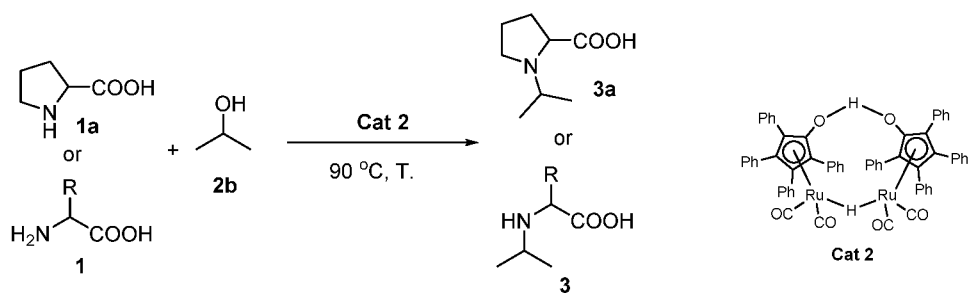
retention of 3fa, the temperature was lowered to 60°C from 90°C while prolonging the reaction time. A quantitative yield of 3fa was obtained, whereas the ee of the final product was improved slightly to 85.6 % (Table 2, entry 11).

5 Continuing exploring the reaction scope, the simplest α -amino acid glycine (1g), also gave quantitative yield of N,N-di-ethyl-glycine (3ga) (Table 2, entry 12). On the other hand, lysine (3h) was unreactive, whereas N6-acetyl-lysine (3i) give 74 % yield of corresponding N2,N2-di-ethylation product 3ia (Table 2, entry 13 - 16). The above results clearly showed that all primary
10 amino acids selectively gave di-N,N-ethylated products when reacted with 2a in neat condition.

EXAMPLE 3: N-alkylation using a secondary alcohol.

Following the results of *N*-ethylation of α -amino acids with 2a, the
15 secondary alcohol isopropanol (2b) was used applied to alkylate α -amino acids (Scheme 2, Table 3).

Under neat conditions, 2a was quantitatively converted to *N*-isopropyl-proline (3ab) (Table 3, entry 1). However, in case of phenylalanine (1d), no significant formation of corresponding *N*-alkyl amino acid was observed due
20 to the poor solubility of 1d in isopropanol (2b). This prompted us to investigate using solvent like H₂O, MeOH (2c) or CF₃CH₂OHⁱ (2d). While H₂O or 2c gave a less satisfactory improvement, surprisingly, the use of 2d gave quantitative yield of *N*-isopropyl-phenylalanine (3db) (Table 3, entry 3 - 9). Only the mono alkylation product was observed in this case, probably
25 due the steric hindrance created after the insertion of the first isopropyl-moiety inhibited the second alkylation step.

Table 3: Direct mono *N*-isopropylation of amino acid (**1**) with isopropanol (**2b**).

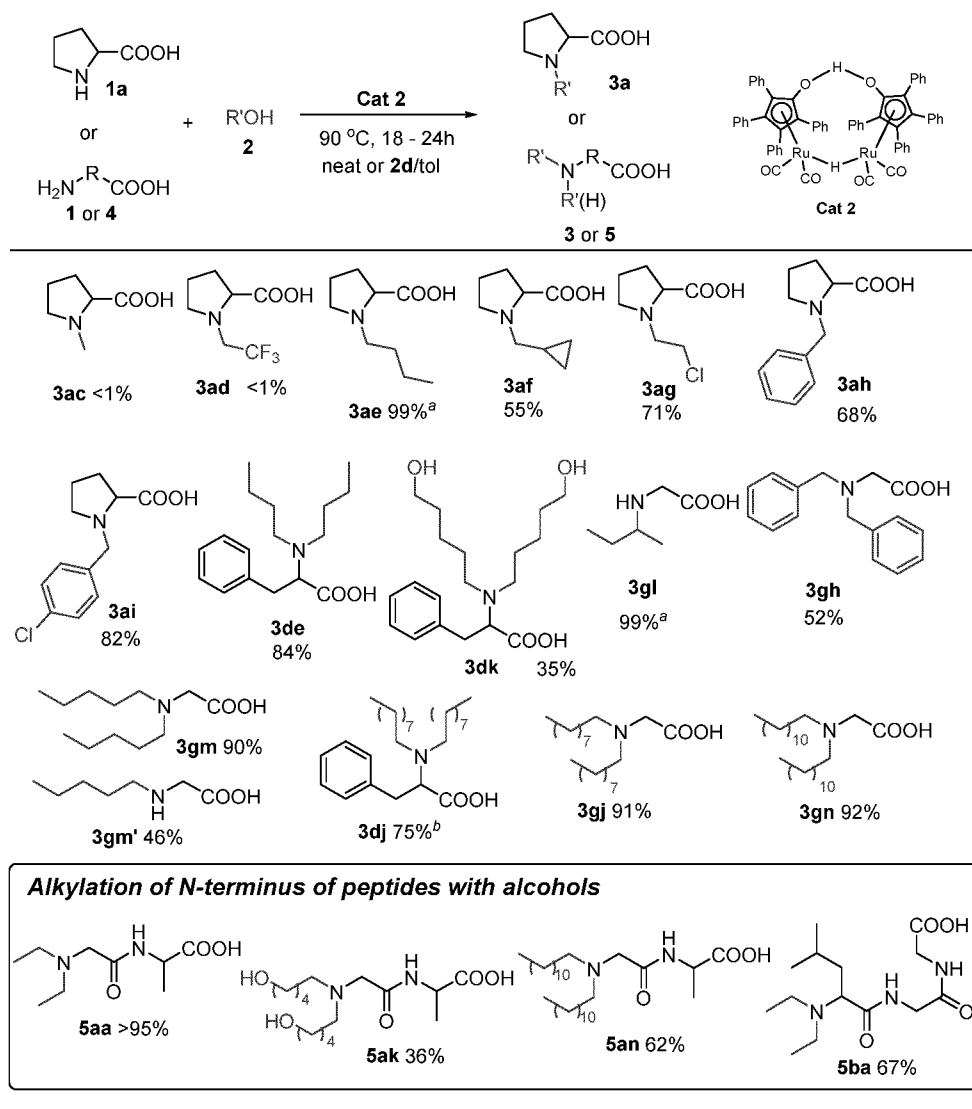
Entry		1 [mmol]	2b [ml]	Sol. [ml]	T. [h]	Conv. 1 [%] ^a	Sele. 3 [%] ^a
1	1a	Proline / 0.2	2	/	18	>99	3cb >99 (99)
2 ^b	1a	Proline / 0.5	2	/	16	88	n.d.
3	1d	Phenylalanine / 0.2	5	/	18	<5	n.d.
4	1d	Phenylalanine / 0.2	4.5	H ₂ O 0.5	45	<1	n.d.
5	1d	Phenylalanine / 0.2	4.5	2c 0.5	45	15	n.d.
6	1d	Phenylalanine / 0.2	4	2c 1	18	12	n.d.
7	1d	Phenylalanine / 0.2	4	2c 1	18	85	n.d.
8	1d	Phenylalanine / 0.2	3	2c 2	18	10	n.d.
9	1d	Phenylalanine / 0.2	4	2d 1	24	>99	3db >99 (99)
10	1b	Leucine / 0.2	5	/	18	<5	n.d.
11	1b	Leucine / 0.2	4	2d 1	24	40	n.d.
12 ^c	1b	Leucine / 0.2	4	2d 1	28	>99	3bb >99 (99)

13	1b	Leucine / 0.2	4	2d 1	42	<5		n.d.
14	1c	Valine / 0.2	5	/	18	<5		n.d.
15	1c	Valine / 0.2	4	2d 1	24	>99	3cb	>99 (99)
16	1e	Serine / 0.2	5	/	18	<5		n.d.
17	1e	Serine / 0.2	4	2d 1	24	15		n.d.
18 ^c	1e	Serine / 0.2	4	2d 1	24	40		n.d.
19 ^d	1e	Serine / 0.2	4	2d 1	24	>99	3eb	>99 (99)
20	1f	Alanine / 0.2	5	/	18	<5		n.d.
21	1f	Alanine / 0.2	4	2d 1	24	67		n.d.
22	1f	Alanine / 0.2	4	2d 1	28	>99	3fb	>99 (99)

General reaction conditions: General procedure 0.2 or 0.5 mmol **1**, 2 - 5 ml **2b**, neat or with solvent (amount shown in the table), 1 mol % **Cat 2**, 16 - 28 h, 90°C, unless otherwise specified, isolated yields in parenthesis; ^aConversion and selectivities are based on crude H NMR; ^b4 mol % **Cat 1b** was used instead of **Cat 2**; ^c2 mol % **Cat 2** was used; ^d100 °C.

Following the same procedure, amino acids **1b**, **1c**, **1e** and **1f** were also successfully mono-isopropylated, forming **3bb**, **3cb**, **3eb** and **3fb** in quantitative yields, respectively (**Table 3**, entry 10 - 22). This method is advantageously used to easily obtain mono-*N*-alkylated amino acids (non-proteinogenic amino acids) e.g. for the synthesis of modified proteins with higher lipophilicity.

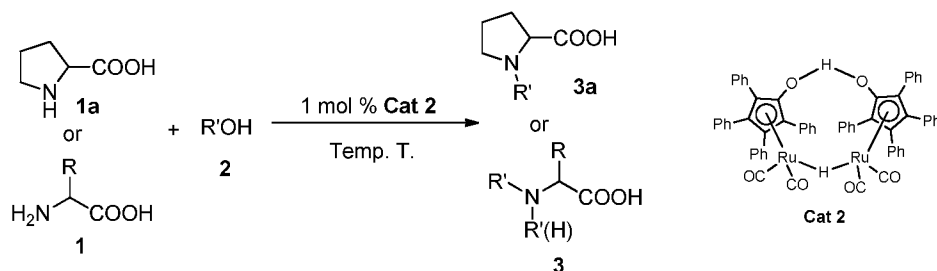
Scheme 2. *N*-alkylation of amino acids and *N*-terminus of peptides with various alcohols.



General reaction conditions: General procedure, 0.2 or 0.5 mmol **1, 1 - 4 ml or 0.6 – 2 mmol **2**, neat or **2d**/tol as solvent, 1 mol % **Cat 2**, 18 - 24 h, 90 °C, isolated yields are shown, unless otherwise specified; ^ayields are base crude H MNR and mass balance; ^b100 °C. For details see Table 4 and 5.**

After exploring the reactivity of various α -amino acids with ethanol (2a) and isopropanol (2b), diverse alcohols were tried (Scheme 2, Table 4). As previously shown, MeOH (2c) or CF₃CH₂OH (2d) did not react with 1a, probably because they were not prone to be dehydrogenated under the present reaction conditions, which allowed for the possibility of using 2c or 2d as solvent. For other alcohol substrates such as 1-butanol (2e), cyclopropylmethanol (2f) and 2-chloroethanol (2g), employed in the reaction with 1a, resulted in quantitative yield of 3ea, 55 % of 3af, and 71 % of 3ag, respectively (Table 4, entry 3 - 5). Also, benzyl alcohol (3h) and 4-chlorobenzyl alcohol (3i) were successfully applied to benzylate 1a with good yields of 68 % and 82 % (Table 4, entry 6 and 7). Here it needs to be pointed out that, the chloro-group on 3ag and 3ai would allow the further functionalization of the obtained amino acid derivatives. Subsequently, the use of other amino acids such as phenylalanine (1d) and glycine (1g) were examined. 1d reacted with 1-butanol (2e) to give 84 % yield of 3de (Table 4, entry 8). Surprisingly, instead of forming a 6-membered heterocycle, the reaction of 1d with 1,5-pentane-diol (2k) gave a 35% yield of the di-alkylated 3dk as the major product (Table 4, entry 9). The introduced 2-hydroxyl groups on 3dk can not only be exploited for further functionalization, but also dramatically increase the hydrophilicity of the product.

Next, the functionalization of 1g with various alcohols was further explored. Upon reaction with 3h the di-benzylation product 3gh was obtained in 52 % yield while when secondary alcohol 2-butanol (2l) was employed to alkylate 1g, selective formation of mono-alkylated product 3gl was obtained in quantitative yield (Table 4, entry 10 and 12).

Table 4: Direct *N*-alkylation of free amino acid (**1**) with various alcohols (**2**).

Entry	1 [mmol]	2	Sol. [ml]	T. [h]	Temp. [°C]	Conv. 1 [%] ^a	Sele. 3 [%] ^a
1	1a Proline / 0.2	2c MeOH / 2 ml	/	18	90	<1	3ac /
2	1a Proline / 0.2	2d CF ₃ CH ₂ OH / 2 ml	/	18	90	<1	3ad /
3	1a Proline / 0.2	2e <i>n</i> BuOH / 2 ml	/	24	90	>99	3ae >99
4	1a Proline / 0.5	2f  / 5 ml	/	18	90	>99	3af (55)
5	1a Proline / 0.2	2g  / 2 ml	/	18	90	>99	3ag (71)
6	1a Proline / 0.2	2h BnOH / 2 ml	/	18	90	>99	3ah (68)
7	1a Proline / 0.5	2i  / 2 mmol	2d /tol. 1/4	20	90	>99	3ai (82)
8	1d Phenylalanine / 0.5	2e <i>n</i> BuOH / 4 ml	2d 2	24	90	>99	3de (84)
9	1d Phenylalanine / 0.5	2j 1-nonanol / 1 ml	2d /tol. 2/2	24	90	>99	3dj (75)
10	1d Phenylalanine / 0.2	2k 1,5-pentanediol / 2 mmol	2d /tol. 2/3	18	100	>99	3dk (35)
12	1g Glycine / 0.5	2h BnOH / 2 mmol	2d /tol. 2/3	18	90	/	3gh (52)
13	1g Glycine / 0.5	2j 1-nonanol / 2 mmol	2d /tol. 2/3	24	90	>99	3gj (91)
14	1g Glycine / 0.2	2l 2-butanol / 4 ml	2d 1	24	90	>99	3gl >99
15	1g Glycine / 0.5	2m 1-pentanol / 2 mmol	2d /tol. 2/3	18	90	>99	3gm (90)
16	1g Glycine / 0.5	2m 1-pentanol / 0.6 mmol	2d /tol. 2/3	18	90	>99	3gm' (46) 3gm (29)
17	1g Glycine / 0.5	2n 1-dodecanol / 2 mmol	2d /tol. 2/3	18	90	>99	3gn (92)

General reaction conditions: General procedure, 0.2 or 0.5 mmol **1**, 0.6 – 2 mmol or 2 - 5 ml **2**, neat or with solvent (amount shown in the table), 1 mol % **Cat 2**, 16 - 24 h, 90 or 100°C, isolated yields in parenthesis; ^aConversion and selectivities are based on crude H NMR.

After this, the possibility of mono-alkylation with primary alcohols was also investigated. Using 4 eq. of 1-pentanol (1m), 1g was successfully di-alkylated giving 90 % of 3gm. When the amount of 1m was decreased to 1.2 eq., 46 % mono-alkylated product 3gm' was obtained, only giving 29 % of the di-alkylated product 3gm (Table 4, entry 15 and 16). As 3gm' is more basic than 3g, without enough steric hindrance on the nitrogen atom, the second alkylation step occurred immediately once 3gm' formed. Thus, mono-*N*-alkylated amino acids can be obtained from primary amino acids and primary alcohols, but with low selectivity.

Long-chain *N*-alkylated amino acids are widely used in surfactants, among which *N*-alkylated amino acids are not well studied because they are relatively difficult to be synthesized⁹. Envisioning the possibility of easily synthesizing long-chain *N*-alkylated amino acids with our methodology, 1-nonanol (2j) was selected to react with 1d and 1g. Di-alkylated compounds 3dj and 3gj were selectively obtained with good to excellent yields of 75 % and 91 %, respectively (Table 4, entry 9 and 13). The reaction also readily proceeded with 1-dodecanol (1n) and 1g as substrate, obtaining the corresponding product 3gn in an excellent yield of 92 % (Table 4, entry 17).

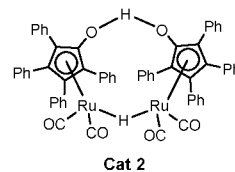
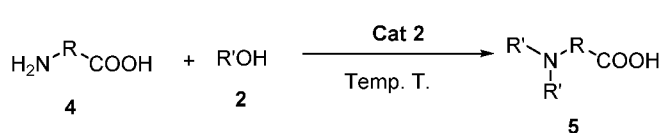
EXAMPLE 4: Direct *N*-alkylation of oligopeptides with alcohol.

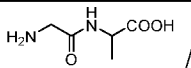
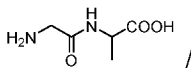
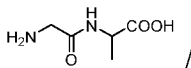
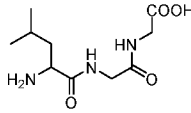
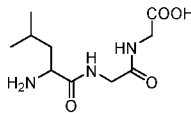
Encouraged by the promising results with the functionalization of amino acids, we attempted the extension of this novel method to simple free oligopeptides, which possess similar physical properties and functional groups as the corresponding amino acids (Scheme 2, Table 5). First, the simple dipeptide glycylalanine (4a) was chosen to react with 2a.

Surprisingly, the corresponding di-ethylated product 5aa was obtained quantitatively (Table 5, entry 1). Based on this protocol, the hydrophobicity or hydrophobicity of peptides can be modulated by introducing either longer

chain alkyl groups or more polar moieties such as hydroxyl groups. Indeed, when 1-dodecanol (2n) was reacted with dipeptide 4a, the corresponding dialkylated product 5an was obtained in 62 % yield (Table 5, entry 3). This type of lipophilic dipeptide can be used for transporting metal ions across
5 membranes. On the other hand, the reaction of diol 2k with 4a, afforded 36 % 5ak, bearing extra hydroxyl groups (Table 5, entry 2).

Following the successful and diverse functionalization of a dipeptide, a tripeptide leucylglycylglycine (4b) was tested in the reaction with 2a, which
10 resulted in the formation of the corresponding di-ethylated product 5ba in 67% yield (Table 5, entry 4 and 5). The above reactions represent the first example of the selective di-alkylation of unprotected oligopeptide substrates on their N-terminus using simple alcohols, allowing for good product yields and easy purification. This methodology can promote the protein N-
15 terminus modification that affecting protein activation, conversion and degradation, further diversifying biological functions .

Table 5: Direct *N*-alkylation of free peptide (**4**) with alcohols (**2**).

Entry	4 [mmol]	2	Sol. [ml]	T. [h]	Temp. [°C]	Conv. 4 [%] ^a	Sele. 5 [%] ^a
1	4a  / 0.2	2a ethanol / 5 ml	/	24	90	>95	5aa >95 (95)
2	4a  / 0.5	2k 1,5-pentanediol / 2 mmol	2d /tol 2/3	24	90	n.d.	5ak (36)
3	4a  / 0.5	2n 1-dodecanol / 2 mmol	2d /tol 2/3	18	90	n.d.	5an (92)
4	4b  / 0.2	2a ethanol / 5 ml	/	24	90	n.d.	5ba (50)
5	4b  / 0.5	2a ethanol / 5 ml	2d 2	18	90	n.d.	5ba (67)

General reaction conditions: General procedure, 0.2 or 0.5 mmol **4, 2**

5 mmol or 5 ml **2**, neat or with solvent (amount shown in the table), 1 mol %

Cat 2, 18 or 24 h, 90°C, isolated yields in parenthesis; ^aConversion and selectivities are based on crude H NMR.

EXAMPLE 5: Synthesis of surfactants from natural amino acids and fatty alcohols.

Realizing that chirality is not an essential property for surfactant, the Fe-based Cat 1b was employed for direct synthesis of surfactants from natural amino acids with fatty alcohols (Figure 1; Table 6).

General reaction conditions: General procedure, 0.5 mmol **1**, 1 ml or 2 mmol **2**, neat or 3 ml **2d** as solvent, 5 mol % Cat 1b, 24 h, 110°C, isolated yields are shown; ^aThe products were isolated as corresponding methyl ester. For details see Table 6.

Glycine (**1g**) and 1-dodecanol (**2n**) were tried under 110 °C for 24 h with 5 mol % Cat 1b. Surprisingly, 54 % mono-N-dodecylglycine (**3gn'**) and 8 % N,N-didodecylglycine (**3gn**) were isolated (Table 6, entry 1). The inherent property of Cat 1b leads to preferentially formation of mono-N-alkyl amino acids, which are one typical type of amino acid based surfactant⁹. After adding KOH and H₂O to **3gn'**, a rich foam formation was clearly seen (Figure 1), indicating its amphiphilic property.

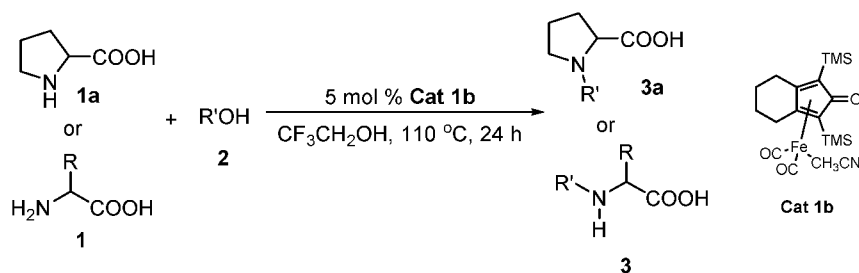
20

Table 6: Direct synthesis of amphiphiles through iron catalyzed *N*-alkylation of free amino acid (**1**) with fatty alcohols (**2**).

Entry	1 [mmol]	2	yield 3 [%]
1	1g Glycine / 0.5	2n 1-dodecanol /	3gn' 54

Glycine (**1g**) and 1-dodecanol (**2n**) were tried under 110 °C for 24 h with 5 mol % Cat **1b**. Surprisingly, 54 % mono-N-dodecylglycine (**3gn'**) and 8 % N,N-didodecylglycine (**3gn**) were isolated (Table 6, entry 1). The inherent property of Cat **1b** leads to preferential formation of mono-N-alkyl amino acids, which are one typical type of amino acid based surfactant⁹. After adding KOH and H₂O to **3gn'**, a rich foam formation was clearly seen (Scheme 3), indicating its amphiphilic property.

Table 6: Direct synthesis of amphiphiles through iron catalyzed *N*-alkylation of free amino acid (**1**) with fatty alcohols (**2**).



Entry	1 [mmol]	2	yield 3 [%]
1	1g Glycine / 0.5	2n 1-dodecanol / 1 ml	3gn' 54 3gn 8
2	1g Glycine / 0.5	2j 1-nonanol / 1 ml	3gj' 51
3	1g Glycine / 0.5	2o 1-decanol / 1 ml	3go' 69
4 ^a	1g Glycine / 0.5	2p 1-tetradecanol / 2 mmol	3gp' 32
4 ^a	1g Glycine / 0.5	2q 1-hexadecanol / 2 mmol	3gq' 38
5 ^a	1g Glycine / 0.5	2r 1-octadecanol / 2 mmol	3gr' 39

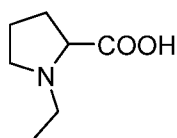
6	1g	Alanine / 0.5	2n	1-dodecanol / 1 ml	3gn'	49
7	1a	Proline / 0.5	2j	1-nonanol / 1 ml	3aj	52

General reaction conditions: General procedure (see page S2), 0.5 mmol **1**, 1 ml or 2 mmol **2**, 3 ml CF₃CH₂OH, 5 mol % **Cat 2b**, 24 h, 110°C, isolated yields are shown; ^aThe products were isolated as corresponding methyl ester.

- 5 Later on, various biomass derived fatty alcohols including 1-nonanol (**2j**), 1-decanol (**2o**), 1-tetradecanol (**2p**), 1-hexadecanol (**2q**) and 1-octadecanol (**2n**) were reacted with **1g**, and 32 – 69 % yields of corresponding mono-N-alkyl glycine **3gj'**, **3go'**, **3gp'**, **3gq'** and **3gr'** were isolated (Table 6, entry 2 - 5). Alanine (**1f**) and proline (**1a**) were reacted with 1-dodecanol (**2n**) and 1-
- 10 nonanol (**2j**), with 5 mol % **Cat 1b**, 49 % of **3fn'** and 52 % of **3aj** were isolated respectively (Table 6, entry 6 and 7). Long-chain alcohols can be easily produced from natural fats and oils. The present findings open up important possibilities for the fully sustainable production of long-chain N-alkyl amino acid based surfactants completely from biomass, with iron based catalyst.

15

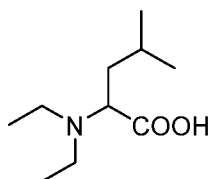
Spectral data of isolated compounds



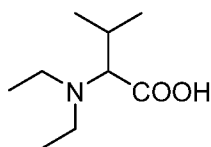
N-ethyl-proline (3aa): Synthesized according to **General procedure**.

- 20 Quantitative yield has been obtained according to crude H NMR. ¹H NMR (400 MHz, D₂O) δ 3.83 – 3.96 (m, 1H), 3.64 – 3.78 (m, 1H), 3.15 – 3.33 (m, 2H), 3.05 – 3.15 (m, 1H), 2.35 – 2.52 (m, 1H), 1.99 – 2.16 (m, 2H), 1.85 – 1.99 (m, 1H), 1.27 (t, J = 7.28 Hz, 3H). ¹³C NMR (100 MHz, MeOD) δ 173.46, 70.13, 55.45, 51.51, 30.32, 24.34, 11.31. HRMS (APCI+, m/z): calculated for

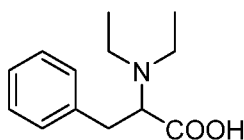
$C_7H_{14}NO_2$ $[M+H]^+$: 144.10191; found: 144.10181. The physical data matches previous report.



- 5 ***N,N*-di-ethyl-leucine (3ba)**: Synthesized according to **General procedure**. Quantitative yield has been obtained according to crude H NMR. 1H NMR (400 MHz, D_2O) δ 3.58 – 3.67 (m, 1H), 3.08 – 3.30 (m, 4H), 1.67 – 1.78 (m, 1H), 1.51 – 1.67 (m, 2H), 1.26 (t, J = 5.24 Hz, 6H), 0.83 – 1.00 (m, 6H). ^{13}C NMR (100 MHz, D_2O) δ 176.34, 68.56, 48.14 (br.s), 38.83, 27.77, 25.52, 23.04, 11.13 (br.s). HRMS (APCI+, m/z): calculated for $C_{10}H_{20}NO_2$ $[M-H]^-$: 186.14886; found: 186.15012.



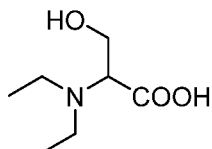
- N,N*-di-ethyl-valine (3ca)**: Synthesized according to **General procedure**. Quantitative yield has been obtained according to crude H NMR. 1H NMR (400 MHz, D_2O) δ 3.45 – 3.52 (m, 1H) 3.05 – 3.35 (m, 4H), 2.22 – 2.40 (m, 1H), 1.15 – 1.40 (m, 6H), 0.99 – 1.10 (m, 3H), 0.86 – 0.99 (m, 3H) ^{13}C NMR (100 MHz, D_2O) δ 174.31, 74.85, 49.85, 45.30, 28.02, 22.22, 18.81, 11.73, 9.62. HRMS (APCI+, m/z): calculated for $C_9H_{20}NO_2$ $[M+H]^+$: 174.14886; found: 174.14879.



***N,N*-di-ethyl-phenylalanine (3da)**: Synthesized according to **General procedure**. Quantitative yield has been obtained according to crude H

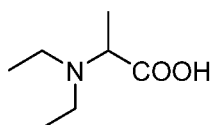
NMR. ^1H NMR (400 MHz, D_2O) δ 7.12 – 7.43 (m, 5H), 3.80 – 3.93 (m, 1H), 3.08 – 3.40 (m, 4H), 2.98 – 3.31 (m, 2H), 1.17 – 1.33 (m, 6H). ^{13}C NMR (100 MHz, D_2O) δ 172.37, 135.44, 129.06, 128.80, 127.28, 67.81, 45.70 (br.s), 33.44, 8.42 (br.s). HRMS (APCI+, m/z): calculated for $\text{C}_{13}\text{H}_{18}\text{NO}_2$ $[\text{M}-\text{H}]^-$:

5 220.13321; found: 220.13433.

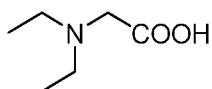


***N,N*-di-ethyl-serine (3ea):** Synthesized according to **General procedure**.

Quantitative yield has been obtained according to crude H NMR. ^1H NMR (400 MHz, D_2O) δ 3.98 – 4.18 (m, 2H), 3.79 – 3.91 (m, 1H), 3.37 – 3.52 (m, 1H), 3.18 – 3.37 (m, 3H), 1.18 – 1.40 (m, 6H). ^{13}C NMR (100 MHz, D_2O) δ 173.91, 69.38, 60.55, 50.29, 47.45, 11.81, 10.96. HRMS (APCI+, m/z): calculated for $\text{C}_7\text{H}_{14}\text{NO}_3$ $[\text{M}-\text{H}]^-$: 160.09682; found: 160.09811.



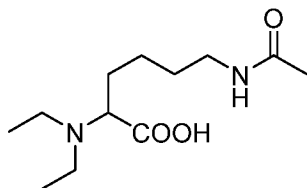
15 ***N,N*-di-ethyl-alanine (3ea):** Synthesized according to **General procedure**. Quantitative yield has been obtained according to crude H NMR. ^1H NMR (400 MHz, D_2O) δ 3.77 – 3.92 (m, 1H), 3.19 – 3.39 (m, 2H), 3.00 – 3.20 (m, 2H), 1.38 – 1.52 (m, 3H), 1.17 – 1.37 (m, 6H). ^{13}C NMR (100 MHz, D_2O) δ 174.31, 61.46, 47.04, 44.97, 11.47, 9.29, 8.32. HRMS (APCI+, m/z): calculated for $\text{C}_7\text{H}_{16}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 146.11756; found: 146.11746.



***N,N*-di-ethyl-glycine (3ga):** Synthesized according to **General**

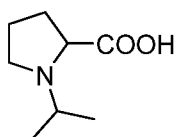
25 **procedure**. Quantitative yield has been obtained according to crude H NMR. ^1H NMR (400 MHz, D_2O) δ 3.66 (s, 2H), 3.22 (q, $J = 7.31$ Hz, 4H), 1.26 (t, $J = 7.32$ Hz, 6H). ^{13}C NMR (100 MHz, D_2O) δ 173.50, 57.36, 52.14, 52.11,

11.13. HRMS (APCI+, m/z): calculated for C₆H₁₄NO₂ [M+H]⁺: 132.10191; found: 132.10180.

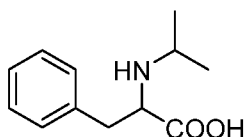


N⁶-acetyl-N²,N²-di-ethyl-lysine (3ha): Synthesized according to **General**

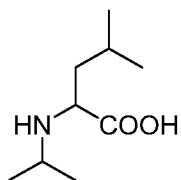
- 5 **procedure.** N⁶-acetyl-lysine (0.094 g, 0.50 mmol) affords **3ha** (0.090 g, 75% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 80:20 to 50:50).. ¹H NMR (400 MHz, D₂O) δ 4.02 – 4.13 (m, 1H), 2.96 – 3.08 (m, 4H), 2.82 – 2.96 (m, 2H), 1.97 (s, 3H), 1.70 – 1.84 (m, 1H), 1.51 – 1.70 (m, 3H), 1.25 – 1.38 (m, 2H), 1.07 – 1.24 (m, 6H). ¹³C NMR (100 MHz, D₂O) δ 181.63, 176.06, 57.47, 54.04, 49.59, 33.75, 25.86, 25.15, 24.49, 11.12. HRMS (APCI+, m/z): calculated for C₁₂H₂₅N₂O₃ [M+H]⁺: 245.18597; found: 245.18593.



- 15 **N-isopropyl-proline (3ab):** Synthesized according to **General procedure.** Quantitative yield has been obtained according to crude H NMR. ¹H NMR (400 MHz, D₂O) δ 3.93 – 4.05 (m, 1H), 3.62 – 3.72 (m, 1H), 3.47 – 3.61 (m, 1H), 3.13 – 3.26 (m, 1H), 2.29 – 2.44 (m, 1H), 2.00 – 2.15 (m, 2H), 1.79 – 1.97 (m, 1H), 1.20 – 1.40 (m, 6H). ¹³C NMR (100 MHz, D₂O) δ 174.53, 65.80, 56.99, 52.28, 29.59, 23.59, 17.39, 17.26. HRMS (APCI+, m/z): calculated for C₈H₁₆NO₂ [M+H]⁺: 158.11756; found: 158.11747.

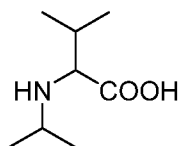


N-isopropyl-phenylalanine (3db): Synthesized according to **General procedure**. Quantitative yield has been obtained according to crude H NMR. ^1H NMR (400 MHz, D_2O) δ 6.87 – 7.24 (m, 5H), 3.24 – 3.34 (m, 1H), 2.71 – 2.80 (m, 1H), 2.45 – 2.65 (m, 2H), 0.74 – 0.93 (m, 6H). ^{13}C NMR (100 MHz, D_2O) δ 181.45, 137.84, 129.08, 128.33, 126.40, 62.55, 45.96, 39.09, 22.57, 19.72. HRMS (APCI+, m/z): calculated for $\text{C}_{12}\text{H}_{18}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 208.13321; found: 208.13312.



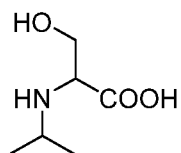
N-isopropyl-leucine (3bb): Synthesized according to **General**

procedure. Quantitative yield has been obtained according to crude H NMR. ^1H NMR (400 MHz, D_2O -NaOH) δ 3.08 – 3.21 (m, 1H), 2.52 – 2.66 (m, 1H), 1.37 – 1.54 (m, 1H), 1.16 – 1.35 (m, 2H), 0.86 – 1.05 (m, 6H), 0.68 – 0.86 (m, 6H). ^{13}C NMR (100 MHz, D_2O) δ 177.15, 61.65, 53.02, 42.23, 27.09, 24.64, 23.83, 21.72, 20.32. HRMS (APCI+, m/z): calculated for $\text{C}_9\text{H}_{20}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 174.14886; found: 174.14872.



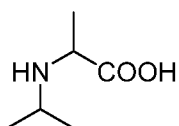
N-isopropyl-valine (3cb): Synthesized according to **General procedure**.

Quantitative yield has been obtained according to crude H NMR. ^1H NMR (400 MHz, D_2O) δ 3.09 – 3.20 (m, 1H), 2.82 – 2.96 (m, 1H), 1.81 – 2.00 (m, 1H), 0.99 – 1.25 (m, 6H), 0.85 – 0.98 (m, 6H). ^{13}C NMR (100 MHz, D_2O) δ 181.45, 69.24, 51.18, 32.88, 23.89, 21.59, 21.17, 20.60. HRMS (APCI+, m/z): calculated for $\text{C}_8\text{H}_{18}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 160.13321; found: 160.13307.



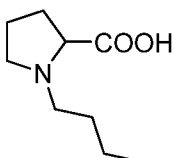
N-isopropyl-serine (3eb): Synthesized according to **General procedure**.

Quantitative yield has been obtained according to crude H NMR. ¹H NMR (400 MHz, D₂O) δ 3.85 – 4.01 (m, 2H), 3.72 – 3.79 (m, 1H), 3.38 – 3.52 (m, 1H), 1.24 – 1.36 (m, 6H). ¹³C NMR (100 MHz, D₂O) δ 174.36, 63.67, 62.42, 53.03, 21.26, 21.24, 20.58. HRMS (APCI+, m/z): calculated for C₆H₁₄NO₃ [M+H]⁺: 148.09682; found: 148.09671.



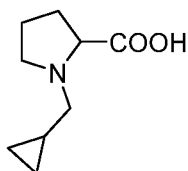
N-isopropyl-alanine (3fb): Synthesized according to **General procedure**.

Quantitative yield has been obtained according to crude H NMR. ¹H NMR (400 MHz, D₂O) δ 3.71 – 3.80 (m, 1H), 3.38 – 3.50 (m, 1H), 1.42 – 1.53 (m, 3H), 1.26 – 1.37 (m, 6H). ¹³C NMR (100 MHz, D₂O) δ 177.76, 57.76, 52.17, 21.17, 20.71, 18.14. HRMS (APCI+, m/z): calculated for C₆H₁₄NO₂ [M+H]⁺: 132.10191; found: 132.10181.



N-n-butyl-proline (3ae): Synthesized according to **General procedure**.

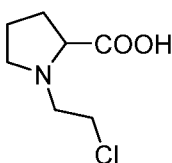
Quantitative yield has been obtained according to crude H NMR. ¹H NMR (400 MHz, MeOD) δ 3.78 – 3.90 (m, 1H), 3.65 – 3.78 (m, 1H), 3.15 – 3.27 (m, 1H), 2.95 – 3.15 (m, 2H), 2.30 – 2.46 (m, 1H), 1.99 – 2.15 (m, 2H), 1.82 – 1.98 (m, 1H), 1.56 – 1.76 (m, 2H), 1.30 – 1.47 (m, 2H), 0.83 – 1.04 (m, 3H). ¹³C NMR (100 MHz, MeOD) δ 173.50, 70.60, 56.38, 55.95, 30.30, 28.94, 24.34, 20.88, 13.96. HRMS (APCI+, m/z): calculated for C₉H₁₆NO₂ [M-H]⁻: 170.11756; found: 170.11876.



N-cyclopropylmethyl-proline (3fb): Synthesized according to **General procedure**. Proline (0.058 g, 0.50 mmol) affords **3fb** (0.046 g, 55% yield).

White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH

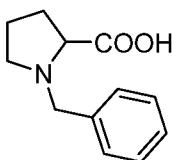
5 80:20 to 80:20). ¹H NMR (400 MHz, CDCl₃) δ 3.83 – 4.01 (m, 1H), 3.67 – 3.83 (m, 1H), 2.98 – 3.14 (m, 1H), 2.70 – 2.95 (m, 2H), 2.22 – 2.38 (m, 1H), 2.05 – 2.22 (m, 1H), 1.86 – 2.05 (m, 2H), 0.95 – 1.09 (m, 1H), 0.49 – 0.70 (m, 2H), 0.18 – 0.40 (m, 2H). ¹³C NMR (100 MHz, CDCl₃) δ 170.64, 68.63, 59.14, 53.96, 29.01, 23.06, 6.63, 4.45, 4.06. HRMS (APCI+, m/z): calculated for C₉H₁₄NO₂ [M-H]⁻: 168.10191; found: 168.10321.



N-(2-chloroethyl)-proline (3ag): Synthesized according to **General procedure**. Proline (0.058 g, 0.50 mmol) affords **3ag** (0.063 g, 71% yield).

15 White solid was obtained after crystallization in MeOH/Et₂O. ¹H NMR (400 MHz, D₂O) δ 4.45 – 4.65 (m, 3H), 3.76 – 3.93 (m, 2H), 3.33 – 3.54 (m, 2H), 2.38 – 2.55 (m, 1H), 2.15 – 2.32 (m, 1H), 1.97 – 2.14 (m, 2H). ¹³C NMR (100 MHz, D₂O) δ 169.50, 66.38, 59.35, 46.16, 41.59, 28.06, 23.07. HRMS (APCI+, m/z): calculated for C₇H₁₃ClNO₂ [M+H]⁺: 178.06293; found: 178.06289.

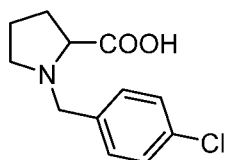
20



N-benzyl-proline (3ah): Synthesized according to **General procedure**.

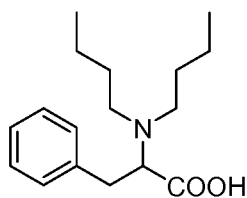
Proline (0.019 g, 0.20 mmol) affords **3ah** (0.028 g, 68% yield). White solid

was obtained after crystallization in Et₂O. ¹H NMR (400 MHz, CDCl₃) δ 9.33 (br.s, 1H), 7.28 – 7.52 (m, 5H), 4.11 – 4.37 (m, 2H), 3.74 – 3.90 (m, 1H), 3.59 – 3.74 (m, 1H), 2.78 – 2.94 (m, 1H), 2.17 – 2.38 (m, 2H), 1.79 – 2.08 (m, 2H). ¹³C NMR (100 MHz, CDCl₃) δ 171.02, 130.73, 130.53, 129.40, 129.04, 67.31, 57.61, 53.31, 28.89, 22.89. HRMS (APCI+, m/z): calculated for C₁₂H₁₆NO₂ [M+H]⁺: 206.11756; found: 206.11742.



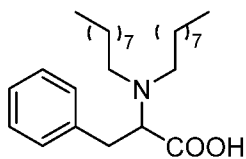
N-(4-chloro-benzyl)-proline (3ai): Synthesized according to **General**

procedure. Proline (0.058 g, 0.50 mmol) affords **3ai** (0.098 g, 82% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 90:10 to 50:50). ¹H NMR (400 MHz, D₂O) δ 7.28 – 7.48 (m, 4H), 4.12 – 4.34 (m, 2H), 3.75 – 3.88 (m, 1H), 3.41 – 3.54 (m, 1H), 3.03 – 3.18 (m, 1H), 2.31 – 2.48 (m, 1H), 1.80 – 2.09 (m, 3H). ¹³C NMR (100 MHz, D₂O) δ 176.56, 137.60, 134.51, 131.87, 131.63, 70.76, 59.87, 56.86, 31.30, 25.17. HRMS (APCI+, m/z): calculated for C₁₂H₁₅ClNO₂ [M+H]⁺: 240.07858; found: 240.07854.



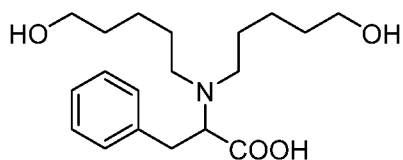
N,N-(di-*n*-butyl)-phenylalanine (3de): Synthesized according to **General procedure.** Phenylalanine (0.083 g, 0.50 mmol) affords **3de** (0.116 g, 84% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 90:10 to 80:20). ¹H NMR (400 MHz, CDCl₃) δ 9.36 (br.s, 1H), 7.10 – 7.35 (m, 5H), 3.83 – 3.94 (m, 1H), 3.42 – 3.56 (m, 1H), 2.88 – 3.07 (m, 3H), 2.70 – 2.85 (m, 2H), 1.52 – 1.67 (m, 2H), 1.36 – 1.52 (m, 2H), 1.05 – 1.27

(m, 4H). ^{13}C NMR (100 MHz, CDCl_3) δ 170.49, 137.88, 128.76, 128.55, 126.65, 67.49, 51.45, 33.82, 26.82, 19.90, 13.49. HRMS (APCI+, m/z): calculated for $\text{C}_{17}\text{H}_{26}\text{NO}_2$ $[\text{M}-\text{H}]^-$: 276.19581; found: 276.19703.

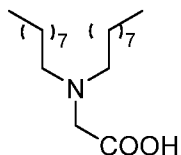


5 ***N,N*-(di-*n*-nonyl)-phenylalanine (3dj)**: Synthesized according to **General procedure**. Phenylalanine (0.083 g, 0.50 mmol) affords **3dj** (0.147 g, 75% yield). White solid was obtained after column chromatography (SiO_2 , Pentane/EtOAc 50:50 to 0:100, then EtOH/MeOH 90/10). ^1H NMR (400 MHz, CDCl_3) δ 8.68 (br.s, 1H), 7.13 – 7.31 (m, 5H). 3.84 – 3.93 (m, 1H), 3.50
10 – 3.58 (m, 1H), 2.88 – 3.05 (m, 3H), 2.67 – 2.79 (m, 2H), 1.53 – 1.67 (m, 2H), 1.38 – 1.52 (m, 2H), 0.98 – 1.35 (m, 24H), 0.75 – 0.89 (m, 6H). ^{13}C NMR (100 MHz, CDCl_3) δ 170.49, 138.02, 128.75, 128.59, 126.67, 67.41, 51.76, 33.71, 31.67, 29.30, 29.05, 29.03, 26.70, 25.07, 22.50, 13.94. HRMS (APCI+, m/z): calculated for $\text{C}_{27}\text{H}_{48}\text{NO}_2$ $[\text{M}-\text{H}]^-$: 418.36796; found: 418.36763.

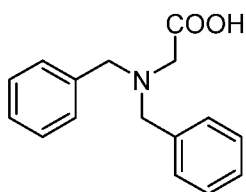
15



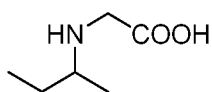
***N,N*-di-(5-hydroxypentyl)-phenylalanine (3dk)**: Synthesized according to **General procedure**. Phenylalanine (0.083 g, 0.50 mmol) affords **3dk**
20 (0.059 g, 35% yield). White solid was obtained after column chromatography (SiO_2 , EtOAc/MeOH 50:50 to 30:70). ^1H NMR (400 MHz, D_2O) δ 7.23 – 7.45 (m, 5H), 3.91 – 4.03 (m, 1H), 3.47 – 3.65 (m, 4H), 3.13 – 3.32 (m, 4H), 2.95 – 3.13 (m, 2H), 1.58 – 1.79 (m, 4H), 1.45 – 1.58 (m, 4H), 1.21 – 1.43 (m, 4H). ^{13}C NMR (100 MHz, D_2O) δ 172.40, 135.73, 128.94, 128.85, 127.27, 68.29,
25 61.15, 33.38, 30.56, 23.33, 22.13. HRMS (APCI+, m/z): calculated for $\text{C}_{19}\text{H}_{32}\text{NO}_4$ $[\text{M}+\text{H}]^+$: 338.23258; found: 338.23176.



N,N-di-*n*-nonyl-glycine (3gj): Synthesized according to **General procedure**. Glycine (0.038 g, 0.50 mmol) affords **3gj** (0.149 g, 91% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 5 90:10 to 70:30). ¹H NMR (400 MHz, CDCl₃) δ 8.78 (br.s, 1H), 3.48 (s, 2H), 2.95 – 3.15 (m, 4H), 1.52 – 1.75 (m, 4H), 1.03 – 1.42 (m, 24H), 0.65 – 0.95 (m, 6H). ¹³C NMR (100 MHz, CDCl₃) δ 167.42, 55.99, 53.92, 31.63, 29.26, 29.03, 28.99, 26.66, 23.71, 22.47, 13.90. HRMS (APCI+, m/z): calculated for C₂₀H₄₄NO₂ [M+H]⁺: 328.32155; found: 328.32095.

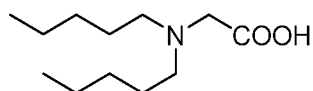


N,N-di-benzyl-glycine (3gh): Synthesized according to **General procedure**. Glycine (0.038 g, 0.50 mmol) affords **3gh** (0.066 g, 52% yield). White solid was obtained after precipitation in MeOH/Et₂O. ¹H NMR (400 15 MHz, DMSO-d₆) δ 7.20 – 7.40 (m, 10H), 3.73 (s, 4H), 3.15 (s, 2H). ¹³C NMR (100 MHz, DMSO-d₆) δ 172.19, 139.00, 128.53, 128.24, 126.98, 56.79, 53.00. HRMS (APCI+, m/z): calculated for C₁₆H₁₈NO₂ [M+H]⁺: 256.13321; found: 256.13325.

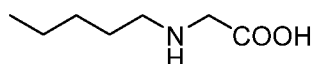


N-2-butyl-glycine (3gl): Synthesized according to **General procedure**. Quantitative yield has been obtained according to crude H NMR. ¹H NMR (400 MHz, D₂O) δ 3.48 – 3.67 (m, 2H), 3.14 – 3.27 (m, 1H), 1.68 – 1.84 (m, 1H), 1.47 – 1.66 (m, 1H), 1.22 – 1.35 (m, 3H). 0.88 – 1.03 (m, 3H). ¹³C NMR

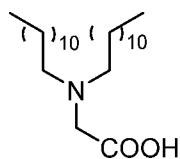
(100 MHz, D₂O) δ 174.11, 58.44, 49.00, 28.27, 17.51, 11.47. HRMS (APCI+, m/z): calculated for C₆H₁₄NO₂ [M+H]⁺: 132.10191; found: 132.10181.



- 5 ***N,N*-di-(*n*-pentyl)-glycine (3gm):** Synthesized according to **General procedure**. Glycine (0.038 g, 0.50 mmol) affords **3gm** (0.097 g, 90% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 90:10 to 70:30). ¹H NMR (400 MHz, D₂O) δ 3.68 (s, 2H), 3.05 – 3.25 (m, 4H), 1.56 – 1.80 (m, 4H), 1.20 – 1.40 (m, 8H), 0.75 – 1.00 (m, 6H). ¹³C NMR (100 MHz, D₂O) δ 173.36, 58.39, 57.49, 30.46, 25.61, 24.08, 15.65. HRMS (APCI+, m/z): calculated for C₁₂H₂₆NO₂ [M+H]⁺: 216.19581; found: 216.19574.

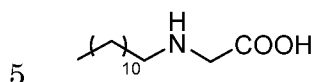


- N-n*-pentyl-glycine (3gm'):** Synthesized according to **General procedure**. Glycine (0.038 g, 0.50 mmol) affords **3gm'** (0.033 g, 46% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 60:40 to 30:70). ¹H NMR (400 MHz, D₂O) δ 3.57 (s, 2H), 2.91 – 3.13 (m, 2H), 1.54 – 1.80 (m, 2H), 1.20 – 1.46 (m, 4H), 0.75 – 0.99 (m, 3H). ¹³C NMR (100 MHz, D₂O) δ 174.07, 51.74, 50.13, 30.40, 27.75, 24.05, 15.62. HRMS (APCI+, m/z): calculated for C₇H₁₆NO₂ [M+H]⁺: 146.11756; found: 146.11751.



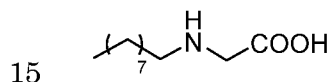
- N,N*-di-*n*-dodecyl-glycine (3gn):** Synthesized according to **General procedure**. Glycine (0.038 g, 0.50 mmol) affords **3gn** (0.189 g, 92% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 90:10 to 70:30). ¹H NMR (400 MHz, CDCl₃) δ 8.36 (br.s, 1H), 3.49 (s, 2H), 2.95 – 3.15 (m, 4H), 1.52 – 1.75 (m, 4H), 1.03 – 1.42 (m, 36H), 0.75 – 0.95 (m,

6H). ^{13}C NMR (100 MHz, CDCl_3) δ 167.51, 56.23, 54.01, 31.80, 29.51, 29.42, 29.38, 29.23, 29.11, 26.73, 23.74, 22.57, 13.99. HRMS (APCI+, m/z): calculated for $\text{C}_{26}\text{H}_{54}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 412.41491; found: 412.41482.



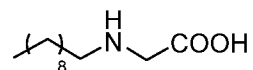
N-dodecyl-glycine (3gn'): Synthesized according to **General procedure**.

Glycine (0.038 g, 0.50 mmol) affords **3gn'** (0.066 g, 54% yield). White solid was obtained after column chromatography (SiO_2 , EtOAc/MeOH 70:30 to 40:60). ^1H NMR (400 MHz, KOH, D_2O) δ 3.00 – 3.20 (m, 2H), 2.38 – 2.58 (m, 2H), 1.36 – 1.57 (m, 2H), 1.12 – 1.35 (m, 18H), 0.73 – 0.90 (m, 3H). ^{13}C NMR (100 MHz, KOH, D_2O) δ 180.96, 55.23, 51.69, 34.63, 32.69, 32.61, 32.54, 32.40, 32.22, 31.93, 30.15, 25.28, 16.42. HRMS (APCI+, m/z): calculated for $\text{C}_{14}\text{H}_{30}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 244.22711; found: 244.22711.



N-nonyl-glycine (3gj'): Synthesized according to **General procedure**.

Glycine (0.038 g, 0.50 mmol) affords **3gj'** (0.052 g, 51% yield). White solid was obtained after column chromatography (SiO_2 , EtOAc/MeOH 70:30 to 40:60). ^1H NMR (400 MHz, KOH, D_2O) δ 2.90 – 3.18 (m, 2H), 2.30 – 2.56 (m, 2H), 1.28 – 1.55 (m, 2H), 1.02 – 1.38 (m, 12H), 0.65 – 0.91 (m, 3H). ^{13}C NMR (100 MHz, KOH, D_2O) δ 181.71, 55.18, 51.50, 34.34, 32.09, 31.96, 31.80, 31.74, 29.78, 25.05, 16.31. HRMS (APCI+, m/z): calculated for $\text{C}_{11}\text{H}_{24}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 202.18016; found: 202.18010.

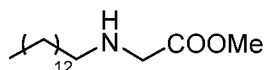


N-decyl-glycine (3go'): Synthesized according to **General procedure**.

Glycine (0.038 g, 0.50 mmol) affords **3go'** (0.075 g, 69% yield). White solid was obtained after column chromatography (SiO_2 , EtOAc/MeOH 70:30 to 50:50). ^1H NMR (400 MHz, KOH, D_2O) δ 3.00 – 3.15 (m, 2H), 2.36 – 2.58 (m,

2H), 1.36 – 1.54 (m, 2H), 1.05 – 1.35 (m, 14H), 0.73 – 0.90 (m, 3H). ^{13}C NMR (100 MHz, KOH, D_2O) δ 181.42, 55.34, 51.67, 34.49, 32.45, 32.30, 32.22, 32.05, 31.99, 30.04, 25.18, 16.35. HRMS (APCI+, m/z): calculated for $\text{C}_{12}\text{H}_{26}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 216.19581; found: 216.19589.

5

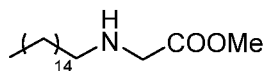


methyl *N*-tetradecylglycinate (methyl ester of **3gp'**: **3gp'-Me**):

Synthesized according to **General procedure** and **Esterification**

procedure. Glycine (0.038 g, 0.50 mmol) affords **3gp'-Me** (0.046 g, 32%

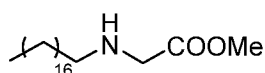
10 yield). Pale oily compound was obtained after column chromatography. ^1H NMR (400 MHz, CDCl_3) δ 3.72 (s, 3H), 3.41 (s, 2H), 2.52 – 2.64 (m, 2H), 1.42 – 1.54 (m, 2H), 1.15 – 1.36 (m, 22H), 0.83 – 0.91 (m, 3H) ^{13}C NMR (100 MHz, CDCl_3) δ 173.02, 51.71, 50.84, 49.67, 31.91, 30.04, 29.68, 29.66, 29.65, 29.64, 29.60, 29.57, 29.51, 29.34, 27.21, 22.68, 14.10. HRMS (APCI+, m/z):
15 calculated for $\text{C}_{17}\text{H}_{36}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 286.27406; found: 286.27430.



methyl *N*-hexadecylglycinate (methyl ester of **3gq'**: **3gq'-Me**):

Synthesized according to **General procedure** and **Esterification**

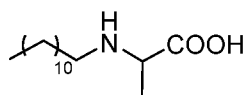
20 **procedure**. Glycine (0.038 g, 0.50 mmol) affords **3gq'-Me** (0.059 g, 38% yield). Pale oily compound was obtained after column chromatography. ^1H NMR (400 MHz, CDCl_3) δ 3.73 (s, 3H), 3.43 (s, 2H), 2.52 – 2.64 (m, 2H), 1.42 – 1.58 (m, 2H), 1.13 – 1.37 (m, 26H), 0.80 – 0.94 (m, 3H) ^{13}C NMR (100 MHz, CDCl_3) δ 172.71, 51.78, 50.65, 49.60, 31.92, 29.85, 29.68, 29.67, 29.65, 29.60,
25 29.57, 29.49, 29.35, 27.19, 22.68, 14.11. HRMS (APCI+, m/z): calculated for $\text{C}_{19}\text{H}_{40}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 314.30536; found: 314.30540.



methyl *N*-octadecylglycinate (methyl ester of **3gr'**: **3gr'-Me**):

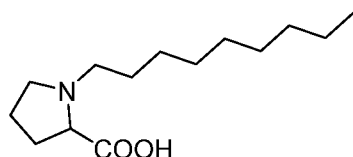
Synthesized according to **General procedure** and **Esterification**

procedure. Glycine (0.038 g, 0.50 mmol) affords **3gr'-Me** (0.067 g, 39%
 5 yield). Pale oily compound was obtained after column chromatography. ¹H
 NMR (400 MHz, CDCl₃) δ 3.71 (s, 3H), 3.40 (s, 2H), 2.53 – 2.64 (m, 2H), 1.42
 – 1.52 (m, 2H), 1.17 – 1.36 (m, 30H), 0.80 – 0.91 (m, 3H) ¹³C NMR (100 MHz,
 CDCl₃) δ 173.01, 51.68, 50.83, 49.67, 31.90, 30.04, 29.67, 29.66, 29.64, 29.59,
 29.57, 29.50, 29.34, 27.21, 22.67, 14.09. HRMS (APCI+, m/z): calculated for
 C₂₁H₄₄NO₂ [M+H]⁺: 342.33666; found: 342.33681.



***N*-dodecyl-alanine (3gn')**: Synthesized according to **General procedure**.

Alanine (0.045 g, 0.50 mmol) affords **3gn'** (0.063 g, 49% yield). White solid
 was obtained after column chromatography (SiO₂, EtOAc/MeOH 70:30 to
 15 40:60). ¹H NMR (400 MHz, KOH, D₂O) δ 3.00 – 3.20 (m, 2H), 2.38 – 2.58 (m,
 2H), 1.36 – 1.57 (m, 2H), 1.12 – 1.35 (m, 18H), 0.73 – 0.90 (m, 3H). ¹³C NMR
 (100 MHz, KOH, D₂O) δ 180.96, 55.23, 51.69, 34.63, 32.69, 32.61, 32.54,
 32.40, 32.22, 31.93, 30.15, 25.28, 16.42. HRMS (APCI+, m/z): calculated for
 C₁₅H₃₂NO₂ [M+H]⁺: 258.24276; found: 258.24302.

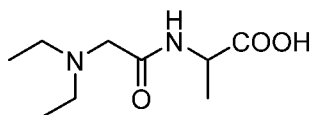


***N*-nonyl-proline (3aj)**: Synthesized according to **General procedure**.

Proline (0.053 g, 0.50 mmol) affords **3aj** (0.063 g, 52% yield). White solid
 was obtained after column chromatography (SiO₂, EtOAc/MeOH 90:10 to
 25 50:50). ¹H NMR (400 MHz, CDCl₃) δ 3.92 – 3.06 (m, 1H), 3.61 – 3.78 (m, 1H),
 3.06 – 3.22 (m, 1H), 2.90 – 3.06 (m, 1H), 2.71 – 2.89 (m, 1H), 2.27 – 2.43 (m,
 1H), 2.13 – 2.27 (m, 1H), 1.88 – 2.09 (m, 2H), 1.60 – 1.79 (m, 2H), 1.05 – 1.43

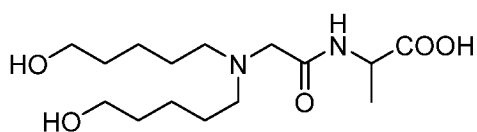
(m, 12H), 0.72 – 0.95 (m, 3H). ^{13}C NMR (100 MHz, CDCl_3) δ 170.24, 69.66, 55.60, 54.77, 31.68, 29.38, 29.27, 29.05, 26.64, 25.66, 23.48, 22.53, 13.97. HRMS (APCI+, m/z): calculated for $\text{C}_{14}\text{H}_{28}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 242.21146; found: 242.21144.

5



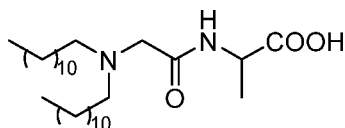
N,N-di-ethyl-glycyl-alanine (5aa): Synthesized according to **General procedure**. Quantitative yield has been obtained according to crude H NMR. ^1H NMR (400 MHz, D_2O) δ 4.08 – 4.22 (m, 1H), 3.88 – 4.03 (m, 2H), 3.17 – 3.31 (m, 4H), 1.30 – 1.36 (m, 3H), 1.21 – 1.30 (m, 6H). ^{13}C NMR (100 MHz, D_2O) δ 179.44, 164.77, 53.19, 51.17, 49.43, 16.92, 8.30. HRMS (APCI+, m/z): calculated for $\text{C}_9\text{H}_{19}\text{N}_2\text{O}_3$ $[\text{M}+\text{H}]^+$: 203.13902; found: 203.13895.

10

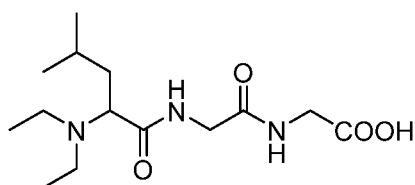


N,N-di-(5-hydroxy-pentyl)-glycyl-alanine (5ak): Synthesized according to **General procedure**. Glycyl-alanine (0.073 g, 0.50 mmol) affords **5ak** (0.036 g, 36% yield). White solid was obtained after column chromatography (SiO_2 , EtOAc/MeOH 60:40 to 30:70). ^1H NMR (400 MHz, D_2O) δ 4.15 – 4.23 (m, 1H), 3.55 – 3.67 (m, 4H), 3.19 – 3.34 (m, 2H), 2.52 – 2.71 (m, 4H), 1.47 – 1.66 (m, 8H), 1.28 – 1.44 (m, 7H). ^{13}C NMR (100 MHz, D_2O) δ 182.21, 175.12, 64.25, 59.67, 57.15, 53.19, 33.76, 28.20, 25.58, 20.44. HRMS (APCI+, m/z): calculated for $\text{C}_{15}\text{H}_{31}\text{N}_2\text{O}_5$ $[\text{M}+\text{H}]^+$: 319.22275; found: 319.22278.

20



***N,N*-di-(*n*-dodecyl)-glycyl-alanine (5an):** Synthesized according to **General procedure**. Glycyl-alanine (0.073 g, 0.50 mmol) affords **5an** (0.149 g, 62% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 90:10 to 50:50). ¹H NMR (400 MHz, CDCl₃) δ 8.20 (br.s, 1H), 4.17 – 4.38 (m, 1H), 3.24 – 3.60 (m, 2H), 2.52 – 2.95 (m, 4H), 1.45 – 1.64 (m, 4H), 1.11 – 1.42 (m, 39H), 0.70 – 0.95 (m, 6H). ¹³C NMR (100 MHz, CDCl₃) δ 177.93, 167.98, 56.32, 54.16, 49.95, 31.87, 29.64, 29.61, 29.58, 29.37, 29.31, 27.15, 25.34, 22.63, 18.39, 14.04. HRMS (APCI+, *m/z*): calculated for C₂₉H₅₉N₂O₃ [M+H]⁺: 483.45202; found: 483.45170.



***N,N*-di-ethyl-glycyl-alanine (5ba):** Synthesized according to **General procedure**. Leucyl-glycyl-glycine (0.123 g, 0.50 mmol) affords **5ba** (0.101 g, 67% yield). White solid was obtained after column chromatography (SiO₂, EtOAc/MeOH 60:40 to 30:70). ¹H NMR (400 MHz, D₂O) δ 3.86 – 4.02 (m, 2H), 3.75 (s, 2H), 3.38 – 3.48 (m, 1H), 2.71 – 2.88 (m, 2H), 2.46 – 2.62 (m, 2H), 1.68 – 1.83 (m, 1H), 1.32 – 1.54 (m, 2H), 0.97 – 1.13 (m, 6H), 0.83 – 0.96 (m, 6H). ¹³C NMR (100 MHz, D₂O) δ 179.02, 177.83, 173.32, 64.90, 46.57, 45.81, 44.84, 40.43, 27.43, 25.57, 23.61, 14.04. HRMS (APCI+, *m/z*): calculated for C₁₄H₂₈N₃O₄ [M+H]⁺: 302.20743; found: 302.20747.

REFERENCES

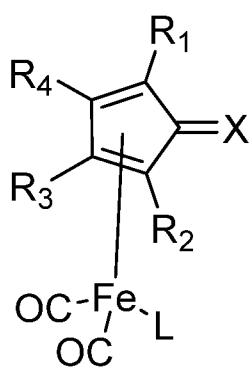
- ¹ Ingrid Wagner, Hans Musso, New Naturally Occurring Amino Acids, *Angew. Chem. Int. Ed. Engl.* **22** (1983) 816-828
- 5 ² G. Xu, X. Fan, A. J. Miller Plant Nitrogen Assimilation and Use Efficiency, *Annu. Rev. Plant Biol.* **2012**, *63*, 153 – 182
- ³ Elinor Scott & Francisc Peter & Johan Sanders Biomass in the manufacture of industrial products—the use of proteins and amino acids, *Appl Microbiol Biotechnol* (2007) *75*:751–762
- 10 ⁴ L. Aurelio, A. B. Hughes, *Amino Acids, Peptides and Proteins in Organic Chemistry. Vol.1 – Origins and Synthesis of Amino Acids, chapter 6, Synthesis of N-Alkyl Amino Acids* **2010**, Wiley-VCH Verlag GmbH & Co. KGaA
- ⁵ Crystal R. Waites, Mark A. Dominick, Thomas P. Sanderson, Beth E. Schilling, *Toxicological sciences*, 2007, *100*, 248–258.
- 15 ⁶ M. C. Cleij, P. Scrimin, P. Tecilla, U. Tonellato *J. Org. Chem.* 1997, *62*, 5592 – 5599.
- ⁷ Pijus K. Sasmal, a Craig N. Streub and Eric Meggers *Chem. Commun.*, 2013, *49*, 1581--1587
- ⁸ D. R. Fandrick et al. (more than 25 authors) *J. Org. Chem.* **2013**, *78*, 3592 – 3615
- 20 ⁹ Yunxiang Li, Krister Holmberg, Romain Bordes *Journal of Colloid and Interface Science* *411* (2013) 47–52 R. Bordes, K. Holmberg *Advances in Colloid and Interface Science* **2015**, *222*, 79 – 91.
- ¹⁰ Santimukul Santra, J. Manuel Perez *Biomacromolecules* 2011, *12*, 3917–3927
- 25 ¹¹ M. Y. Pavlov, R. E. Watts, Z. Tan, V. W. Cornish, M. Ehrenberg, A. C. Forster Slow peptide bond formation by proline and other *N*-alkylamino acids in translation *P.N.A.S.* **2009**, *106*, 50 – 54.
- ¹² A recent example derived protected amino acids through C-H activation, see T. J. Osberger, D. C. Rogness, J. T. Kohrt, A. F. Stepan, M. C. White doi:10.1038/nature18941
- 30 ¹³ Yasufumi Ohfuné 1, Natsuko Kurokawa 1, Naoki Higuchi 2, Masayuki Saito 2, Masaki Hashimoto 2, Takaharu Tanaka 2 *CHEMISTRY LETTERS*, 1984, *13*, 441-444.
- ¹⁴ Graham Barrett, *Chemistry and Biochemistry of the Amino Acids* chapter 11, reactions of amino acids, page 360, **1985**, Chapman and hall ltd, London
- 35 ¹⁵ M. H. S. A. Hamid, P. A. Slatford, J. M. J. Williams, *Adv. Synth. Catal.* **2007**, *349*, 1555 – 1575.
- ¹⁶ G. Guillena, D. J. Ramón, M. Yus Hydrogen autotransfer in the *N*-alkylation of amines and related compounds using alcohols and amines as electrophiles. *Chem. Rev.* **2010**, *110*, 1611 – 1641.
- 40 ¹⁷ G. E. Dobereiner, R. H. Crabtree, *Chem. Rev.* **2010**, *110*, 681 – 703
- ¹⁸ A. J. A. Watson, J. M. J. Williams, *Science* **2010**, *329*, 635–636
- ¹⁹ S. Bähn, S. Imm, L. Neubert, Dr. M. Zhang, H. Neumann, M. Beller. *ChemCatChem* **2011**, *3*, 1853 – 1864
- ²⁰ C. Gunanathan, D. Milstein, *Science* **2013**, *341*, 1229712.
- 45 ²¹ Q. Yang, Q. Wang, Z. Yu *Chem. Soc. Rev.*, **2015**, *44*, 2305 - 2329
- ²² K. Barta, P. C. Ford, *Acc. Chem. Res.* **2014**, *47*, 1503 – 1512.
- ²³ U. R. Kreutzer *J. Am. Oil. Chem. Soc.* **1984**, *61*, 343 – 348.

- 24 Industrial production of amines mostly based nucleophilic substitution mechanism. For example, methylamines can be produced under 350-500 °C and 15-30 bar pressure using aluminum-based heterogeneous catalyst from ammonia and methanol. See: Weissermel, K.; Arpe, H.-J. *Industrial Organic Chemistry*, Fourth
5 Completely Revised Edition, Wiley-VCH: Weinheim, 2003, p. 51.
- 25 C.-P. Xu, Z.-H. Xiao, B.-Q. Zhuo, Y.-H. Wang, P.-Q. Huang *Chem. Commun.*, 2010, 46, 7834–7836
- 26 V. N. Tsarev, Y. Morioka, J. Caner, Q. Wang, R. Ushimaru, A. Kudo, H. Naka, S. Saito *Org. Lett.* 2015, 17, 2530–2533
- 10 27 T. Yan, B. L. Feringa, K. Barta, *Nat. Commun.* 2014, 5, 5602.
- 28 B. L. Conley, M. K. Pennington-Boggio, E. Boz, and T. J. Williams *Chem. Rev.* 2010, 110, 2294–2312
- 29 One example reported N-alkylation of anilines with amine with Shvo Catalyst, see D. Hollmann, S. Bahn, A. Tillack, M. Beller *Angew. Chem. Int. Ed.* 2007, 46,
15 8291–8294.
- 30 C. P. Casey, J. B. Johnson *J. AM. CHEM. SOC.* 2005, 127, 1883-1894
- 31 I. A. Shuklov, N. V. Dubrovina, A. Börner *Synthesis* 2007, 19, 2925 – 2943.
- 32 O. Boutureira, G. J. L. Bernardes *Chem. Rev.* 2015, 115, 2174–2195
- 33 Zon W Lai, Agnese Petrera², Oliver Schilling Protein amino-terminal
20 modifications and proteomic approaches for N-terminal profiling *Current Opinion in Chemical Biology* 2015, 24, 71–79

Claims

1. A method for the *N*-alkylation of an unprotected amino acid or the N-terminus of an unprotected oligopeptide substrate, comprising reacting said unprotected amino acid or oligopeptide substrate with an alcohol in the presence of a homogeneous transition metal catalyst, preferably wherein the transition metal catalyst is a Fe- or Ru-based catalyst.

2. Method according to claim 1, wherein the Fe-based catalyst is of one of the following formula's :

**Formula A**

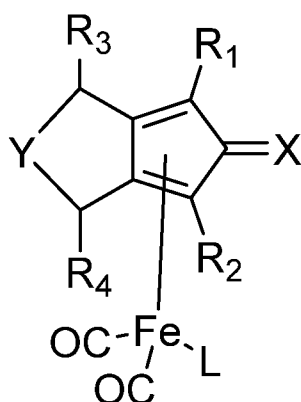
wherein R_1 and R_2 are independently selected from the group consisting of: alkyl, aryl, $-\text{CH}_2\text{Ph}$ and silyl moieties (e.g. TMS (trimethylsilyl), TBDMS (tertbutyldimethylsilyl), TIPS (triisopropylsilyl) or TBDPS (tertbutyldiphenylsilyl))

R_3 and R_4 are independently selected from the group consisting of: hydrogen, optionally substituted alkyl and optionally substituted, aryl (with broad range of substitution e.g. $-\text{OCH}_3$, CN , NO_2 , COOR , Alkyl)

L is selected from the group consisting of CO, acetonitrile phosphine, phosphite, phosphoramidite, primary or secondary amine, primary, secondary or tertiary alcohol, preferably L is CO or acetonitrile

X is O, NH or N-R₅ where R₅ is an alkyl or aryl, preferably X is O.

5



Formula B

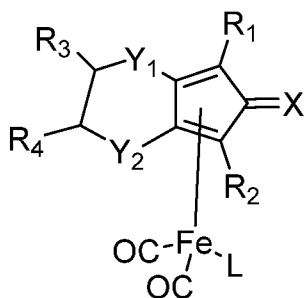
wherein R₁ and R₂ are independently selected from the group
 10 consisting of: alkyl, aryl, -CH₂Ph and silyl moieties ((e.g. TMS (trimethylsilyl), TBDMS (tertbutyldimethylsilyl), TIPS (triisopropylsilyl) or TBDPS (tertbutyldiphenylsilyl));

R₃ and R₄ are independently selected from the group consisting of:
 hydrogen, optionally substituted alkyl (with broad range of substitution),
 15 and optionally substituted aryl (with broad range of substitution e.g. H, -OCH₃, CN, NO₂, COOR, Alkyl), silyl, or OH, O-R; NH₂, NH-R (where R can be alkyl, aryl, sulphonyl, tosyl, mesyl);

Y = CH₂; CH₂-CH₂; CH₂-O-; CH₂-NH-; O; NH or NHR₅, wherein R₅=alkyl or aryl, sulphonyl, tosyl, mesyl;

20 L is selected from the group consisting of CO, acetonitrile phosphine, phosphite, phosphoramidite, primary or secondary amine, primary, secondary or tertiary alcohol, preferably L is CO or acetonitrile

X is O, NH or N-R₆ where R₆ is an alkyl or aryl, preferably X is O.

**Formula C**

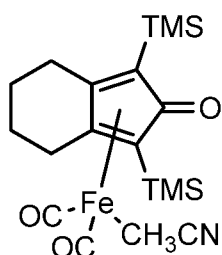
wherein R_1 and R_2 are independently selected from the group
 5 consisting of alkyl, aryl, $-\text{CH}_2\text{Ph}$ and silyl moieties (e.g. TMS [trimethylsilyl], TBDMS [tertbutyldimethylsilyl], TIPS [triisopropylsilyl] or TBDPS [tertbutyldiphenylsilyl])

R_3 and R_4 are independently selected from the group consisting of
 alkyl (with broad range of substitution), silyl, Aryl (with broad range of
 10 substitution on the aromatic ring: e.g. H, $-\text{OCH}_3$, CN, NO_2 , COOR, Alkyl); or
 OH, $\text{O}-R_5$ (where R_5 =alkyl, aryl); NH_2 , $\text{NH}-R_6$ (where R_6 =alkyl, aryl,
 sulphonyl), or wherein R_3 and R_4 are connected to form a cyclic structure,

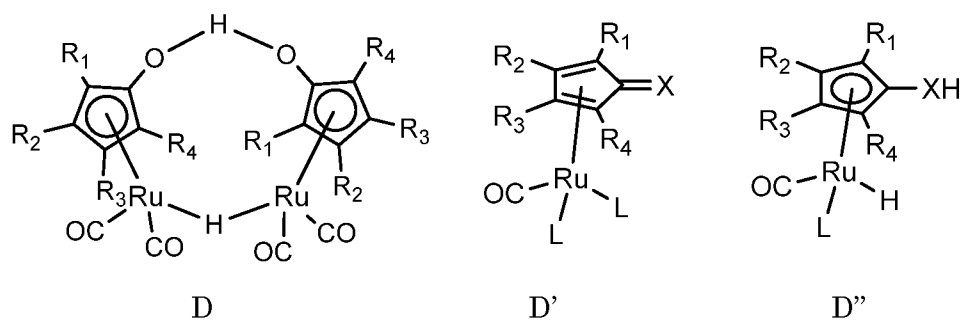
Y_1 and $Y_2 = \text{CH}_2$; O; NH; NHR_7 (where R_7 =alkyl, aryl, sulphonyl,
 tosyl, mesyl); CH_2-CH_2 ; $\text{CH}_2-\text{O}-$; or $\text{CH}_2-\text{NH}-$

15 L is selected from the group consisting of CO, acetonitrile
 phosphine, phosphite, phosphoramidite, primary or secondary amine,
 primary, secondary or tertiary alcohol, preferably L is CO or acetonitrile

3. Method according to claim 2, wherein the Fe-based catalyst is

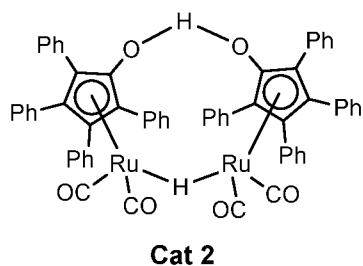
**Cat 1b**

4. Method according to claim 1, wherein the Ru-based catalyst is of one of the following formula's :



wherein R₁, R₂, R₃ and R₄ are independently selected from the group consisting of alkyl, aryl, -CH₂Ph, optionally with one or more substitution on the aromatic ring, preferably wherein R₁, R₂, R₃ and R₄ are (optionally substituted) aryl or -CH₂Ph.

5. Method according to claim 4, wherein the Ru-based catalyst is of the formula D, preferably wherein the catalyst is



6. Method according to any one of the preceeding claims, wherein the amino acid or oligopeptide substrate consists of or comprises α- amino acids.

7. Method according to any one of the preceeding claims, wherein the oligopeptide substrate consists of from two to eight amino acids, preferably two to five amino acids, more preferably wherein the oligopeptide is a dipeptide or a tripeptide.

5

8. Method according to any one of the preceeding claims, wherein the amino acid or oligopeptide substrate consists of or comprises an amino acid residue having an uncharged or hydrophobic side chain, preferably wherein the amino acid residue is selected from the group consisting of Ala, Gly, Pro, Val, Leu, Ile, Gln, Asn.

10

9. Method according to any one of the preceeding claims, wherein the alcohol is a long-chain primary alcohol, preferably a C₈-C₁₈ fatty alcohol.

10. Method according to any one of claims 1- 8, wherein the alcohol is an optionally substituted linear, branched or aromatic C₁-C₆ alcohol, preferably selected from the group consisting of ethanol, isopropanol, 1-butanol, 2-chloroethanol, 2-butanol, cyclopropylmethanol, benzylalcohol and 4-chlorobenzyl alcohol.

15

11. Method according to any one of the preceeding claims, wherein the alcohol is derived from biomass.

12. *N*-alkylated amino acid or oligopeptide selected from the group consisting of the compound of formula *N,N*-di-(5-hydroxypentyl)-phenylalanine (3dk), *N,N*-di-(*n*-pentyl)-glycine (3gm), *N,N*-(di-*n*-nonyl)-phenylalanine (3dj), *N,N*-di-*n*-nonyl-glycine (3gj), *N,N*-di-dodecylglycine (3gn), *N,N*-di-(5-hydroxy-pentyl)-glycyl-alanine (5ak), *N,N*-di-(*n*-dodecyl)-glycyl-alanine (5an), *N*-nonyl-glycine (3gi'), *N*-decyl-glycine (3go'), *N*-dodecyl-alanine (3gn'), *N*-tetradecylglycine (3gp'), *N*-hexadecylglycine (3gq'), *N*-octadecylglycine (3gr'), *N*-dodecyl-alanine (3fn') and *N*-nonyl-proline (3aj).

25

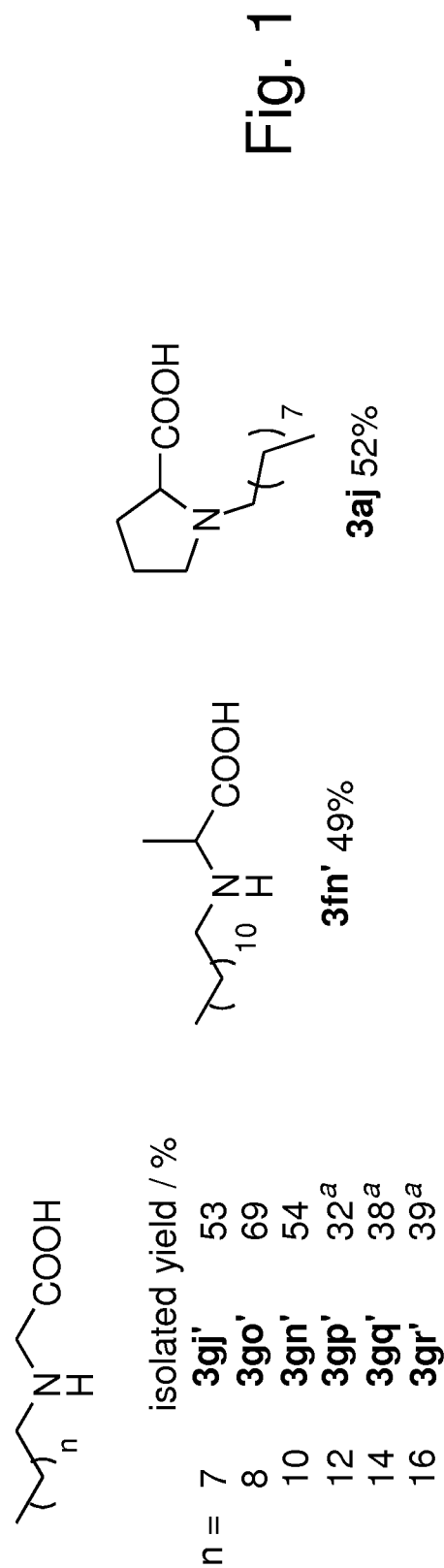
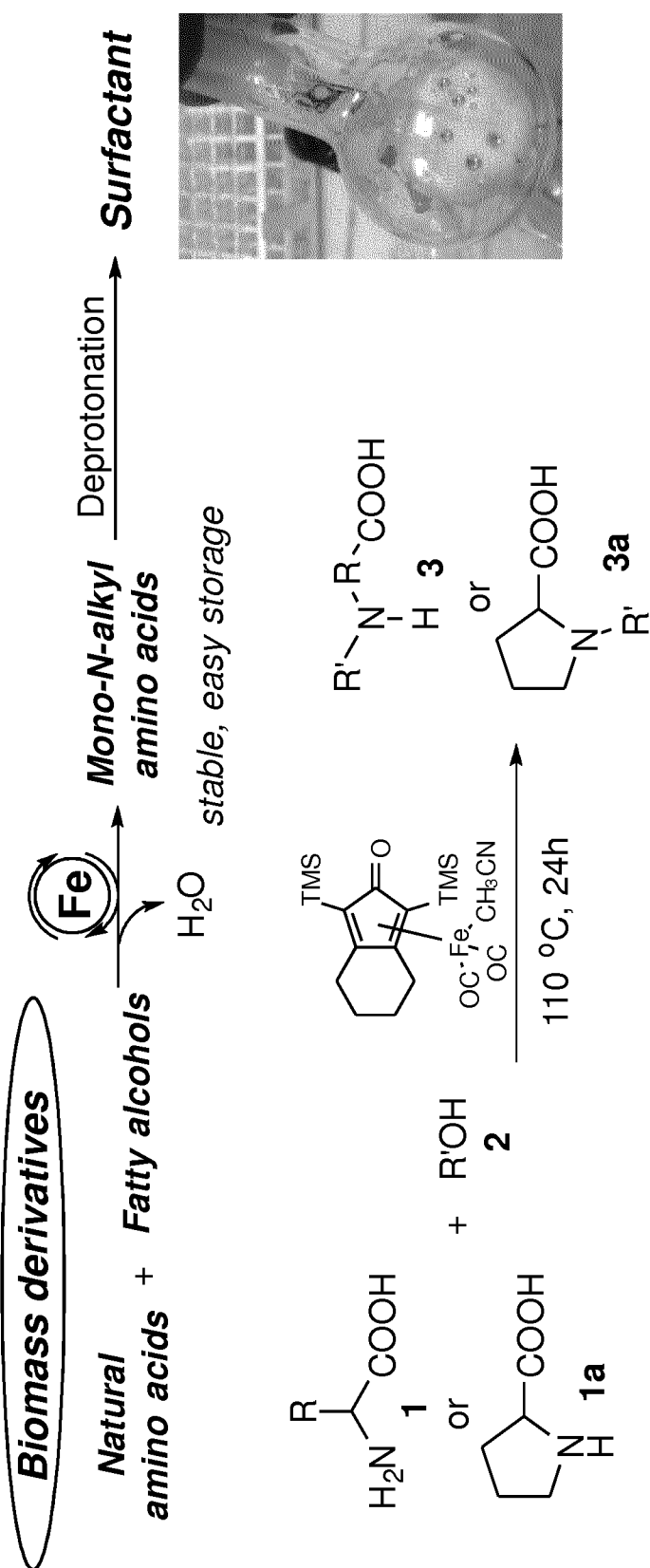
13. A method according to claim 1 for the synthesis of an amino acid based surfactant, preferably a long-chain *N*-alkyl amino acid, comprising the *N*-alkylation of an unprotected amino acid or the N-terminus of an oligopeptide substrate by reacting said unprotected amino acid or

5 oligopeptide substrate with a long chain (fatty) alcohol in the presence of a homogeneous transition metal catalyst.

14. Method according to claim 13, comprising mono-*N*-alkylation using a Fe-based catalyst, preferably a catalyst mentioned in claim 2 or 3.

15. Method according to claim 13 or 14, comprising the synthesis of an amino acid-based surfactant from natural building blocks/renewable sources.

16. Use of an *N*-alkylated amino acid or oligopeptide according to claim 12 as surfactant.



INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2018/058488

A. CLASSIFICATION OF SUBJECT MATTER
 INV. C07D207/16 C07K1/00 C07K5/06
 ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 C07D C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

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

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LEONARD J. ET AL.: "A Survey of the Borrowing Hydrogen Approach to the Synthesis of some Pharmaceutically Relevant Intermediates", ORGANIC PROCESS RESEARCH AND DEVELOPMENT, vol. 19, 4 August 2015 (2015-08-04), pages 1400-1410, XP002770469, scheme 14	1
Y	-----	1-16
Y	YAN T. ET AL.: "Iron catalysed direct alkylation of amines with alcohols", NATURE COMMUNICATIONS, vol. 5, 5602, 26 November 2014 (2014-11-26), pages 1-7, XP002770470, DOI: 10.1038/ncomms6602 cited in the application the whole document ----- -/-	1-16



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

13 June 2018

Date of mailing of the international search report

03/07/2018

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040,
 Fax: (+31-70) 340-3016

Authorized officer

Bérillon, Laurent

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2018/058488

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	TUAN THANH DANG ET AL: "Efficient Ruthenium-Catalyzed N-Methylation of Amines Using Methanol", ACS CATALYSIS, vol. 5, no. 7, 2 July 2015 (2015-07-02), pages 4082-4088, XP055264644, US ISSN: 2155-5435, DOI: 10.1021/acscatal.5b00606 the whole document	1-16
Y	YAN T. ET AL.: "Benzylamines via Iron-Catalysed Direct Amination of Benzyl Alcohols", ACS CATALYSIS, vol. 6, 3 December 2015 (2015-12-03), pages 381-388, XP002770471, cited in the application the whole document	1-16
Y	GABRIELA GUILLENA ET AL: "Hydrogen Autotransfer in the N -Alkylation of Amines and Related Compounds using Alcohols and Amines as Electrophiles", CHEMICAL REVIEWS, vol. 110, no. 3, 10 March 2010 (2010-03-10), pages 1611-1641, XP055262106, US ISSN: 0009-2665, DOI: 10.1021/cr9002159 the whole document	1-16
X	DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 15 January 1998 (1998-01-15), XP002781900, retrieved from STN Database accession no. 199870-59-2 compound with CAS registry number 199870-59-2	12
X	DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 17 May 1986 (1986-05-17), XP002781901, retrieved from STN Database accession no. 102116-00-7 compound with CAS registry number 102116-00-7	12
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2018/058488

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 16 November 1984 (1984-11-16), XP002781902, retrieved from STN Database accession no. 50997-13-2 compound with CAS registry number 50997-13-2	12
X	----- DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 16 November 1984 (1984-11-16), XP002781903, retrieved from STN Database accession no. 20933-87-3 compound with CAS registry number 20933-87-3	12
X	----- DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 16 November 1984 (1984-11-16), XP002781904, retrieved from STN Database accession no. 78961-17-8 compound with CAS registry number 78961-17-8	12
X	----- DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 16 November 1984 (1984-11-16), XP002781905, retrieved from STN Database accession no. 18749-74-1 compound with CAS registry number 18749-74-1	12
X	----- DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 12 August 2005 (2005-08-12), XP002781906, retrieved from STN Database accession no. 859981-28-5 compound with CAS registry number 859981-28-5 ----- -/--	12

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2018/058488

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 16 November 1984 (1984-11-16), XP002781907, retrieved from STN Database accession no. 42373-45-5 compound with CAS registry number 42373-45-5	12
X	----- DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 16 November 1984 (1984-11-16), XP002781908, retrieved from STN Database accession no. 35168-40-2 compound with CAS registry number 35168-40-2	12
X	----- DATABASE REGISTRY [Online] CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US; 16 November 1984 (1984-11-16), XP002781909, retrieved from STN Database accession no. 65527-68-6 compound with CAS registry number 65527-68-6 -----	12