OUKA
The University of Osaka
Institutional Knowledge Archive

Title	Synthetic study of protein and selenoprotein based on selenium chemistry
Author(s)	武居, 俊樹
Citation	大阪大学, 2018, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/69343
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

Synthetic study of protein and selenoprotein based on selenium chemistry

Toshiki Takei

2018

Acknowledgments

This studies presented in this thesis have been performed at Institute for Protein Research (IPR) in Osaka University. The author received the countless helps and information from a lot of people. However, the author especially would like to express his gratitude to following persons.

The author would like to express his sincere gratitude to his supervisor, Prof. Hironobu Hojo whose comments, supports, suggestions and friendships were innumerably valuable throughout the course of his study.

The author would like to thank Prof. Michio Iwaoka and Mr. Shingo Shimodaira in Tokai University for providing selenocysteine derivative and helps for his study.

The author would like to thank Assoc. Prof. Kenta Arai in Tokai University for valuable suggestions and supports of oxidative folding of Se-Ins.

The author would like to thank Asst. Prof. Satoshi Watanabe, Asst. Prof. Masaki Okumura, and Asst. Prof. Yuta Amagai in Institute of Multidisciplinary Research for Advanced Materials (IMRAM) in Tohoku University for their many collaborative researches, crystallization of Se-Ins, X-ray crystal analysis, phosphorylation assay and measurement of degradation rate by IDE.

The author would also like to thank Prof. Toshifumi Takao and Mr. Tomoshige Andoh in IPR for valuable suggestions and friendships as well as for measurement of dismutase activity of SOD-1 and mass spectrum.

The author would like to thank Assoc. Prof. Toru Kawakami and Asst. Prof. Yuya Asahina in Laboratory of Protein Organic Chemistry for their advices, guidance and friendships.

The author would also like to thank all members in Laboratory of Protein Organic Chemistry for their friendships.

Finally, the author would also like to express his gratitude to his parents for their moral support and warm encouragements.

Abbreviations

AAA	amino acid analysis
Ac	acetyl
Acm	acetamidomethyl
AEMTS	2-aminoethyl methanethiosulfonate
Boc	*tert*-butoxycarbonyl
Bu^t	*tert*-butyl
CD	circular dichroism
DCE	1,2-dichloroethane
DCM	dichloromethane
DIC	*N*,*N*'-diisopropylcarbodiimide
DIEA	*N*,*N*'-diisopropylethylamine
Dmb	dimethoxybenzyl
DMF	*N*,*N*'-dimethylformamide
DMSO	dimethyl sulfoxide
DODT	3,6-dioxa-1,8-octanedithiol
DPDS	2,2'-dipyridyldisulfide
DTT	dithiothreitol
EDT	1,2-ethanedithiol
eq.	equivalent
ESI	electrospray ionization
Et	ethyl
Fmoc	9-fluorenylmethoxycarbonyl
GSH	reduced glutathione
GSSG	oxidized glutathione
HBTU	*N*-[(1H-benzotriazole-1-yl)(dimethylamino)methylene]-*N*-methylmethanaminiumhexafluorophosphate *N*-oxide
HOBt	1-hydroxybenzotriazole
HOOBt	3-hydroxy-1,2,3-benzotriazin-4(3H)-one
HPLC	high-performance liquid chromatography
*i*Noc	4-pyridylmethoxycarbonyl
MALDI	matrix-assisted laser desorption/ionization
MBHA	4-methylbenzhydrylamine
MeCN	acetonitrile
MESNa	sodium-2-mercaptoethanethiosulfonate

MPA	3-mercaptopropionic acid
MPM	4-methoxyphenylmethyl
MS	mass
NAC	*N*-alkylcysteine
NCL	native chemical ligation
NMP	*N*-methyl-2-pyrrolidone
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
Pys	2-pyridylsulfanyl
RP	reverse-phase
SDS	sodium dodecyl sulfate
SPPS	solid-phase peptide synthesis
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid
TfOH	trifluoromethanesulfonic acid
TIS	triisopropylsilane
TOF	time of flight
Tris	tris(hydroxymethyl)aminomethane
Trt	trityl

Contents

Acknowledgements

Abbreviations

Chapter 1. General Introduction

1-1. Selenium and its biological importance 1

1-2. Biosynthesis of selenoprotein 2

1-3. Glutathione peroxidese (GPx) 4

1-4. Preparation of Sec containing peptide and its application for the synthesis of protein 6

1-5. Purposes of this study 9

1-6. References 9

Chapter 2. Preparation of Selenoinsulin as a Long-Lasting Insulin Analogue

2-1. Introduction 12

2-2. Results and discussion 13

2-3. Experimental section 19

2-4. References 28

Chapter 3. One-Pot Four-Segment Ligation Using Seleno- and Thioesters: Synthesis of Superoxide Dismutase

3-1. Introduction 31

3-2. Results and discussion 32

3-3. Experimental section 39

3-4. References 54

Chapter 4. Summary

4-1. Summary 57

4-2. Lists of Publication 58

4-3. Related Papers 58

1. General Introduction

1-1. Selenium and its biological importance

Selenium, which is an essential trace element, is incorporated into proteins genetically as a selenocysteine (Sec, U) residue (Figure 1-1). Sec containing protein (selenoprotein) plays essential roles for health in human and other mammals, because these proteins are now known to be implicated in various important cellular and organismal functions, such as immune function[1] and mammalian development[2] and so on. However, the importance of selenium atom for human had not been recognized for a long time.

Cysteine (Cys, C) Selenocysteine (Sec, U)

Figure 1-1. Chemical structure of cysteine and selenocysteine.

Until the early 20th century, selenium and its inorganic compounds had been regarded only as a toxic compound, because eating cereal grain containing high level of selenium caused the livestock death. However, the view changed after the discovery in which factor 3, a mixture of inorganic selenium compounds isolated from hog kidney, effectively prevented experimental necrosis of rat liver.[3] In the cases of other diseases, such as white muscle disease in cattle and mulberry heart disease in pigs, factor 3 was also effective and the results suggested that selenium might participate in redox reactions. Based on these findings, scientists attributed the cause of these diseases to the insufficient selenium supply and recognized selenium as an essential trace element. However, the precise mechanism of Se atom to prevent these diseases was still unknown.

The direct evidence which shows a function of Se atom at the molecular level was first reported by Flohe and co-workers. They proved the existence of selenium atom in the glutathione peroxidase (GPx), which is one of the antioxidant enzymes, by X-ray crystal analysis.[4] Their analysis showed that selenium atom is covalently bound to the GPx. A few years later, it was found that selenium atom is incorporated as the selenocysteine[5] and the

encoded codon was identified as UGA.[6] Through these studies, selenium has recognized as not only a toxic compound but also an essential trace element for humans and other mammals and Sec began to be called 21st amino acid. Currently, 25 kinds of selenoproteins are found in human (more than 50 selenoproteins known in several species) (Table 1-1).

Table 1-1. Identified selenoproteins and their functions in human.

Selenoproteins	Function	Selenoproteins	Function
GPx families (5 types)	glutathione peroxidase	Selenoprotein I	unknown
Iodothyronine deiodinases (3 types)	thyroid hormone–activating iodothyronine deiodinase	15 kDa Selenoprotein	putative role in quality control of protein folding
Thioredoxin reductases (3 types)	reducion of thioredoxin	Selenoprotein M	unknown
Methionine-R-sulfoxide reductase	reduction of oxidized methionine residue	Selenoprotein K	putative role in ER–associated degradation
Selenophosphate synthetase 2	involved in synthesis of selenoproteins	Selenoprotein S	putative role in ER–associated degradation
Selenoprotein W	unknown	Selenoprotein O	unknown
Selenoprotein T	unknown	Selenoprotein N	putative role during muscle development
Selenoprotein H	unknown	Selenoprotein P	Se transport
Selenoprotein V	unknown		

1-2. Biosynthesis of selenoprotein

Sec incorporation into proteins is regulated by several specific proteins. In eukaryote, the first stage is the synthesis of selenocysteinyl-tRNA$^{[Ser]Sec}$ and the second is its use for translation in ribosome. Biosynthetic pathway of eukaryotic selenocysteinyl-tRNA$^{[Ser]Sec}$ is shown in Figure 1-2.[7]

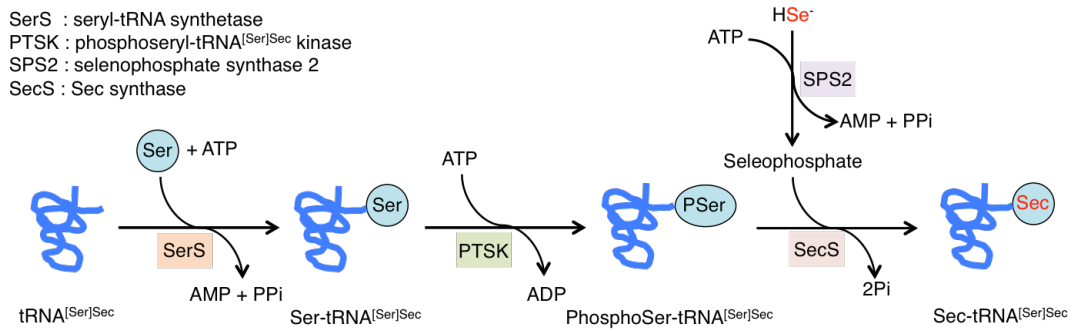

Figure 1-2. Biosynthetic route of Sec-tRNA$^{[Ser]Sec}$ in eukaryote.

It starts from the aminoacylation of tRNA$^{[Ser]Sec}$ by Ser in the presence of ATP catalyzed by the seryl-tRNA synthethase (SerS). After the acylation, the side chain hydroxy group of Ser is

phosphorylated using ATP by phosphorseryl-tRNA kinase (PSTK) and the generated phosphoseryl-tRNA$^{[Ser]Sec}$ is converted to dehydroalanyl-tRNA$^{[Ser]Sec}$ by Sec synthase (SecS) via β-elimination. Finally, selenophosphate, synthesized from HSe^- and ATP catalyzed by selenophosphate synthase 2 (SPS2), is coupled to dehydroalanyl-tRNA$^{[Ser]Sec}$ under the SecS catalysis. Selenocysteinyl-tRNA$^{[Ser]Sec}$ is then used for translation in ribosome complex as well as other aminoacyl-tRNAs (Figure 1-3).

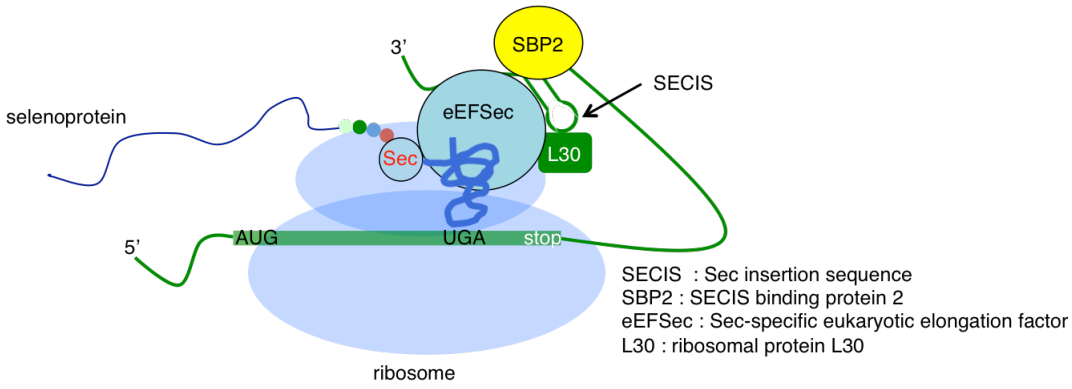

Figure 1-3. Mechanism of Sec incorporation into selenoproteins in eukaryote.

The characteristic feature for Sec incorporation is the presence of Sec insertion sequence (SECIS) element, which locates at 3'-untransrated mRNA region.[8] Sec is incorporated to protein at UGA, stop codon, by SECIS and other functional protein complex even though translation of normal protein is stopped by this codon. When translation of selenoprotein reached at UGA codon, the complex composed of the SECIS, SECIS binding protein 2 (SBP2), Sec-specific translation elongation factor (eEFSec) and ribosomal protein L30, is formed. Several reports demonstrated that SBP2 acts as a regulator for the selenoprotein synthesis,[9, 10] as the knockdown of SBP2 in mammalian cell decreases the total expression level of selenoproteins, whereas the overexpression of SBP2 increases Sec incorporation. eEFSec acts as recruiter for selenocysteinyl-tRNA$^{[Ser]Sec}$ to SECIS-SBP2 complex,[11] because eEFSec only binds to selenocysteinyl-tRNA$^{[Ser]Sec}$ unlike the canonical elongation factor involved in 20 kinds of amino acids incorporation. Further study indicated that ribosomal protein L30 has been predicted to constitute a part of the basal Sec insertion machinery.[12] After Sec incorporation, the translation of selenoproteins is stopped by actual stop codon. Recently, another SECIS-binding proteins, eukaryotic initiation factor 4a3 (eIF4a3)[13] and nucleolin[14], were

identified and proposed their roles in selenoprotein synthesis. However, function of these proteins in the selenoprotein translation are still unclear. Further studies are required to elucidate the physiological roles of each SECIS-binding protein.

1-3. Glutathione peroxidase (GPx)

One of the most studied selenoproteins is GPxs, which act as antioxidant enzyme. In mammals, there are eight GPxs in which five GPxs, GPx1, GPx2, GPx3, GPx4 and GPx6, have one Sec in their active site. The active site of other three GPxs, GPx5, GPx7 and GPx8, is replaced by Cys. GPx1 is the most abundant selenoprotein not only in the GPx family, but also in the entire selenoproteins in mammal, and their redox-catalytic activity has been studied (Figure 1-4).[15]

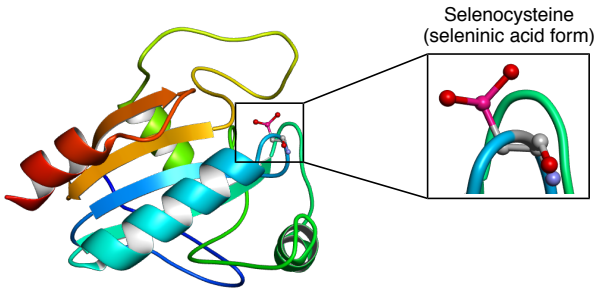

Figure 1-4. X-ray structure of GPx1 (PDB code : 1gp1).[16]

GPx1 catalyzes the glutathione-dependent reduction of hydroperoxides (ROOH), for example hydrogen peroxide (H_2O_2), with glutathione reductase (GR) in cytosol. Three active states in the catalytic cycle are proposed and the detail is elucidated as shown in Figure 1-5.

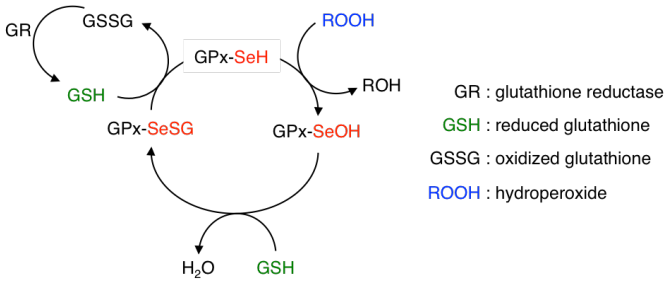

Figure 1-5. Plausible catalytic cycle of GPx.

As mutational study of Sec to Cys in the GPx showed that the enzymatic activity was approximately 1000-fold lower than native GPx,[17] Sec is important for the redox reaction. Actually, Sec has some characteristic physical properties in comparison with Cys in points that it has lower pKa (Cys: 8.25, Sec: 5.24-5.63)[18, 19] and redox potential value ($(Cys)_2$: - 215 mV, $(Sec)_2$: - 381 mV).[20] Based on the reaction mechanism as an antioxidant, several GPx mimics, which are organoselenium compounds and peptides, have been designed and investigated.[21, 22]

The author previously reported the structure activity relationship (SAR) study between the GPx activity and a catalytic triad,[23] which is a hydrogen bond network around Sec proposed by X-ray crystal analysis in GPxs, composed of Sec (Cys), Gln and Trp.[16] Although the structure of catalytic triad is well conserved in GPx family, there was little evidence for their relevance to the GPx activity. To model the catalytic triad structure, the author designed the model selenopeptides in silico and synthesized them. As a result, selenopeptide with both Gln and Trp revealed highest GPx-like activity than the ones in which either of the residue was substituted to Ala.

As a promising GPx mimics, Ebselen was previously synthesized by Weiss et al. as an antioxidant (Figure 1-6).[24] When the Se-N bond is reduced in the presence of excess thiol, the generated selenol (SeH) exhibits the GPx-like activity just like that of GPx and enters the catalytic cycle.[25] In addition, ^{75}Se-labeled studies demonstrated that the selenium atom in Ebselen is not used for the synthesis of GPx in human body and the Ebselen is excreted from the body,[26, 27] accounting for its lack of toxicity. Therefore, the Ebselen is currently under clinical phase III in Japan for antioxidant.

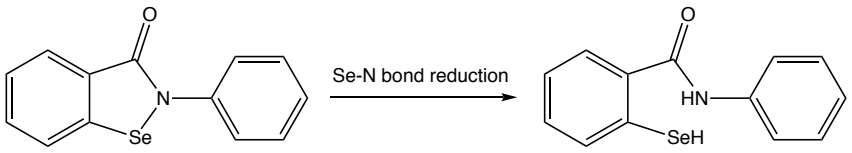

Figure 1-6. Chemical structure of Ebselen and its active form.

1-4. Preparation of Sec containing peptide and its application for the synthesis of protein

Sec containing peptide or protein would be useful for the understanding of protein function as well as for the development of new functional peptide or protein. Therefore, substitution of Cys to Sec in protein has been studied by using recombinant DNA technology. As an example in the recent studies, Sec-substituted cytochrome P450 enzymes were expressed by some research groups.[28] The P450 enzymes, which are heme-dependent monooxygenases, catalyze a variety of physiologically important oxidative reactions, such as C-H bond hydroxylation, hetero atom oxidation and so on. A coordination of Cys thiolate to the heme iron is believed to control the reactivity, however, the precise correlation is not clear. To explore the effect, the P450 enzymes containing the selenolate-ligated heme were expressed by using the Cys auxotrophic strain of *E. coli*, which cannot synthesize Cys due to the mutation of *CysE* gene, or by using the engineered gene containing SECIS of formate dehydrogenase from *E. Coli*. These results demonstrated that the Sec-substitution changes the reaction rate of C-H bond hydroxylation, and it also suggested that the Cys thiolate controls the reactivity in the native enzyme. However, the expression by the Cys auxotrophic strain gave the mixture with non-substituted P450 enzyme. In addition, the use of the engineered gene containing SECIS required the mutation of two amino acid residues other than the Cys residue due to the insertion of SECIS. Thus, the preparation of homogeneous Sec-substituted protein by recombinant DNA technology is still challenging.

On the other hand, chemical synthesis potentially yields Sec containing peptide and protein easily and in high purity. Chemical synthesis of selenoglutathione, which is an analog of glutathione, was reported in solution method.[29,30] Glutathione is a tripeptide composed of Glu, Cys and Gly, and widely distributes in cells to keep the redox condition. To reduce the reactive oxygen species (ROS), several antioxidant enzymes, such as superoxide dismutase (SOD), catalase and GPx, and small antioxidant compounds, glutathione, tocopherol (Vitamin E) and ascorbic acid, work in the cell. Glutathione also works for the quality control of disulfide containing protein by disulfide reshuffling. Recently, selenoglutathione was utilized as a reagent for oxidative folding.[31] The folding rate in the presence of selenoglutathione was faster than that in the presence of glutathione. The result indicated that selenoglutathione could rescue the kinetically trapped folding intermediate to the native folded state, due to its lower redox potential than glutathione. The synthesis of selenoanalogs of biologically active peptides, such as oxytocin and somatostatin, were also reported in solution method.[32,33] For example,

deamino-1-selenooxytocin, which is one of the selenoanalogs, revealed higher antidiuretic activity than native oxytocin and deamino-oxytocin.

The solid-phase peptide synthesis (SPPS)[34] developed by R. B. Merrifield realized an easy and quick synthesis of peptides. The general procedure of SPPS is the removal of protecting group at the terminal amino group followed by coupling of N-protected amino acid to the resin (Figure 1-7).

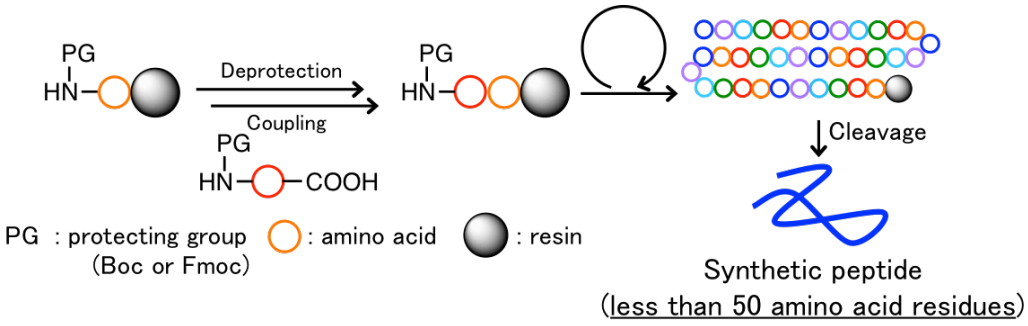

Figure 1-7. Outline of solid-phase peptide synthesis (SPPS).

These steps are repeated until the completion of the chain assembly from C terminal to N terminal direction. Metalloselenonein, in which all of seven Cys are substituted with Sec, was synthesized Boc-assisted SPPS by automated peptide synthesizer.[35] Native metallothionein is involved in the provision of physiological metal ions for the metalloenzymes. It also acts for the detoxification of heavy metals by binding them to thiol groups on the Cys. The substitution to Sec showed that three copper ions bind to metaloselenonein per molecule compared with six copper ions per molecule in the native one. The less binding activity was explained by the ionic radius between sulfur and selenium. The report indicated probability that Sec-substitution of Cys in the peptide (protein) could change its function. It has to be noted that the synthesis of Sec-containing peptide by the Fmoc method is difficult, due to the sensitivity of Sec to the piperidine treatment. It causes the dehydroalanine formation via β-elimination and subsequent the formation of piperidine adduct by Michael addition.[36] Thus, the repeated piperidine treatments to Sec-containing peptide in the synthesis significantly reduces the synthetic yield.

In the synthesis of protein with more than 100 amino acid residues, segment condensation method should be applied, because SPPS can be usually applied to the synthesis of peptide with less than 50 amino acid residues due to the increase of the side products. Kent and co-workers reported that thioester based segment condensation method, the native chemical

ligation (NCL) method (Figure 1-8 a).[37]

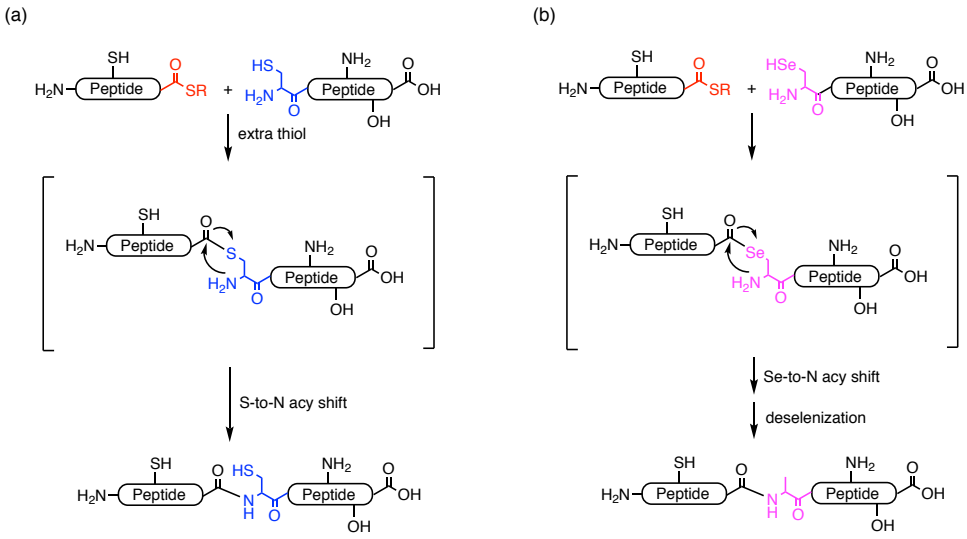

Figure 1-8. Outline of NCL method (a) and deselenization based Sec-mediated NCL (b).

The NCL method can condense peptide thioester and peptide with Cys at its N-terminus. The reaction proceeds via transthioesterification between two peptides, and the formed intermolecular thioester bond spontaneously rearranges via *S*-to-*N* acyl shift to form native amide bond. The NCL method is widely used for the chemical synthesis of protein, although Cys is necessary at the ligation site. To expand the utility of NCL method, conversion of Cys to Ala by desulfurization[38] and the use of β-thiol containing amino acid derivatives were reported. However, if Ala is selected as a ligation site and if the protein has native Cys residue(s) in the sequence, the native Cys residue(s) has to be protected to avoid the undesired desulfurization. That is, the additional deprotection step is needed.

Sec has been used for the NCL reaction, due to its similar structure with Cys (Sec-mediated NCL). In addition, the reaction rate increases compared to that by using Cys due to the high nucleophilicity of Se. Further, Dawson et al. demonstrated that Sec can be easily converted to Ala after ligation by deselenization following the same procedure for the desulfurization of Cys.[39] The deselenization selectively proceeds under the mild radical reaction without desulfurization of Cys due to the weaker Se-C bond relative to the S-C bond (Figure 1-8 b). Therefore, Cys needs no protection during the Sec-mediated NCL method to give the synthetic protein efficiently. Thus, the Sec-mediated NCL with deselenization has an advantage over the NCL at Cys site combined with desulfurization.

In spite of the usefulness of the Sec for the ligation reaction, the synthesis of Sec-containing proteins is usually difficult, as the Selenium atom is easily removed by β-elimination during the SPPS and ligation reactions as described above. Therefore, a few reports on the synthesis or semi-synthesis of Sec-substituted proteins of, such as RNase A,[40] BPTI,[41] azurin[42] and selenoprotein M (SELM) as well as selenoprotein W (SELW),[43] by Sec-mediated NCL exist. Further study will give us some insights for breakthrough to synthesize and design the Sec containing peptide-based functional compound.

1-5. Purposes of this study

Due to the increasing recognition that selenium is an essential trace element for humans, the function of selenoproteins has attracted much attention. Therefore, chemists have been attempting to utilize selenium for the chemical protein synthesis as well as for development of new functional compounds. However, the synthetic procedure for the selenopeptide and selenoprotein has not been practical, especially, by the Fmoc method. This results in the premature understanding and application of selenoprotein. To overcome this problem, the author first describes the SPPS of selenopeptides by the Fmoc method and its use for novel selenoinsulin synthesis in chapter 2. The utilization of selenium can probably provide more potent synthetic method for the protein, even though the Sec-mediated NCL and subsequent the deselenization has already been developed. To expand the usefulness of selenium for the chemical protein synthesis, the author also describes the reactivity of a peptide aryl selenoester in a ligation chemistry and development of one-pot four-segment ligation by using the selenoester in chapter 3.

1-6. References

[1] M. P. Rayman, *Lancet* **2012**, *379*, 1256–1268.

[2] J. Köhrle, *Cell. Mol. Life Sci.* **2000**, *57*, 1853–1863.

[3] K. Schwarz, C. M. Foltz, *J. Am. Chem. Soc.* **1957**, *79*, 3292–3293.

[4] L. Flohé, W. A. Günzler, H. H. Schock, *FEBS Lett.* **1973**, *32*, 132–134.

[5] J. E. Cone, R. M. Del Rio, J. N. Davis, T. C. Stadtman, *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 2659–2663.

[6] I. Chambers, J. Frampton, P. Goldfarb, N. Affara, W. McBain, P. R. Harrison, *EMBO J.*

1986, *5*, 1221–1227.

[7] V. M. Labunskyy, D. L. Hatfield, V.N. Gladyshev, *Physiol. Rev.* **2014**, *94*, 739–777.

[8] M. J. Berry, L. Banu, Y. Y. Chen, S. J. Mandel, J. D. Kieffer, J. W. Harney, P. R. Larsen, *Nature* **1991**, *353*, 273–276.

[9] L. V. Papp, J. Lu, F. Striebel, D. Kennedy, A. Holmgren, K. K. Khanna, *Mol. Cell. Biol.* **2006**, *26,* 4895–4910.

[10] S. C. Low, E. Grundner-Culemann, J. W. Harney, M. J. Berry, *EMBO J.* **2000**, *19*, 6882–6890.

[11] R. M. Tujebajeva, P. R. Copeland, X. M. Xu, B. A. Carlson, J. W. Harney, D. M. Driscoll, D. L. Hatfield, M. J. Berry, *EMBO Rep.* **2000**, *1,* 158–163.

[12] L. Chavatte, B. A. Brown, D. M. Driscoll, *Nat. Struct. Mol. Biol.* **2005**, *12*, 408–416.

[13] M. E. Budiman, J. L. Bubenik, A. C. Miniard, L. M. Middleton, C. A. Gerber, A. Cash, D. M. Driscoll, *Mol. Cell* **2009**, *35*, 479-489.

[14] A. C. Miniard, L. M. Middleton, M. E. Budiman, C. A. Gerber, D. M. Driscoll, *Nucleic Acids Res.* **2010**, *38*, 4807-4820.

[15] S. Toppo, L. Flohé, F. Ursini, S. Vanin, M. Maiorino, *Biochim. Biophys. Acta* **2009**, *1790*, 1486–1500.

[16] O. Epp, R. Ladenstein, A. Wendel, *Eur. J. Biochem.* **1983**, *133*, 51–69.

[17] C. Rocher, J. L. Lalanne, J. Chaudiëre, *Eur. J. Biochem.* **1992**, *205*, 955–960.

[18] R. E. Huber, R. S. Criddle, *Arch. Biochem. Biophys.* **1967**, *122*, 164–173.

[19] A. P. Arnold, K. S. Tan, D. L. Rabenstein, *Inorg. Chem.* **1986**, *25*, 2433–2437.

[20] D. Besse, F. Siedler, T. Dierks, H. Kessler, L. Moroder, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 883–885; *Angew. Chem.* **1997**, *109*, 885-887.

[21] Y. Nakamura, Q. Feng, T. Kumagai, K. Torikai, H. Ohigashi, T. Osawa, N. Noguchi, E. Niki, K. Uchida, *J. Biol. Chem.* **2002**, *277*, 2687–2694.

[22] S. Yoshida, F. Kumakura, I. Komatsu, K. Arai, Y. Onuma, H. Hojo, B. G. Singh, K. I. Priyadarsini, M. Iwaoka, *Angew. Chem. Int. Ed.* **2011**, *50*, 2125–2128.

[23] T. Takei, Y. Urabe, Y. Asahina, H. Hojo, T. Nomura, K. Dedachi, K. Arai, M. Iwaoka, *J. Phys. Chem. B* **2014**, *118*, 492–500.

[24] R. Lesser, R. Weiss, *Ber. Dtsch. Chem. Ges.* **1924**, *57*, 1077–1082.

[25] H. Sies, *Free Radical Biol. Med.* **1993**, *14*, 313–323.

[26] R. Cantineau, G. Tihange, A. Plenevaux, L. Christiaens, M. Guillaume, A. Welter, N. Dereu, *J. Labelled Compd. Radiopharm.* **1985**, *23*, 59–65.

[27] A. Müller, H. Gabriel, H. Sies, R. Terlinden, H. Fischer, A. Römer, *Biochem. Pharmacol.* **1988**, *37,* 1103–1109.

[28] a) Y. Jiang, S. Sivaramakrishnan, T. Hayashi, S. Cohen, P. Moënne-Loccoz, S. Shaik, P. R. Ortiz de Montellano, *Angew. Chem. Int. Ed.* **2009**, *48*, 7193–7195; b) C. Aldag, I. A. Gromov, I. Garacia-Rubio, K. V. Koenig, I. Schlichting, B. Jaun, D. Hilvert, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 5481–5486; c) S. Sivaramakrishnan, H. Ouellet, J. Du, K. J. McLean, K. F. Medzihradszky, J. H. Dawson, A. W. Munro, P. R. Ortiz de Montellano, *Biochemistry* **2011**, *50*, 3014–3024; d) S. Sivaramakrishnan, H. Ouellet, H. Matsumura, S. Guan, P. Moënne-Loccoz, A. L Burlingame, P. R. Ortiz de Montellano, *J. Am. Chem. Soc.* **2012**, *134*, 6673–6684; e) A. Vandemeulebroucke, C. Aldag, M. T. Stiebritz, M. Reiher, D. Hilvert, *Biochemistry* **2015**, *54*, 6692–6703; f) E. L. Onderko, A. Silakov, T. H. Yosca, M. T. Green, *Nat. Chem.* **2017**, *9*, 623–628.

[29] D. Theodoropoulos, I. L. Schwartz, R. Walter, *Biochemistry* **1967**, *6*, 3927–3932.

[30] T. Tamura, T. Oikawa, A. Ohtaka, N. Fujii, N. Esaki, K. Soda, *Anal. Biochem.* **1993**, *208*, 151–154.

[31] J. Beld, K. J. Woycechowsky, D. Hilvert, *Biochemistry* **2007**, *46*, 5382–5390.

[32] R. Walter, W. Y. Chan, *J. Am. Chem. Soc.* **1967**, *89*, 3892–3898.

[33] B. Hartrodt, K. Neubert, B. Bierwolf, W. Blech, H.-D. Jakubke, *Tetrahedron Lett.* **1980**, *21*, 2393–2396.

[34] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

[35] T. Oikawa, N. Esaki, H. Tanaka, K. Soda, *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 3057–3059.

[36] D. Besse, L. Moloder, *J. Pept. Sci.* **1997**, *3*, 442–453.

[37] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779.

[38] L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533.

[39] N. Metanis, E. Keinan, P. E. Dawson, *Angew. Chem. Int. Ed.* **2010**, *49*, 7049–7053.

[40] R. J. Hondal, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2001**, *123*, 5140–5141.

[41] R. Quaderer, A. Sewing, D. Hilvert, *Helv. Chim. Acta* **2001**, *84*, 1197–1206.

[42] S. M. Berry, M. D. Gieselman, M. J. Nilges, W. A. van der Donk, Y. Lu, *J. Am. Chem. Soc.* **2002**, *124*, 2084–2085.

[43] L. Dery, P. S. Reddy, S. Dery, R. Mousa, O. Ktorza, A. Talhami, N. Meatnis, *Chem. Commun.* **2017**, *8*, 1922–1926.

2. Preparation of Selenoinsulin as a Long-Lasting Insulin Analogue

2-1. Introduction

Insulin, a small globular protein (5.8 kDa), comprises two peptide chains, the A-chain (Ins-A, 21 amino acid residues) and B-chain (Ins-B, 30 amino acid residues). The native structure in a monomeric active state is stabilized by two interchain disulfide bridges, Cys^{A7}-Cys^{B7} and Cys^{A20}-Cys^{B19}, in addition to one intrachain disulfide linkage, Cys^{A6}-Cys^{A11} (Figure 2-1).[1]

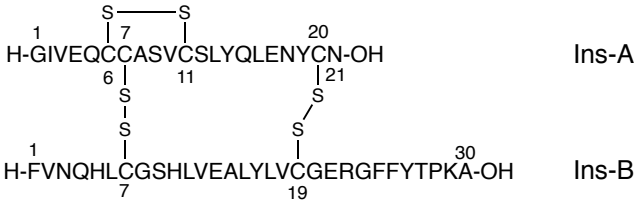

Figure 2-1. Primary sequence and disulfide pattern of wild-type bovine pancreatic insulin (BPIns).

Considerable efforts have been directed toward development of various insulin analogues[2] which imitate either bolus secretion of insulin for expeditiously reducing postprandial blood glucose levels[3] or basal secretion of insulin to control the glucose level for an entire day.[4] The latter long-acting analogues have been designed so that insulin forms infusible precipitates or soluble oligomers (hexamer or dihexamer) under physiological conditions and slowly releases active insulin monomers.

In contrast, the insulin-degrading enzyme (IDE) is a possible alternative target for diabetes therapy. IDE, which is involved in clearance of insulin and amyloid β (Aβ),[5] is found in the liver and kidneys. Recent research has revealed that synthetic IDE inhibitors increase circulation of insulin by preventing its degradation in the liver, thus resulting in improvement of the postprandial glucose tolerance.[6] However, other research suggests that IDE inhibitors could induce accumulation of Aβ in the brain,[7] and would lead to Aβ-mediated cognitive impairment. Hence, the design of long-lasting insulin analogues resistant against IDE would be desirable.[8]

In this study, we have attempted a new approach to a long-lasting insulin analogue by exploiting the unique chemical properties of a diselenide bond. Namely, introduction of two

juxtaposed selenium atoms to the insulin analogue could lead to a higher kinetic and thermodynamic stability than that of the wild-type without affecting the bioactivity. This new strategy is based primarily on the higher rotational barrier of a Se-Se bond (ca. 4 kcal mol^{-1}) than that of an S-S bond (ca. 3 kcal mol^{-1}),[9] and secondarily on the differing redox potentials of diselenides (E'_0 -381 mV) and disulfides (E'_0 -215 mV),[10] even though the structural features of a Se-Se and S-S bond are very similar to each other.

The lower redox potential of a Se-Se bond is also advantageous for preparation of a diselenide insulin analogue (selenoinsulin) because formation of a Se-Se bridge should occur independently of the presence of additional cysteine residues, and thus facilitate the correct oxidative folding of cysteine-rich peptides. This advantage was already fully confirmed by the quantitative oxidative production of the natural isomer of endothelin[11] and both the natural and non-natural isomers of apamin.[12] More recently, it was successfully used in the generation of the correct natural diselenide/ disulfide isomers of many cysteine-rich peptides (for reviews see Ref. [13]). The obtained Sec-substituted peptides were found to retain the native structures with biological activities comparable to those of the wild-types, thus confirming the isomorphous structural character of the Se-Se versus the S-S bridges.[13a,c,14] By employing bovine pancreatic insulin (BPIns; Figure 2-1) as a model protein, we designed a new selenoinsulin analogue in which Cys^{A7} and Cys^{B7} are replaced with two selenocysteine (Sec; U) residues. We chose this strategy not only because it would bring significant conformational stabilization around the solvent-exposed Cys^{A7}-Cys^{B7} bridge by substitution with a more robust Se-Se bridge, but also because it would inhibit undesired formation of the swap species having an incorrect Cys^{A6}-Cys^{B7} bridge.[15]

2-2. Results and discussion

Preparation and characterization of selenoinsulin

We first prepared the two selenopeptides, Se-Ins-A and Se-Ins-B, by applying Fmoc method. The synthesis of Se-Ins-A was started from Fmoc-Asn(Trt)-CLEAR acid resin and the peptide chain was elongated by the conventional Fmoc method except for Sec residue. For the introduction of Sec, Fmoc-Sec(MPM)-OH was utilized. Once Sec was introduced to the resin, the solvent was changed from *N*-methyl-2-pyrrolidone (NMP) to dichloromethane (DCM) in order to suppress the β-elimination by decreasing the solvent polarity. Although the peptide

chain elongation succeeded without β-elimination, the fully reduced and unprotected form of the Se-Ins-A could not be obtained as a single product, probably because of facile formation of either the intrachain Se-S or S-S linkage at various positions. Thus, the obtained peptide was cleaved from the resin by TFA and simultaneously protected by 2,2'-dipyridyldisulfide (DPDS) to give the Se-Ins-A (**1**) as a 2-pyridylsulfanyl (Pys)-protected derivative with one Se-S linkage in 9% yield (Figure 2-2).

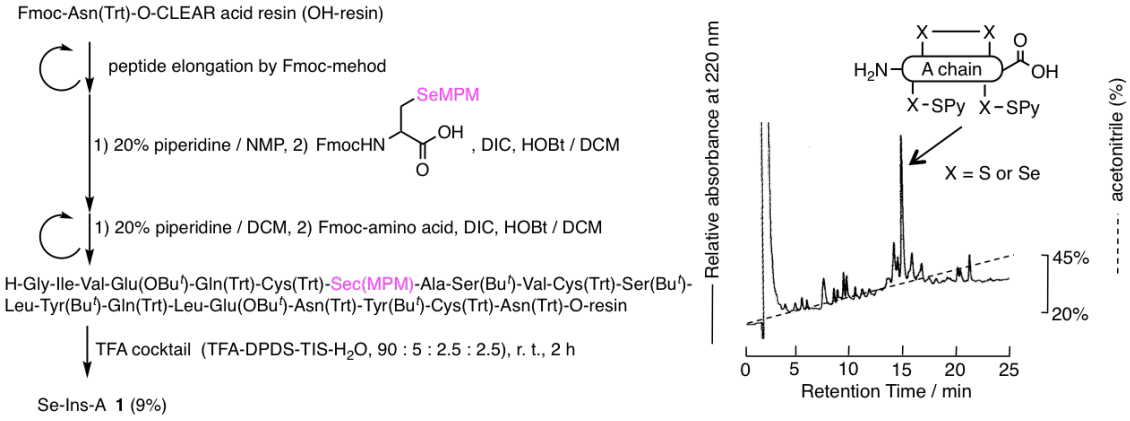

Figure 2-2. Preparation of Se-Ins-A and the RP-HPLC profiles. Elution conditions: column, Mightysil RP-18 GP II (4.6 × 150 mm, Kanto Chemical, Japan) at the flow rate of 1 ml/ min; eluent, A, 0.1% TFA, B, acetonitrile containing 0.1% TFA.

According to reverse-phase (RP) HPLC analysis, **1** contained a single component, but the exact position of the Se-S linkage could not be determined. To obtain the Se-Ins-B, we also synthesized it by Fmoc method in the same manner. After the peptide chain assembly, the obtained peptide was cleaved from the resin by TFA and the crude was then treated under the high-TfOH condition.[16] After purification by RP-HPLC, desired Se-Ins-B (**2**) with intramolecular selenenyl sulfide was obtained in 6% yield (Figure 2-3).

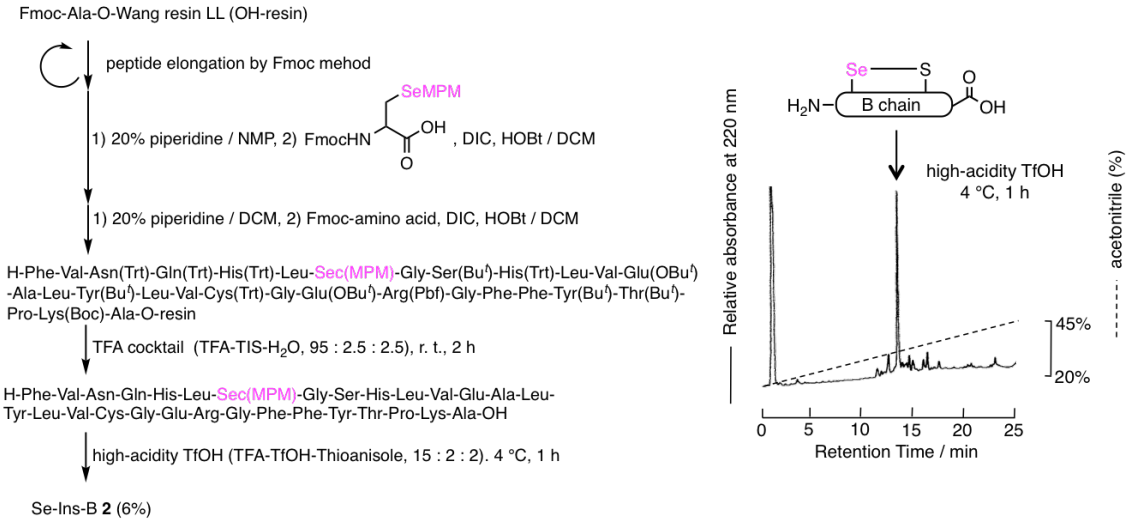

Figure 2-3. Preparation of Se-Ins-B and the RP-HPLC profiles. Elution conditions are the same as those of Figure 2-2.

The oxidative chain assembly was initiated by treatment of the 1:1 mixture of **1** and **2** with DTT (four equivalents with respect to the sum of Sec and Cys residues), which according to its redox potential[10] should remove the Pys groups from the Cys residues as well as reduce the Se-S linkages between the Sec and Cys residues, at 4 °C in a sodium bicarbonate buffer solution at pH 10.0 containing 0.75 M urea to prevent aggregation of **1**. The reduced peptides **1** and **2**, with free thiol (SH) and selenol (SeH) functional groups, were subsequently allowed to assemble spontaneously under thermodynamic control by air oxygen. After 1 minute, an aliquot of the solution was taken, and the reaction was quenched with aqueous 2-aminoethyl methanethiosulfonate (AEMTS),[17] which rapidly converted the free SH and SeH groups into $-SSCH_2CH_2NH_3^+$ and $-SeSCH_2CH_2NH_3^+$, respectively. The folding intermediates present in the reaction solution were then analyzed by RP-HPLC (Figure 2-4) and MALDI-TOF-MS analysis.

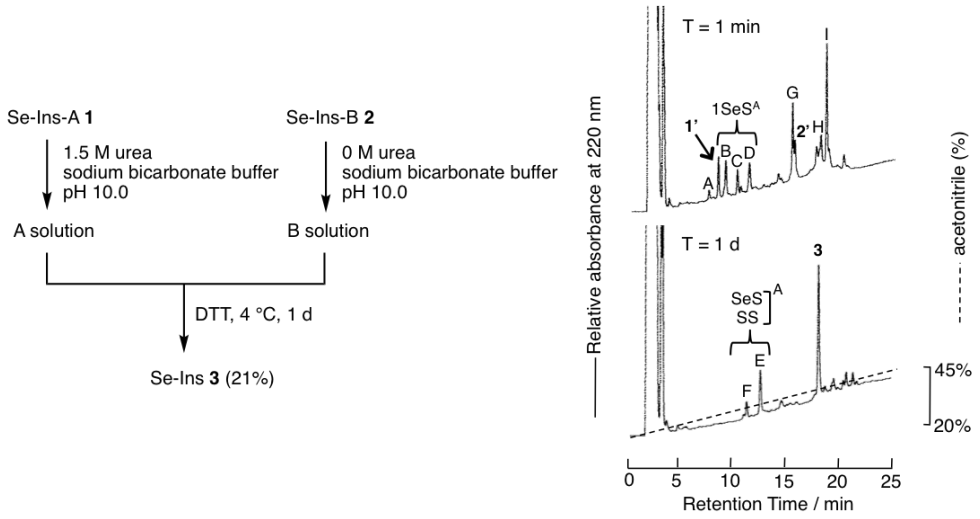

Figure 2-4. Oxidative folding of Se-Ins and the RP-HPLC profiles. Elution conditions are the same as those of Figure 2-2. Folding of Se-Ins was initiated by addition of DTT under the conditions of [**1**] = [**2**] = 184 mM, [DTT] = 4 mM, and [urea]=0.75 M at 4 °C. The symbols indicate. A=fully reduced **1**, B–D = $1SeS^A$, E and F = fully oxidized **1** having one S-S and one Se-S linkage, G=reduced **2**, **3**=native Se-Ins, H=oxidized **2**, I=$[Se\text{-}Ins\text{-}B]_2$ = a homodimer of Se-Ins-B, **1'** = Se-Ins-A-DTT_2 conjugate with two SH groups, and **2'** = Se-Ins-B-DTT_2 conjugate.

Various species, that is, 1SeS of **1** (peaks B, C, and D) and reduced forms of **1** and **2** (peaks A and G), were identified along with oxidized **2** (peak H). A homodimer of Se-Ins-B (peak I) was also observed in a significant amount. After 24 hours, most of the species initially observed converged to the major product (**3**), concomitantly generating the fully oxidized **1** (peaks E and F) and some unidentified species as minor products (Figure 2-4). The major product was then isolated and characterized as correctly folded selenoinsulin (Se-Ins) by single-crystal X-ray analysis. The yield of the isolated product was up to 21 % as determined by amino-acid analysis. By optimizing the reaction conditions (see experimental section), the yield could be slightly increased to 24 % by use of TCEP, instead of DTT, as a reductant to initiate the chain-assembly reaction. The yield was further increased to 27% when the chain assembly was carried out in the presence of less urea (0.5 M). It is worth noting that this preparation is the first example to show that the Cys-to-Sec substitution method[10–13] is effective for the oxidative folding of heterodimeric proteins. The crystal structure of Se-Ins thus isolated was determined at 1.45 Å resolution by the single anomalous dispersion (SAD) method using anomalous scattering of the selenium atoms (Figure 2-5).

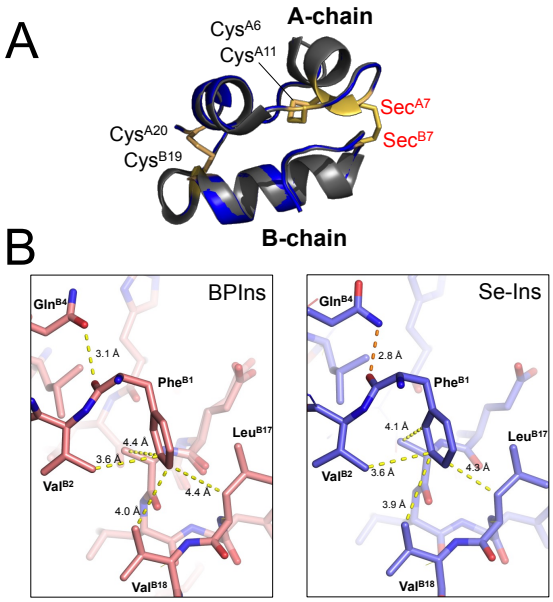

Figure 2-5. A) Superposition of X-ray crystal structures of Se-Ins (blue; PDB code: 5azz) and BPIns (gray; PDB code: 2bn3). B) Comparison of the N-terminal regions of B-chains between BPIns (left panel) and Se-Ins (right panel). Shorter distances between Phe1 and its adjacent residues in the B-chain of Se-Ins suggest that the N-terminal half of the B-chain is more tightly packed in Se-Ins.

Superposition of the X-ray crystal structures of Se-Ins and BPIns (PDB code: 2bn3)[18] demonstrates that Se-Ins has a nearly identical structure to that of BPins with an RMSD of 0.985 Å over all identical atoms in both structures (Figure 2-5 A) though small but significant changes have been induced by the diselenide bond substitution as described below. The two Sec residues form a diselenide bridge at the native position.

Bioassay and IDE degradation assay of Se-Ins

The results of the X-ray analysis suggest that the function of Se-Ins should be comparable to that of native BPIns. This function was confirmed by experiments using cultured cells (Figure 2-6). Upon stimulation by Se-Ins, the phosphorylation levels of Akt and GSK3β were significantly increased in HeLa cells, as by stimulation with BPIns.[19] The data indicate that Se-Ins, as well as the synthetic folded BPIns, retain the physiological activity of insulin even after replacement of an interchain disulfide with a diselenide bridge.

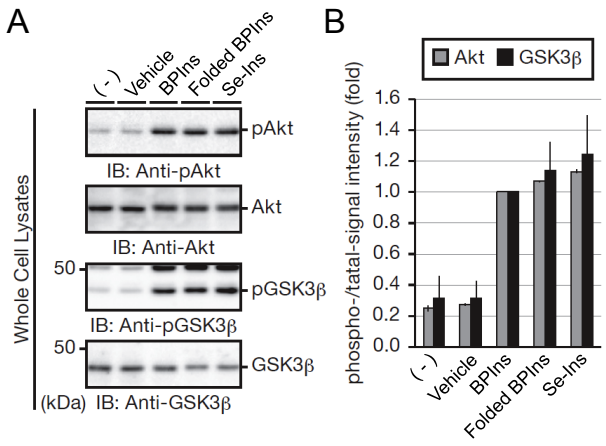

Figure 2-6. Se-Ins or refolded BPIns activates insulin signaling pathway. A) Whole-cell lysates analyzed by immunoblotting. B) The quantitative data. Bars are shown as means ± SEM (n=3). The signal intensity of phosphorylated Akt and phosphorylated GSK3β was normalized against that of Akt and GSK3β, respectively.

The rate of degradation by IDE was subsequently compared for Se-Ins and BPIns. By treating the insulin samples with IDE at 30 °C in a Tris buffer solution at pH8.0, a number of digested peptide fragments were produced as revealed by RP-HPLC analysis (see experimental section). Of interest, the degradation rate of Se-Ins was about eight times slower than that of BPIns in terms of the half lifetime ($\tau_{1/2} \approx 8$ h for Se-Ins vs. ≈ 1 h for BPIns), thus strongly suggesting a long-lasting nature of Se-Ins in vivo (Figure 2-7).

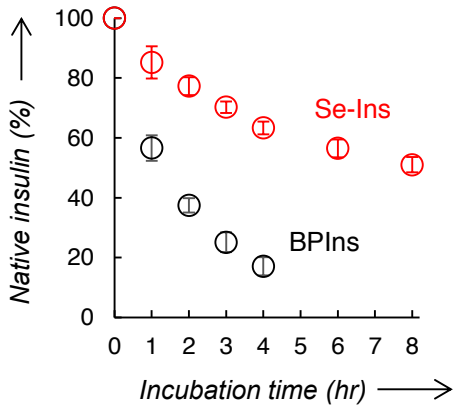

Figure 2-7. Degradation of BPIns and Se-Ins by insulin-degrading enzyme (IDE). Quantitative analyses of native BPIns (black) and Se-Ins (red). Values are the mean ± SD of thee independent experiments.

Though not fully evidenced, the following scenarios could account for this observation. Firstly, replacement of the solvent-exposed Cys^{A7}-Cys^{B7} bridge with a diselenide bond, which is more resistant to reduction and internal bond rotation, should confer extra structural robustness on the insulin fold, thus resulting in the enhanced resistance to IDE degradation. The second scenario could be drawn based on the local stabilization around the N- and C-terminal regions of Se-Ins. The crystal structure of the IDE-Ins complex showed that the N- and C-terminal regions of each chain are partially unfolded at the initial step of the degradation by IDE.[20] Local conformational changes induced by the Se-Se bond substitution enhances the interactions between the N-terminus (Phe^{B1}) and its neighboring residues in the Se-Ins B-chain (Figure 2-5 B). We also note that water-mediated interactions involving Thr^{B27}, Ile^{A2}, and Tyr^{A19} are further stabilized at the interface of the N-terminal region of the A-chain and the C-terminal region of the B-chain in Se-Ins (see experimental section). Such local stabilization found in the N- and C-terminal regions likely allows Se-Ins to prevent the partial unfolding by IDE to a significant extent.

Summary of this work

In summary, we have applied the Cys-to-Sec substitution methodology to chain assembly/folding of BPIns and have succeeded in the synthesis of a new [$C7U^A$,$C7U^B$] variant of insulin in reasonable yields upon isolation (up to 27 %). The molecular structure, determined by X-ray crystallography, reveals that the replacement of the Cys^{A7}-Cys^{B7} bridge with a diselenide bond exerts only a marginal effect on the native insulin structure. In accord with this, Se-Ins exhibited a bioactivity comparable to that of BPIns in the phosphorylation assays of Akt and GSK3β using HeLa cells. Meanwhile, the degradation rate by IDE was greatly decreased ($\tau_{1/2} \approx 8$ h for Se-Ins vs. ≈ 1 h for BPIns), presumably because of the intrinsic stability of the diselenide bond as well as to the slight but significant changes in the local configuration induced by the replacement. Thus, Se-Ins could be a new class of long-acting insulin analogues for diabetes therapy.

2-3. Experimental section

Materials and Methods

Bovine pancreatic insulin (BPIns) of commercialized product as a standard was

purchased from Sigma Aldrich, Japan and used without purification. AEMTS was synthesized according to the literature methods.[17] All other general reagents were commercially available and used without further purification. *N*-(9-Fluorenylmethoxycarbonyl)-*Se*-(*p*-methoxyphenylmethyl)selenocysteine, Fmoc-Sec(MPM)-OH, was synthesized by the literature method.[21]

General procedure for the synthesis of Se-Ins-A (1) and Se-Ins-B (2)

The SPPS of selenopeptides was carried out by the Fmoc method. Once selenocysteine (Sec; U) was introduced to a resin, solvent was changed from *N*-methyl-2-pyrrolidone (NMP) to dichloromethane (DCM). Completion of the couplings was assessed by the Kaiser test. The obtained Se-Ins-A-resin and Se-Ins-B-resin were treated with TFA cocktail A (TFA-2,2'-dipyridyldisulfide-triisopropylsilane-H_2O, 90 : 5 : 2.5 : 2.5) and TFA cocktail B (TFA-triisopropylsilane-H_2O, 95 : 2.5 : 2.5), respectively, and the mixture was vortexed for 2 h at room temperature. The crude of MPM-protected Se-Ins-B was further treated with high-acidity TfOH (TFA-thioanisole-TfOH = 15:2:2). After purification by RP-HPLC, the obtained selenopeptides, Se-Ins-A (**1**) and Se-Ins-B (**2**), were characterized by MALDI-TOF-MS, RP-HPLC, and amino acid analysis (AAA). The AAA sample was prepared by hydrolyzing the solution with 6 M HCl at 150 °C for 2 h in an evacuated sealed tube. Details of procedures for the synthesis of **1** and **2** are given below.

SPPS of selenoinsulin A-chain (1)

The standard Fmoc-SPPS protocol using *O*-(1H-benzotriazol-1-yl)-*N*,*N*,*N*',*N*'-tetramethyluronium hexafluorophosphate (HBTU) as a condensing reagent was employed. Fmoc-Asn(Trt)-CLEAR acid resin (710 mg, 0.2 mmol) was swelled with NMP for 1 h at room temperature. After removing the NMP, the resin was treated with 20% piperidine/NMP for 5 min with vortex mixing. After the deprotection reaction was repeated with fresh 20% piperidine/NMP for 15 min, the resin was washed with NMP (×5). Fmoc-Cys(Trt)-OBt, which was prepared by mixing Fmoc-Cys(Trt)-OH (470 mg, 0.8 mmol), 0.5 M HBTU/DMF (1.5 mL) and DIEA (0.3 mL, 1.6 mmol) for 3 min, was added to the resin. The mixture was vortexed for 12 min at room temperature. After the coupling, the resin was washed with NMP (×3). The unreacted amino groups were acetylated by using 10% Ac_2O and 5% DIEA in NMP for 5 min. By applying a similar peptide-elongating method, Fmoc-Ala-Ser(Bu^t)-Val-Cys(Trt)-Ser(Bu^t)-Leu-Tyr(Bu^t)-Gln(Trt)-Leu-Glu(OBu^t)-Asn(Trt)-Ty

r(Bu^t)-Cys(Trt)-Asn(Trt)-CLEAR acid resin was yielded. A part of the obtained resin (50 μmol) was treated with 20% piperidine/DCM and then washed with DCM. Fmoc-Sec(MPM)-OBt, which was prepared by mixing Fmoc-Sec(MPM)-OH (64 mg, 130 μmol), DIC (31 μL, 200 μmol) and HOBt (26 mg, 200 μmol) in DCM (0.40 mL) for 20 min at room temperature, was added to the resin, and the mixture was vortexed for 30 min at room temperature. After the coupling, the resin was washed with 50% MeOH/DCM (×1) and DCM (×3). The unreacted amino groups were acetylated by using 10% Ac_2O and 5% DIEA in DCM for 5 min. The resin was washed with DCM (×3) and treated with 20% piperidine/DCM for 5 min. The deprotection reaction was repeated with fresh 20% piperidine/DCM for 15 min. The resin was washed with DCM (×5). Fmoc-Cys(Trt)-OBt, which was prepared by mixing Fmoc-Cys(Trt)-OH (120 mg, 200 μmol), DIC (63 μL, 400 μmol) and HOBt (54 mg, 400 μmol) in DCM (0.80 ml) for 20 minutes at room temperature, was added to the resin, and the mixture was vortexed for 30 min at room temperature. Fmoc-Gln(Trt)-OH (190 mg, 200 μmol), Fmoc-Glu(OBu^t)-OH (93 mg, 200 μmol), Fmoc-Val-OH (70 mg, 200 μmol), Fmoc-Ile-OH (73 mg, 200 μmol), and Fmoc-Gly-OH (62 mg, 200 μmol) were sequentially introduced to the resin by the DIC–HOBt method in DCM. The N-terminal Fmoc group was finally deprotected. The resulting resin was washed with DCM (×3), Et_2O (×3) and dried in vacuo to yield H-Gly-Ile-Val-Glu(OBu^t)-Gln(Trt)-Cys(Trt)-Sec(MPM)-Ala-Ser(Bu^t)-Val-Cys(Trt)-Ser(Bu^t)-Leu-Tyr(Bu^t)-Gln(Trt)-Leu-Glu(OBu^t)-Asn(Trt)-Tyr(Bu^t)-Cys(Trt)-Asn(Trt)-CLEAR acid resin (390 mg). A portion of the obtained resin (50 mg, 6.5 μmol) was treated with TFA cocktail A (1 mL) for 2 h at room temperature. After the removal of TFA by N_2 stream, the deprotected peptide was precipitated with Et_2O, washed with Et_2O (×3) and dried in vacuo. The residue was purified by RP-HPLC to give Se-Ins-A **1** derivative, (H-Gly-Ile-Val-Glu-Gln-Cys-Sec-Ala-Ser-Val-Cys-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Tyr-Cys-Asn-OH)+2Pys (540 nmol, 9%). MALDI-TOF-MS (m/z) found: 2603.43, calcd for $[M+H]^+$: 2602.91. Amino acid analysis: $Asp_{1.95}Ser_{1.75}Glu_{3.75}Gly_1Ala_{0.92}Val_{1.47}Ile_{0.46}Leu_{1.84}Tyr_{1.79}$.

SPPS of selenoinsulin B-chain (2)

By applying similar procedure, the crude resin of Se-Ins-B, H-Phe-Val-Asn(Trt)-Gln(Trt)-His(Trt)-Leu-Sec(MPM)-Gly-Ser(Bu^t)-His(Trt)-Leu-Val-Glu(OBu^t)-Ala-Leu-Tyr(Bu^t)-Leu-Val-Cys(Trt)-Gly-Glu(OBu^t)-Arg(Pbf)-Gly-Phe-Phe-Tyr(Bu^t)-Thr(Bu^t)-Pro-Lys(Boc)-Ala-Wang resin LL (330 mg) was yielded. A portion of the obtained resin (59 mg, 8.9 μmol) was treated with TFA cocktail B (1 mL) for 2 h at room temperature. After

the removal of TFA by N_2 stream, the crude peptide was precipitated with Et_2O, washed with Et_2O (×3) and dried in vacuo. The obtained crude peptide was dissolved in high-acidity TfOH cocktail, and the solution was then incubated at 4 °C for 1 h. The resulting crude peptide was precipitated with Et_2O, washed with Et_2O (×3), dried in vacuo, and purified by RP-HPLC to give oxidized Se-Ins-B **2**,

H-Phe-Val-Asn-Gln-His-Leu-Sec-Gly-Ser-His-Leu-Val-Glu-Ala-

Leu-Tyr-Leu-Val-Cys-Gly-Glu-Arg-Gly-Phe-Phe-Tyr-Thr-Pro-Lys-Ala-OH (540 nmol, 6.1%). MALDI-TOF-MS (m/z) found: 3445.09, calcd for $[M+H]^+$: 3444.61. Amino acid analysis: $Asp_{1.02}Thr_{0.99}Ser_{1.03}Glu_{2.99}Pro_{1.71}Gly_{3.10}Ala_{2.05}Val_{2.79}Leu_{4.06}Tyr_{2.15}Phe_3Lys_{1.49}His_{2.03}Arg_{1.09}$.

Two-chain oxidative folding of Se-Ins (3)

A solution of Se-Ins-A (**1**) (91.8 nmol) in a sodium bicarbonate buffer (25 mM, pH 10.0, 245 μL) containing urea (1.5 M) was mixed with a solution of Se-Ins-B (**2**) (91.8 nmol) in the same buffer (245 μL) without urea. The oxidative folding reaction was initiated by addition of DTT^{red} (0.3 mg, 2 μmol) in a sodium bicarbonate buffer solution (25 mM, pH 10.0, 10 μL) containing urea (0.75 M) to the solution containing **1** and **2**. The mixture was incubated at 4 °C. To monitor the reaction progress, small aliquots (5 μL) were taken from the mixture, quenched by addition of aqueous AEMTS solution (7 mg/mL. 200 μL), diluted with 0.1% TFA (800 μL) in water, and analyzed the HPLC system equipped with a sample solution loop (1 mL) and a Mightysil RP-18 GP II 4.6 × 150 reverse phase column, which was equilibrated with a 80:20 (v/v) mixture of TFA (0.1%) in water (eluent A) and TFA (0.1%) in acetonitrile (eluent B) at a flow rate of 1 mL/min. After injection of the sample solution (700 μL) onto the HPLC system, a solvent gradient was applied: a ratio of eluent B linearly increased from 20% to 45% in 0–25 min. The folding intermediates produced were detected by the absorbance at 220 nm. After completion of the reaction, generated Se-Ins was isolated through a C18 column by HPLC and lyophilized. The Se-Ins **3** was obtained as white powder (19.1 nmol, 21%). MALDI-TOF-MS (m/z) found: 5827.7, calcd for $[M+H]^+$: 5827.3. Amino acid analysis: $Asp_{3.14}Thr_{0.91}Ser_{2.57}Glu_{6.34}Gly_{4.14}Ala_{3.25}Val_{5.02}Ile_{0.62}Leu_{6.19}Tyr_{4.12}Phe_3Lys_{1.64}His_{1.99}Arg_{1.00}$. The synthesis was repeated at least three times. The yields were 18 ± 3% under the same conditions and 12 to 27% under various reaction conditions.

Two-chain oxidative folding of wild-type BPIns.

Solutions of the A-chain (R^A) and B-chain (R^B), which were obtained by the

reduction of native BPIns with DTT^{red}, were prepared by dissolving R^A or R^B in a sodium bicarbonate buffer solution (25 mM at pH 10.0) containing urea (1 or 2 M, respectively) and EDTA (1 mM). The solutions were diluted with the same buffer solution to adjust the concentrations. The solutions were chilled at 4 °C for 15–30 min. The folding was initiated by vigorously mixing the R^A solution (75 μL) with the R^B solution (75 μL) for 15 s, and the mixture was immediately diluted with a sodium bicarbonate buffer solution (25 mM at pH 10.0, 100 μL) containing EDTA (1 mM) at 0°C. After a certain period of time, the aliquot (20 μL) was transferred into an aqueous AEMTS solution (8 mg/mL, 200 μL) chilled at 4 °C in a micro-centrifuge tube. After 15 min, the sample solution was diluted with an aqueous TFA solution (0.1%, 800 μL) and stored at –30 °C. The sample solution (1 mL) was analyzed by the HPLC system equipped with a sample solution loop (1 mL) and a Tosoh TSKgel ODS-100V 4.6 × 150 reverse phase (RP) column, which was equilibrated with a 80:20 (v/v) mixture of TFA (0.1%) in water (eluent A) and TFA (0.1%) in acetonitrile (eluent B) at a flow rate of 1.0 ml/ min. A solvent gradient (i.e., a ratio of eluent B linearly increased from 20% to 36% in 0–15 min, from 36% to 39% in 15–20 min, from 39% to 40% in 20–23 min) was applied. The folding intermediates produced were detected by the absorbance at 280 nm.

X-ray crystal structure analysis of Se-Ins

Se-Ins obtained above was purified by RP-HPLC and lyophilized. The lyophilized Se-Ins was dissolved in 20 mM HCl containing 0.57 mM $ZnCl_2$ to obtain the sample of 9 mg/ml. After centrifuge at 15,000 rpm for 15 min, the supernatant was used for crystallization. Crystals of Se-Ins were grown by sitting-drop vapor diffusion in a mixture of 0.5 M NaCl, 0.1 M Sodium citrate (pH 5.6), and 2% (v/v) ethylene imine polymer at 20 °C. Crystals grew within a few day and were cryo-protected in 0.6 M NaCl, 0.1 M sodium citrate (pH 5.6), 2% (v/v) ethylene imine polymer, and 35% (v/v) ethylene glycol. The X-ray diffraction data were collected on beamlines BL1A and 17A at the Photon Factory and were processed with XDS.[22] The structure was determined by the single anomalous dispersion (SAD) method, using the anomalous scattering of Se atoms in selenocysteine. The substructure determination and initial automated model building were performed with AutoSol in PHENIX.[23] The models were further built with COOT[24] and refined with PHENIX. The final models were validated with MolProbity.[25] Structural figures were prepared with PyMOL.[26]

Cell culture and Insulin stimulation assay

HeLa cells (obtained from ATCC) were grown in Dulbecco's modified Eagle's medium (DMEM) (Nacalai tesque, Kyoto, Japan) supplemented with 10% fetal calf serum. For the insulin stimulation assay, 1×10^5 cells were plated in 35 mm plates, cultured for 24 h and serum-starved with DMEM for additional 16 h. Cells were then stimulated with 1 μM native BPIns, folded BPIns or Se-Ins and incubated at 37 °C for 2.5 min. 10 μM HCl was used as a vehicle control. Whole cells were lysed in 1 × sample buffer and cell pellets were homogenized by sonication and then denatured by boiling. The lysates thus prepared were separated by SDS-PAGE and protein bands were visualized by immunoblotting analyses using *anti*-phosphorylated-Akt (Cell signaling), *anti*-Akt (Cell signaling), *anti*-phosphorylated-GSK3β (Cell signaling) and anti-GSK3β (Chemicon) antibodies. The signal intensities were measured using ImageJ software (NIH).

Degradation of BPIns and Se-Ins by insulin-degrading enzyme

Insulin-degrading enzyme (IDE) was purchased by Novoprotein Scientific Inc. Native BPIns (2.1 μM) and Se-Ins (2.1 μM) were mixed with 21 nM IDE and incubated at 30°C in buffer containing 0.1 M Tris-HCl pH 8.0. The digested mixtures were analyzed by RP-HPLC (GL Science) equipped with a COSMOSIL 5C18-AR-II column (Nacalai Tesque), and eluted using a linear gradient of CH_3CN in 0.05% TFA from 5% to 65% at a rate of 1% min^{-1}, with detection at an absorbance of 220 nm.

Table 2-1. Optimization of the oxidative chain assembly conditions

entry	Quantity of Se-Ins-A (**1**) (nmol)	Quantity of Se-Ins-B (**2**) (nmol)	Urea (M)	Reaction initiators	Temp (°C)	pH	Reaction time (h)	Isolated yield (%)
1	91.8 (500 μM)	91.8 (500 μM)	0.75	DTT^{red} (3mM)	4	10	3	14
2	91.8 (500 μM)	91.8 (500 μM)	0.75	DTT^{red} (6 mM)	4	10	6	19
3	91.8 (500 μM)	91.8 (500 μM)	0.75	DTT^{red} (12 mM)	4	10	24	21
4	91.8 (500 μM)	91.8 (500 μM)	0.75	DTT^{red} (24 mM)	4	10	24	16
5	91.8 (200 μM)	91.8 (200 μM)	0.75	GSH:GSSG (10 : 1 mM)	4	10	72	12
6	91.8 (210μM)	91.8 (210μM)	0.75	TCEP (1.4 mM)	4	10	72	24
7	138 (200 μM)	138 (200 μM)	0.5	DTT^{red} (4.8 mM)	4	10	24	19
8	500 (200 μM)	500 (200 μM)	0.5	DTT^{red} (4.8 mM)	4	10	24	27
9	1000 (200 μM)	1000 (200 μM)	0.5	DTT^{red} (4.8 mM)	4	10	24	21

Table 2-2: Data collection and refinement statistics for Se-Ins.

Data collection	
Space group	$I2_13$
Cell dimensions	
$a=b=c$ (Å)	77.99
Resolution (Å)	39-1.45
(outer shell)	(1.48-1.45)
Refs. total/unique	311312/27239
Completeness (%)[a]	100 (99.8)
Redundancy	11.4
$I/\sigma(I)$[a]	12.2 (2.6)
R_{merge} (%)[a, b]	12.4 (93.3)
$CC_{1/2}$[a,c]	0.99 (0.80)
Refinement	
R_{work}/R_{free} (%)[d, e]	16.83/19.16
R.m.s. deviations	
bond length (Å)	0.017
bond angles (°)	1.063
Ramachandran plot	
favored (%)	100
allowed (%)	0
outliers (%)	0

[a]Values in parentheses are for the high-resolution shell.

[b]$R_{merge} = \Sigma| I_h - \langle I_h \rangle | / \Sigma \langle I_h \rangle$, where I_h is the observed intensity and $\langle I_h \rangle$ is the average intensity over symmetry equivalent measurements.

[c]$CC_{1/2}$: The correlation coefficient between average intensities of random half data sets

[d]$R_{work} = \Sigma|F_o - F_c|\Sigma|F_o|$, where F_o and F_c are observed and calculated structure factors, respectively.

[e]R_{free} is the same as R_{work}, but calculated for 5% of randomly chosen reflections that were omitted in the refinement.

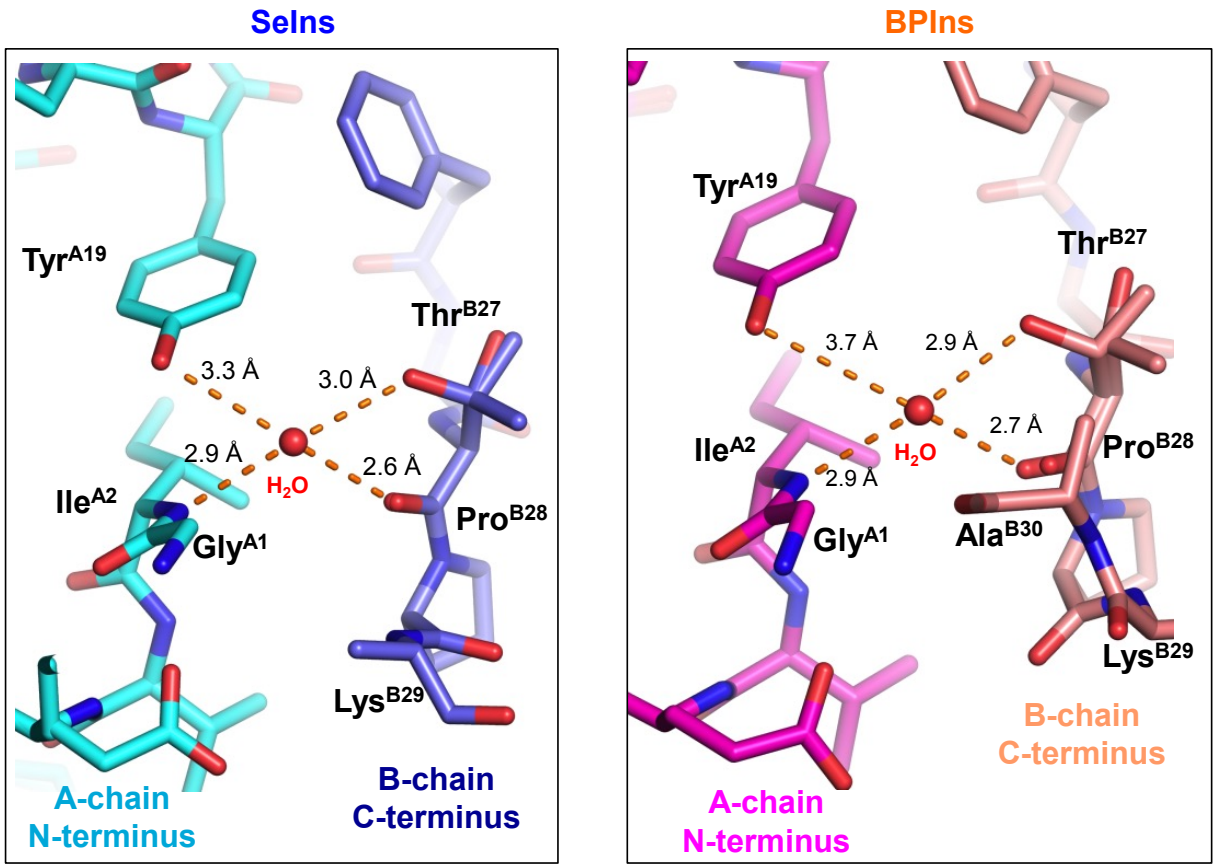

Figure 2-8. Close up view of the interaction between the C terminal region of the B-chain and the N-terminal region of the A- chain in Se-Ins (left panel) and BPIns. Hydrogens bonds are represented by orange broken lines.

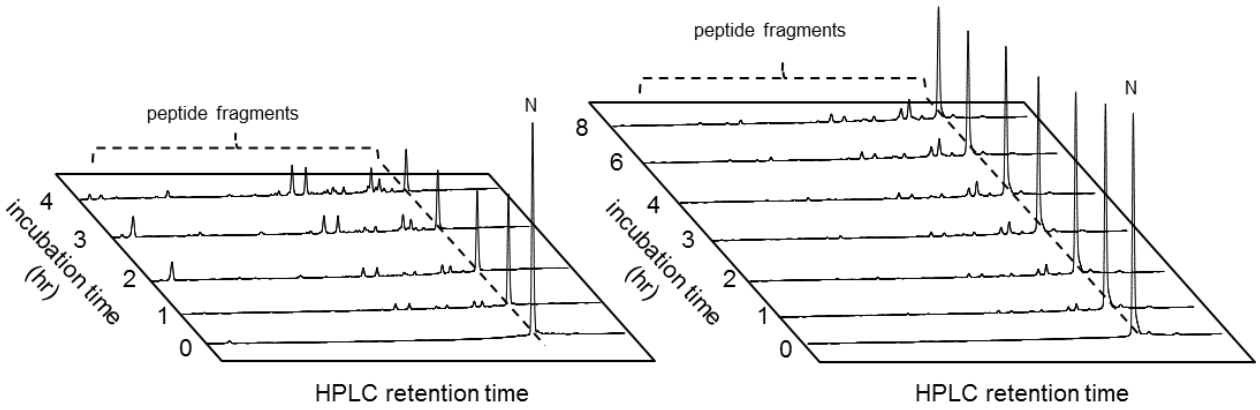

Figure 2-9. Degradation of BPIns (left) and Se-Ins (right) by insulin-degrading enzyme (IDE). Time course of BPIns digestion by IDE. BPIns (2.1 mM) was digested by IDE (21 nM) at 30°C. The resulting samples were analyzed by RP-HPLC using a COSMOSIL $5C_{18}$-AR2 column. "N" indicates the native-fold species that is not degraded by IDE at the indicated incubation time.

2-4. References

[1] A. P. Ryle, F. Sanger, L. F. Smith, R. Kai, *Biochem. J.* **1955**, *60*, 541 – 556.

[2] J. M. Tibaldi, *Am. J. Med.* **2014**, *127*, S25 – S38.

[3] a) T. M. Chapman, S. Noble, K. L. Goa, *Drugs* **2002**, *62*, 1945 – 1981; b) N. A. Reynolds, A. J. Wagstaff, *Drugs* **2004**, *64*, 1957 – 1974 ; c) G. Dailey, *Expert Rev. Endocrinol. Metab.* **2011,** *6*, 527 – 542; d) K. Esposito, A. Capuano, D. Giugliano, *Expert Opin. Biol. Ther.* **2012**, *12,* 1541 – 1550.

[4] a) A. McElduff, *Diabetes Care* **2012**, *35*, 1968 – 1969; b) M. A. Tambascia, F. G. Eliaschewitz, *Diabetol. Metab. Syndr.* **2015**, *7*, 1 – 7; c) H. A. Blair, G. M. *Keating, Drugs* **2016**, *76*, 363 – 374.

[5] a) W. C. Duckworth, R. G. Bennett, F. G. Hamel, *Endocr. Rev.* **1998**, *19*, 608 – 624; b) R. G. Bennett, W. C. Duckworth, F. G. Hamel, *J. Biol. Chem.* **2000**, *275*, 36621 – 36625; c) I. V. Kurochkin, *Trends Biochem. Sci.* **2001**, *26*, 421 – 425.

[6] a) M. A. Leissring, E. Malito, S. Hedouin, L. Reinstatler, T. Sahara, S. O. Abdul-Hay, S. Choudhry, G. M. Maharvi, A. H. Fauq, M. Huzarska, P. S. May, S. Choi, T. P. Logan, B. E. Turk, L. C. Cantley, M. Manolopoulou, W.-J. Tang, R. L. Stein, G. D. Cuny, D. J. Selkoe, *PLoS*

ONE **2010**, *5*, e10504; b) J. P. Maianti, A. McFedries, Z. H. Foda, R. E. Kleiner, X. Q. Du, M. A. Leissring, W.-J. Tang, M. J. Charron, M. A. Seeliger, A. Saghatelian, D. R. Liu, *Nature* **2014**, *511*, 94 – 98; c) J. Charton, M. Gauriot, Q. Guo, N. Hennuyer, X. Marechal, J. Dumont, M. Hamdane, V. Pottiez, V. Landry, O. Sperandio, M. Flipo, L. Buee, B. Staels, F. Leroux, W.-J. Tang, B. Deprez, R. Deprez-Poulain, *Eur. J. Med. Chem.* **2014**, *79*, 184 – 193; d) R. Deprez-Poulain, N. Hennuyer, D. Bosc, W. G. Liang, E. EnØe, X. Marechal, J. Charton, J. Totobenazara, G. Berte, J. Jahklal, T. Verdelet, J. Dumont, S. Dassonneville, E. Woitrain, M. Gauriot, C. Paquet, I. Duplan, P. Hermant, F.-X. Cantrelle, E. Sevin, M. Culot, V. Landry, A. Herledan, C. Piveteau, G. Lippens, F. Leroux, W.-J. Tang, P. Endert, B. Staels, B. Deprez, *Nat. Commun.* **2015**, *6,* 8250; e) J. Charton, M. Gauriot, J. Totobenazara, N. Hennuyer, J. Dumont, D. Bosc, X. Marechal, J. Elbakali, A. Herledan, X. Wen, C. Ronco, H. Gras-Masse, A. Heninot, V. Pottiez, V. Landry, B. Staels, W. G. Liang, F. Leroux, W.-J. Tang, B. Deprez, R. Deprez-Poulain, *Eur. J. Med. Chem.* **2015**, *90*, 547 – 567; f) T. B. Durham, J. L. Toth, V. J. Klimkowski, J. X. C. Cao, A. M. Siesky, J. Alexander-Chacko, G. Y. Wu, J. T. Dixon, J. E. McGee, Y. Wang, S. Y. Guo, R. N. Cavitt, J. Schindler, S. J. Thibodeaux, N. A. Calvert, M. J. Coghlan, D. K. Sindelar, M. Christe, V. V. Kiselyov, M. D. Michael, K. W. Sloop, *J. Biol. Chem.* **2015**, *290*, 20044 – 20059; g) W.-J. Tang, *Cell Press* **2016**, *27*, 24 – 34.

[7] a) R. B. DeMattos, K. R. Bales, D. J. Cummins, J.-C. Dodart, S. M. Paul, D. M. Holtzman, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8850 – 8855; b) R. E. Tanzi, L. Bertram, *Cell* **2005**, *120*, 545 – 555; c) Y. Zhang, D. H. S. Lee, *Neuroscientist* **2011**, *17*, 163 – 173; d) S. Costes, P. C. Butler, *Cell Metab.* **2014**, *20*, 201 – 203.

[8] Very recently, a similar approach to an insulin analogue with enhanced thermodynamic stability by replacing the Cys^{A6}-Cys^{A11} SS linkage with a chemically unreactive SC bond was reported: J. A. Karas, N. A. Patil, J. Tailhades, M.-A. Sani, D. B. Scanlon, B.E. Forbes, J. Gardiner, F. Separovic, J.D. Wade, M.A. Hossain, *Angew. Chem. Int. Ed.* **2016**, *55*, 14743 – 14747; *Angew. Chem.* **2016**, *128*, 14963 – 14967.

[9] J. J. BelBruno, *Heteroat. Chem.* **1996**, *7*, 39 – 43.

[10] The redox potentials of diselenides (E'_0 -381 mV) and seleno- sulfides (E'_0 -326mV) were measured against DTT (E'_0 -323 mV), and that of disulfides (E'_0 -215 mV) was measured against glutathione (E'_0 -215 mV) as determined in a model octapeptide by Moroder and associates: D. Besse, F. Siedler, T. Dierks, H. Kessler, L. Moroder, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 883 – 885; *Angew. Chem.* **1997**, *109*, 915 – 917.

[11] S. Pegoraro, S. Fiori, S. Rudolph-Bçhner, T. X. Watanabe, L. Moroder, *J. Mol. Biol.* **1998**,

284, 779 – 792.

[12] a) S. Pegoraro, S. Fiori, J. Cramer, S. Rudolph-Bçhner, L. Moroder, *Protein Sci*. **1999**, *8*, 1605 – 1613; b) S. Fiori, S. Pegoraro, S. Rudolph-Bçhner, J. Cramer, L. Moroder, *Biopolymers* **2000**, *53*, 550 – 564.

[13] a) L. Moroder, *J. Pept. Sci*. **2005**, *11*, 187 – 214; b) L. Moroder, H.-J. Musiol, M. Gçtz, C. Renner, *Biopolymers* **2005**, *80*, 85 – 97; c) M. Muttenthaler, P. F. Alewood, *J. Pept. Sci*. **2008**, *14*, 1223 – 1239; d) N. Metanis, J. Beld, D. Hilvert, *Patai's Chemistry of Functional Groups* (Ed.: Z. Rappoport), Wiley, **2011**; e) H.-J. Musiol, L. Moroder, *Chimica Oggi/Chem.Today* **2012**, *30*, 14 – 19.

[14] M. Muttenthaler, S. T. Nevin, A. A. Grishin, S. T. Ngo, P. T. Choy, N. L. Daly, S.-H. Hu, C. J. Armishaw, C.-I. A. Wang, R. J. Lewis, J. L. Martin, P. G. Noakes, D. J. Craig, D. J. Adams, P. F. Alewood, *J. Am. Chem. Soc*. **2010**, *132*, 3514 – 3522.

[15] a) Q. Hua, S. N. Gozani, R. E. Chance, J. A. Hoffmann, B. H. Frank, M. A. Weiss, *Nat. Struct. Biol*. **1995**, *2*, 129 – 138; b) Q. Hua, W. Jia, B. H. Frank, N. F. B. Phillips, M. A. Weiss, *Biochemistry* **2002**, *41*, 14700 – 14715.

[16] H. Yajima, N. Fujii, *J. Am. Chem. Soc*. **1981**, *103*, 5867 – 5871.

[17] T. W. Bruice, G. L. Kenyon, *J. Protein Chem*. **1982**, *1*, 47 – 58.

[18] M. H. Nanao, G. M. Sheldrick, R. B. Ravelli, *Acta Crystallogr. Sect. D* **2005**, *61*, 1227 – 1237.

[19] C. M. Taniguchi, B. Emanuelli, C. R. Kahn, *Nat. Rev. Mol. Cell Biol*. **2006**, *7*, 85 – 98.

[20] M. Manolopoulou, Q. Guo, E. Malito, A. B. Schilling, W.-J. Tang, *J. Biol. Chem*. **2009**, 284, 14177 – 14188.

[21] M. Iwaoka, R. Ooka, T. Nakazato, S. Yoshida, S. Oishi, *Chem. Biodiversity* **2008**, *5*, 359 – 374.

[22] W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 125 – 132.

[23] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I.W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R.W. Grosse-Kunstleve, A. J. McCoy, N.W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D* **2010**, *66*, 213 – 221.

[24] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126 – 2132.

[25] V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L.W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D* **2010**, *66*, 12 – 21.

[26] W. L. DeLano, DeLano Scientific LLC, Palo Alto, CA, USA, San Carlos, CA, USA, **2008**.

3. One-Pot Four-Segment Ligation Using Seleno- and Thioesters: Synthesis of Superoxide Dismutase

3-1. Introduction

The synthesis of (glyco)proteins has been efficiently carried out by chemoselective ligation methods. Among them is the native-chemical ligation (NCL) method,[1] developed by Kent and co-workers, which has been widely used. In this method, a peptide with a C-terminal thioester and a peptide with an N- terminal cysteine residue are chemoselectively condensed by an intermolecular thioester exchange reaction, and a subsequent intramolecular *S*-to-*N* acyl shift reaction forms the native peptide bond. As the protection-free ligation is realized, the NCL method has various advantages, such as the simple preparation of the segments and the ease of workup after the ligation.

Recently, selenoesters have been used for the NCL method and demonstrated excellent synthetic results.[2] In particular, even in the case of the extremely sluggish ligation at the Pro-Cys site, selenoesters produced the desired ligation product in good yield, which is due to the good leaving-group ability of the selenol compared to the corresponding thiol.[2a] Payne et al. showed that selenoesters reacts extraordinarily fast with selenocystine to form the ligation product.[2c] By using the difference in the reactivity of the selenoester and thioester surrogates, Melnyk et al. recently realized the one-pot ligation of three segments.[2d] In contrast to these successful achievements, the selenoester has seldom been used as an acyl donor in the direct aminolysis reaction for forming a peptide bond, although it was developed a fairly long time ago.[3]

We have been developing an efficient ligation method for (glyco)protein synthesis called the thioester method, and it is based on the direct aminolysis of the amino- and thiol-protected peptide alkyl thioester by the other segments, using silver ions to activate the thioester.[4] In addition, we found that if the aryl thioester is used, the ligation proceeds in the absence of silver ions.[4c] As the alkyl thioester is stable to aminolysis in the absence of silver ions, the aryl thioester can be selectively ligated with a peptide alkyl thioester, which leads to the development of a one-pot three-segment ligation using the thioester method.[4d] If the selenoester and aryl thioester are sufficiently different in their reactivity for aminolysis, a one-pot four-segment ligation could be realized without requiring any deprotection steps during the ligation (Figure 3-1).

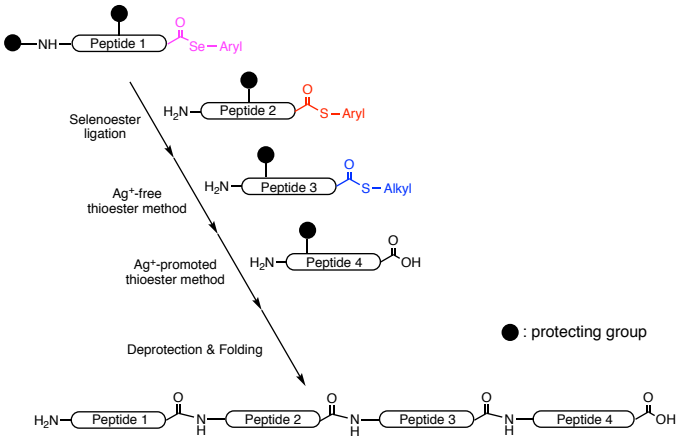

Figure 3-1. One-pot four segment ligation using the thioester method.

The one-pot procedure is important as the purification of the intermediates can be omitted. The purification leads to the loss of the product because of the nonspecific adsorption of the peptide to the column, and it is particularly serious if the target proteins are hydrophobic. In addition, the purified samples have to be lyophilized, which reduces the entire efficiency of the synthetic process. The one-pot ligation was first developed by Kent et al. using three segments,[5a] and since then various one-pot three-segment ligations have been developed.[5] A one-pot four-segment ligation has also been demonstrated in a few syntheses by using the NCL reaction, but in these cases, deprotection of either the terminal amino group or thioester precursor after each ligation reaction is required.[6] Herein, we describe the efficient synthesis of the selenoester by the Fmoc-method using the *N*-alkylcysteine (NAC) as the *N*-to-*S* acyl shift device,[7] the analysis of the amiolysis reaction of the peptide selenoester, and a novel one-pot four-segment ligation to obtain the human superoxide dismutase-1 (SOD-1) which is composed of 153 amino-acid residues.

3-2. Results and discussion

Preparation of a peptide aryl selenoester and the model ligation study

We first examined the reactivity difference between the selenoester and aryl thioester using model peptides. As in the case of the peptide thioester, the synthesis of the selenoester by the Fmoc method is difficult because of the sensitivity of the selenoester linkage to the

piperidine used for the Fmoc group removal. Thus, the selenoesters have previously been prepared by the Boc (*tert*-butoxycarbonyl) method,[8] by the thioester-selenoester exchange reaction,[2a] by the Fmoc synthesis on the Cl-Trt-resin and subsequent coupling with the selenophenol,[2c] and by side-chain anchoring to the resin with subsequent selenoester formation at the C-terminus.[2e] We found that the *N*-alkylcysteine (NAC) method,[7] which was developed for the synthesis of the peptide thioester by the Fmoc method, can be similarly applied for the synthesis of the selenoester as shown in Figure 3-2.

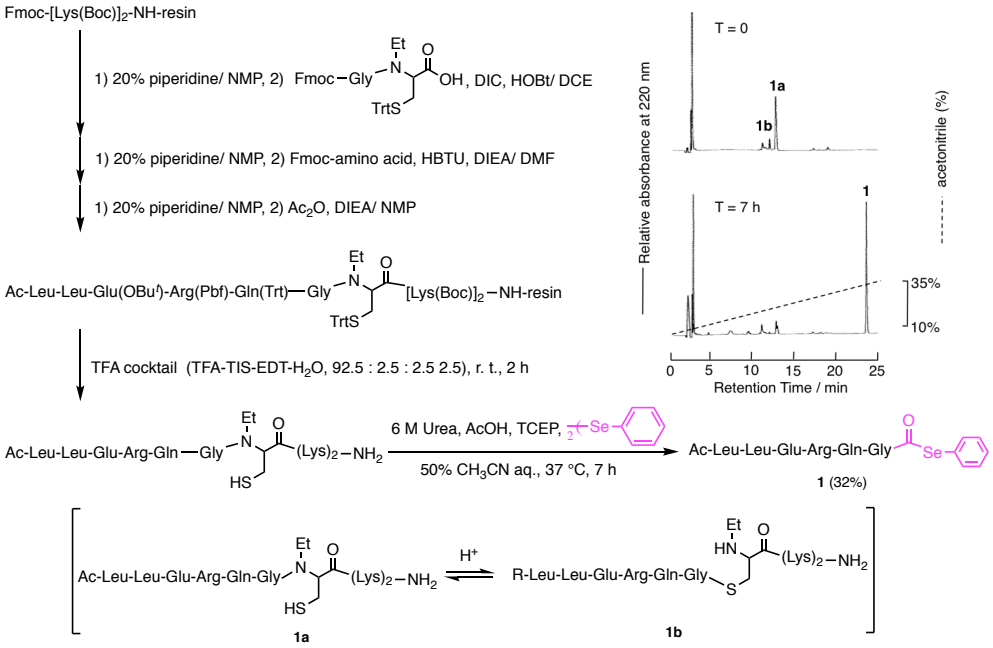

Figure 3-2. Synthesis of selenoester by the NAC method. Elution conditions: column, Mightysil RP-18 GP II (4.6 × 150 mm, Kanto Chemical, Japan) at the flow rate of 1 ml/ min; eluent, A, 0.1% TFA, B, acetonitrile containing 0.1% TFA.

The synthesis was started by the introduction of Fmoc-Gly-(Et)Cys(Trt)-OH to the H-[Lys(Boc)]$_2$-Rink amide MBHA-resin. The peptide chain was elongated by the conventional Fmoc method and the obtained peptide was cleaved from the resin by TFA. The peptide with the NAC was then converted into the selenoester in the presence of diphenyldiselenide and tris(carboxyethyl)-phosphine (TCEP). The reaction was almost complete within 7 hours at 37 °C to give Ac-LLERQG-SeC$_6$H$_5$ (**1**) in 32 % yield upon isolation, thus showing the efficiency of this method. The aryl thioester, LLERQG-SC$_6$H$_4$-*m*-OH (**2**), was also prepared by the NAC method in a 47 % yield (see the experimental section).

We then examined the ligation of **1** and **2**. No reaction occurred by simply mixing the two peptides in dimethyl sulfoxide (DMSO). When 10 equivalents of *N*,*N*-diisopropylethylamine (DIEA) was added to the peptide, the reaction immediately occurred and was complete within 15 minutes to give the desired peptide **3** as shown in Figure 3-3 b.

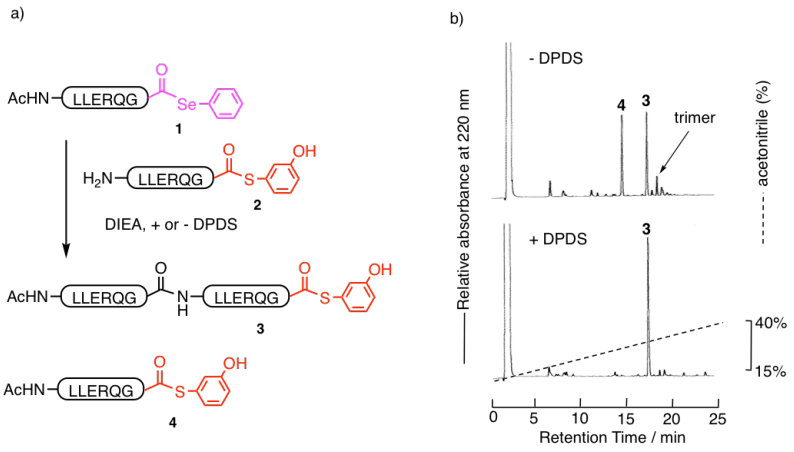

Figure 3-3. a) Ligation of selenophenyl ester and thiophenyl ester. b) HPLC profile of the crude reaction mixture after 15 min. Elution conditions are the same as those of Figure 3-2.

However, Ac-LLERQG-SC_6H_4-*m*-OH (**4**) was also obtained in almost the same amount. In addition, the trimer Ac-$(LLERQG)_3$-SC_6H_4-*m*-OH was also obtained in a small amount. According to these results, the following reaction route can be elucidated: 1. The selenoester **1** reacts with the **2** and selenophenol is liberated. 2. Because of its high nucleophilicity, the selenophenol attacks **2** and liberates the 3-hydroxythiophenol, which attacks **1** to yield **4**. 3. The reaction of the selenoester, derived from **2**, and the thioester **2** forms the product responsible for the trimer formation. These reactions continue until the liberated selenophenol is oxidized to form the diphenyldiselenide. Thus, to suppress these side reactions, the liberated selenophenol has to be immediately trapped after the ligation. We examined several compounds as a scavenger of the selenophenol and found that the 2,2'-dipyridyldisulfide (DPDS) functions well. The reaction was performed again in the presence of excess DPDS, and the desired product **3** was successfully formed in excellent purity as shown in Figure 3-3 b.

Synthesis of SOD-1 by using a one-pot four segment ligation

Having established the conditions for selective activation, the selenoester was then

employed in the one-pot four-segment ligation using human SOD-1[9] as the model. In the initial synthetic plan (see the experimental section), the entire sequence was divided into four segments (see the solid arrows in Figure 3-4).

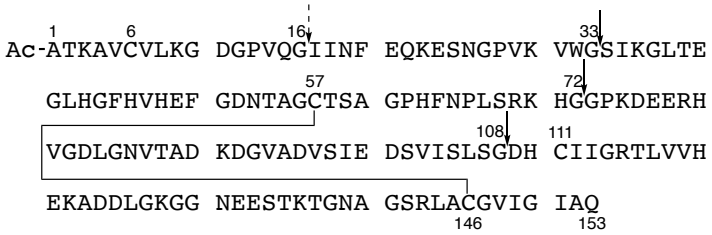

Figure 3-4. Amino-acid sequence of human SOD-1. The arrows indicate the site of the segment coupling. The dotted arrow indicates the additional site of ligation for the entire polypeptide assembly.

Each segment was prepared by solid-phase methods. The side-chain amino group was protected by the 4-pyridylmethoxycarbonyl (*i*Noc) group.[10] Because of its high hydrophilicity, the *i*Noc group does not decrease the solubility of the peptide segments, as has been shown by the efficient syntheses of various glycoproteins.[4d, 11] In addition, it is stable during solid-phase peptide synthesis, deprotection by trifluoroacetic acid (TFA), high-performance liquid chromatography (HPLC) purification, and cleanly cleaved by zinc dust treatment in acetic acid after the ligation. The N-terminal segment (1–33) was prepared as the selenoester **5**, the second segment (34–72) as the aryl thioester **6**, and the third segment (73–108) as the alkyl thioester **7**, all by the NAC method (see the experimental section). The C-terminal segment (109–153) **8** was prepared as the peptide acid starting from the Fmoc-Gln(Trt)-OCH_2-Wang resin. With all the segments in hand, the ligation reaction was attempted according to Figure 3-7 (see the experimental section). However, the ligation between the **5** and **6** did not proceed at all. The analysis of the reaction mixture by HPLC showed that **5** precipitated after the addition of DIEA, thereby retaining the selenoester moiety. Thus, we changed the synthetic strategy by dividing the N-terminal segment into two portions, (1–16) **9** and (17–33) **10**. The revised synthetic route is shown in Figure 3-5.

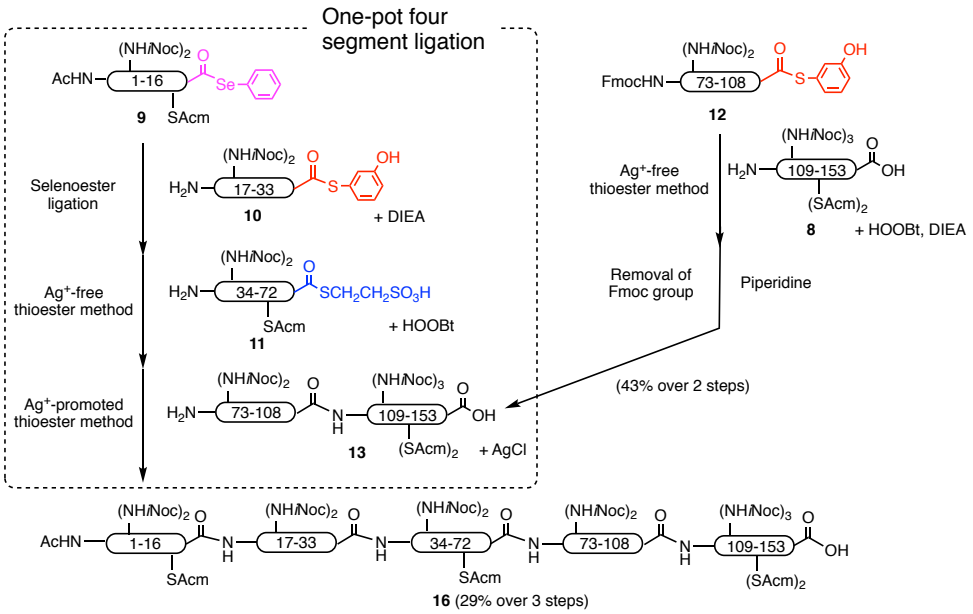

Figure 3-5. Synthesis of SOD-1 by a one-pot four segment ligation.

Following this strategy, the alkyl thioester **11** was newly synthesized. The aryl thioester **12** was also synthesized and used for the ligation with **8** by the silver-ion free thioester method[4c] prior to the one-pot synthesis. The reaction was complete within 10 hours to give the desired product. Piperidine was then added to the reaction mixture to remove the terminal Fmoc group, and the product purified by HPLC to give the C-terminal half **13**.

The one-pot four-segment ligation was then performed. The N-terminal segments **9** and **10** were dissolved in DMSO containing DPDS, and DIEA was added to initiate the ligation. The reaction was complete in one hour and the desired Ac-SOD-1 (1–33), (Acm)(iNoc)$_4$SC$_6$H$_4$-*m*-OH (**14**), was obtained in high purity as shown in Figure 3-6 b.

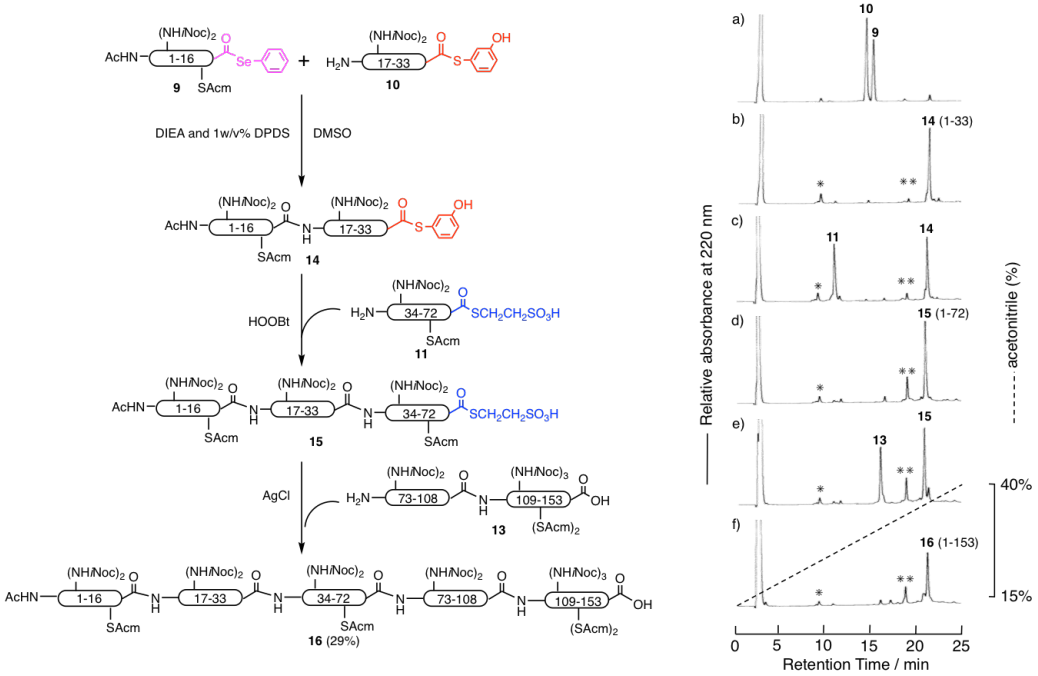

Figure 3-6. Analysis of the progress of the one-pot ligation reaction: a,b) ligation of **9** and **10** at T0 and at 1 h, c,d) ligation of **14** and **11** at T0 and at 3 h, and e,f) ligation of **15** and **13** at T0 and at 12 h. Elution conditions : column, YMC-Pack PROTEIN = RP (4.6 × 150 mm, YMC Co, LTD., Japan) at the flow rate of 1 mL/ min; eluent, A, 0.1% TFA, B, acetonitrile containing 0.1% TFA. Detection of peptide **14** to **16** was performed by ESI mass measurement. The * and ** denote the hydrolyzed products of **9** and **14**, respectively.

Without purification, **11** and 3-hydroxy-1,2,3-benzotriazin-4(3H)-one (HOOBt) were added to initiate the second coupling by the silver-ion-free thioester method. Although the same coupling did not proceed at all in the previous strategy because of the precipitation of the N-terminal selenoester, the combination of the aryl and alkyl thioesters worked (Figure 3-6 d). The reason seems to be the decrease in the pH value of the solution as a result of the addition of HOOBt, which increased the solubility of the aryl thioester. Still, the yield of the ligation product Ac-SOD-1 (1–72), $(Acm)_2$-$(iNoc)_6SC_2H_4SO_3H$ (**15**), was moderate and might indicate that the solubility of the (1–33) segment is marginal. To adjust the ratio of the C-terminal segment **13** to the ligated product (1–72) **15**, the addition of **13** was reduced to 0.5 equivalents relative to the starting segment **10**. The silver ions were then added to initiate the third coupling. The reaction proceeded well to give the polypeptide having the entire sequence of SOD-1. After the HPLC purification, the ligation product **16** was obtained in 29% yield based on the amount of **10** used. The acetamidomethyl (Acm) groups were then removed by silver-ion treatment in aqueous

acetic acid (see experimental section). After 7 hours, the electrospray ionization (ESI) mass analysis of the main peak obtained by HPLC analysis showed that the Acm groups are completely removed to form the peptide **17**. Without isolation of **17**, a solution containing 3-mercaptopropionic acid was added. The precipitated silver salt was removed by centrifugation and Zn dust was added to initiate the *i*Noc-group removal. The reaction also efficiently proceeded within 1 hour (see the experimental section). The mixture was loaded onto an HPLC column to obtain the fully deprotected SOD-1 **18** in 67% yield over two steps. The folding reaction was then performed in the redox buffer containing the reduced and oxidized glutathione for 12 hours to produce the folded SOD-1 **19** (see the experimental section).

A small amount of the **19** was treated with Asp-N in ammonium acetate buffer at pH 5 and 37 °C for 5 hours. The reaction mixture was directly analyzed by ESI mass. The results indicated that Cys^{6} and Cys^{111} have free thiol groups, whereas Cys^{57} and Cys^{146} form a disulfide bond, thus showing that the synthetic SOD-1 is correctly folded (see the experimental section). Finally, zinc and copper ions were coordinated following the reported procedure to obtain the human SOD-1 **20**.[12] The ESI mass analysis showed that the increase in the mass number of the SOD-1 corresponded to the incorporation of the copper and zinc ions (see the experimental section).

The secondary structure of the synthetic SOD-1 in the apo and holo states was analyzed by a circular dichroism (CD) spectroscopy (see the experimental section). The spectrum and their change between the two states agreed well with that of the reported data,[13] thus further supporting the correct folding of the synthetic SOD-1.

The dismutase activity of the synthetic SOD-1 was measured using the method of McCord et al.,[9] wherein the inhibition ratio of the superoxide generation, caused by the oxidation of xanthine by xanthine oxidase, was analyzed. As shown in Figure 3-18 (see the experimental section), the activity was comparable to that of the commercially available SOD-1, thus showing that the synthetic SOD-1 is fully active.

Summary of this work

In conclusion, a peptide selenoester was efficiently prepared by the Fmoc method using the NAC method as the *N*-to-*S* acyl shift component. The selenoester exhibited perfect selectivity for the aryl thioester in the aminolysis reaction in the presence of DIEA in DMSO. Combined with the one-pot three-segment ligation by the thioester method, the one-pot

four-segment ligation was demonstrated to obtain the entire sequence of the SOD-1 composed of 153 amino acid residues, thus showing the efficiency of this method. This method can be easily used for the synthesis of other (glyco)proteins. Along this line, further syntheses will be pursued.

3-3. Experimental section

General information

Fmoc-Gly-(Et)Cys(Trt)-OH and Fmoc-Lys(*i*Noc)-OH were prepared by previously described methods.[4d,7] The ESI mass spectra were recorded using a LCQ DECA XP Plus (Thermo Fisher, MA). The ESI mass spectrum of the Asp-N digest of synthetic human SOD-1 was recorded by Nano frontier (Hitachi, Tokyo) to assign disulfide pairing. Circular dichroism spectrum was measured using a J-820 spectropolarimeter (Jasco, Tokyo). Recombinant human SOD-1 was purchased from Abcam Inc. (Boston, MA) and the dismutase activity was detected by SOD activity detection kit from Wako Pure Chemical Industries (Osaka). The amino acid composition was determined using a LaChrom amino acid analyzer (Hitachi, Tokyo) after hydrolysis with 6 M HCl at 180 °C for 25 min in an evacuated sealed tube. The content of the peptides in the powders was estimated based on the amino acid analysis. The microwave-assisted SPPS was carried out using a Liberty Blue peptide synthesizer (CEM, NC) in 0.1 or 0.25 mmol scale. The brief protocol of 0.25 mmol scale is as follows: (1) Fmoc removal by 10% piperazine 0.1 M HOBt at 75 ºC 15 sec and at 90 ºC 50 sec, (2) DMF washing (4 times), (3) Coupling using Fmoc-amino acid, DIC and HOBt (5 eq. each to the amino groups on resin) at 75 ºC for 15 sec and at 90 ºC for 110 sec, (4) washing with DMF once. For the introduction of His and Cys residue, the coupling was performed at 25 ºC for 120 sec and at 50 ºC for 480 sec. For the introduction of Arg residue, HOBt was substituted to Oxyma and the reaction was performed at 75 ºC for 15 sec and 90 ºC at 110 sec. Double coupling was performed for Arg residue.

Synthesis of model peptides

Ac-LLERQG-SeC_6H_5 1

Fmoc-Rink amide MBHA resin (150 mg, 50 μmol) was treated with 20% piperidine/1-methyl-2-pyrrolidinone (NMP) for 1 min with vortex mixing. The reaction was

repeated with fresh reagent for 5 min. Fmoc-Lys(Boc)-OH (96 mg, 0.2 mmol), dissolved in 0.45 M 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) in DMF (0.42 mL, 0.19 mmol) and *N,N'*-diisopropylethylamine (DIEA) (70 μL, 0.4 mmol), was added to the resin and vortexed for 12 min at 50 °C. Following the same procedure, another Lys residue was introduced and the Fmoc-$[Lys(Boc)]_2$-NH-resin was obtained. After removal of the Fmoc group, a solution of Fmoc-Gly-(Et)Cys(Trt)-OH (70 mg, 0.1 mmol), *N,N'*-diisopropylcarbodiimide (DIC) (31 μL, 0.2 mmol) and 1-hydroxybenzotriazole (HOBt) (29 mg, 0.2 mmol) in 1,2-dichloroethane (DCE) (0.4 mL) was added to the resin, and the mixture was vortexed overnight at room temperature. To the obtained Fmoc-Gly-(Et)Cys(Trt)-Lys(Boc)-Rink amide MBHA resin, peptide chain was elongated by microwave-assisted peptide synthesizer. After the completion of peptide chain assembly, the obtained resin was divided in half. The terminal amino group of the resin (25 μmol) was acetylated by 10% Ac_2O and 5% DIEA in NMP for 10 min at room temperature. The resulting resin was washed with MeOH (x3) and dried in vacuo to yield Ac-Leu-Leu-Glu(OBu^t)-Arg(Pbf)-Gln(Trt)-Gly-(Et)Cys(Trt)-$[Lys(Boc)]_2$-NH-resin (170 mg, 25 μmol). A part of the resin (13 mg, 1.8 μmol) was treated with the TFA cocktail (TFA-triisopropylsilane-H_2O- 1,2-ethanedithiol, 92.5 : 2.5 : 2.5 : 2.5, 0.2 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 50% MeCN aq. containing 6 M urea and 20% AcOH (0.5 mL). Diphenyldiselenide $(PhSe)_2$ (13 mg, 42 μmol) and tris(2-carboxyethyl)phosphine•HCl (TCEP•HCl) (25 mg, 87 μmol) was added to the reaction mixture and the solution was vortexed for 7 h at 37 °C. Diphenyldiselenide was extracted with ether (x3), and the mixture was purified by RP-HPLC to give the peptide **1** (0.51 mg, 570 nmol, 32%). ESI mass, found: m/z 897.6, calcd for $[M+H]^+$: 897.4. Amino acid analysis: $Glu_{2.32(2)}Gly_{1(1)}Leu_{2.01(2)}Arg_{0.99(1)}$.

H-LLERQG-SC_6H_4-*m*-OH 2

The remaining resin in the synthesis of **1** was washed with MeOH (x3) and dried in vacuo to yield H-Leu-Leu-Glu(OBu^t)-Arg(Pbf)-Gln(Trt)-Gly-(Et)Cys(Trt)-$[Lys(Boc)]_2$-NH-resin (180 mg, 25 μmol). A part of the resin (12.0 mg, 1.7 μmol) was treated with the TFA cocktail (TFA-triisopropylsilane-H_2O-1,2-ethanedithiol, 92.5 : 2.5 : 2.5 : 2.5, 0.2 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream

and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 50% MeCN aq. containing 6 M urea and 5% AcOH (0.5 mL). 3-Hydroxybenzenethiol (25 μL) was added to the reaction mixture and the solution was vortexed for 7 h at 37 °C. 3-Hydroxybenzenethiol was extracted with ether (x3), and the mixture was purified by RP-HPLC to give the peptide **2** (0.66 mg, 800 nmol, 47%). ESI mass, found: m/z 823.6, calcd for $[M+H]^+$: 823.4. Amino acid analysis: $Glu_{2.25(2)}Gly_{1(1)}Leu_{2.03(2)}Arg_{0.99(1)}$.

Ligation study of peptide 1 and 2

Ligation in the absence of 2,2'-dipyridyldisulfide (DPDS)

Peptide **1** (90 μg, 100 nmol) and **2** (83 μg, 100 nmol) were dissolved in DMSO (10 μL) containing DIEA (0.17 μL, 1.0 μmol). The mixture was vortexed for 15 min at room temperature. The reaction was monitored by RP-HPLC as shown in Figure 3-3 (b). After 1 h, the reaction mixture was purified by RP-HPLC to give the product, Ac-$(LLERQG)_2$-SC_6H_4-*m*-OH **3** (36 μg, 23 nmol, 23%). ESI mass, found: m/z 1561.9, calcd for $[M+H]^+$: 1561.8. Amino acid analysis: $Glu_{4.53(4)}Gly_{2(2)}Leu_{3.84(4)}Arg_{1.85(2)}$. Ac-LLERQG-$SC_6H_4$-*m*-OH **4** was also obtained in 27% yield (23 μg, 27 nmol). ESI mass, found: m/z 865.7, calcd for $[M+H]^+$: 865.4. Amino acid analysis: $Glu_{2.25(2)}Gly_{1(1)}Leu_{1.89(2)}Arg_{0.89(1)}$.

Ligation in the presence of DPDS

Peptide **1** (90 μg, 100 nmol) and **2** (83 μg, 100 nmol) were dissolved in DMSO (10 μL) containing 1w/v% DPDS and DIEA (0.17 μL, 1.0 μmol). The mixture was vortexed for 15 min at room temperature. The reaction was monitored by RP-HPLC following ether extraction of DPDS as shown in Figure 3-3 b. After 1h, the reaction mixture was purified by RP-HPLC to give the product, Ac-$(LLERQG)_2$-SC_6H_4-*m*-OH **3** (81 μg, 52 nmol, 52%). ESI mass, found: m/z 1561.9, calcd for $[M+H]^+$: 1561.8. Amino acid analysis: $Glu_{4.48(4)}Gly_{2(2)}Leu_{3.88(4)}Arg_{1.89(2)}$.

Synthesis of the peptide segments of the human superoxide dismutase-1 for initial ligation strategy (four segment strategy)

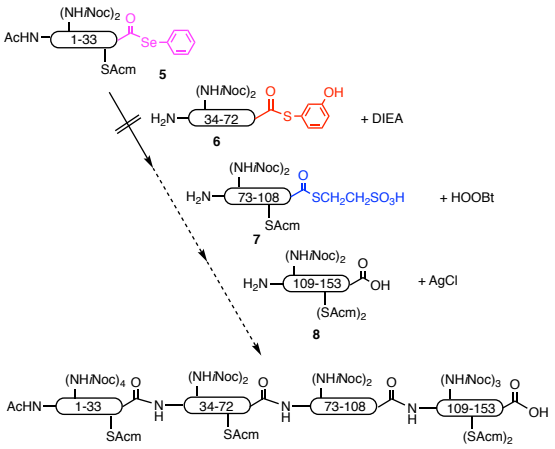

Figure 3-7. Initial strategy to synthesize SOD-1.

Ac-[Lys(*i*Noc)^{3,9,23,30}, Cys(Acm)^{6}]-SOD-1 (1-33)-SeC$_6$H$_5$ 5

Starting from Fmoc-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-Rink amide MBHA-resin (0.2 mmol), peptide chain was elongated manually using the HBTU-DIEA method to obtain Fmoc-Val-Trp(Boc)-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-NH-resin. After removal of the Fmoc group, Fmoc-Lys(*i*Noc)-OH (200 mg, 0.4 mmol), dissolved in 0.45 M HBTU in DMF (0.84 mL, 0.38 mmol) and DIEA (140 µl, 0.8 mmol), was added to the resin and vortexed for 12 min at 50 °C. The peptide chain was then elongated by microwave-assisted peptide synthesizer except for Asp^{11}-Gly^{12} and Lys^{3,9,23}. For the introduction of Asp^{11}-Gly^{12}, Fmoc-Asp(OBu^t)-(Dmb)Gly-OBt, which was prepared by mixing Fmoc-Asp(OBu^t)-(Dmb)Gly-OH (250 mg, 0.4 mmol), DIC (93 µL, 0.6 mmol), and HOBt (82 mg, 0.6 mmol) in DCE (2.0 mL) for 30 min at room temperature, was reacted overnight at room temperature. Lys^{3,9,23} were introduced as described above. After the completion of the chain assembly, the resulting resin was washed with MeOH (x3) and dried in vacuo to yield Ac-Ala-Thr(Bu^t)-Lys(*i*Noc)-Ala-Val-Cys(Acm)-Val-Leu-Lys(*i*Noc)-Gly-Asp(OBu^t)-(Dmb)Gly -Pro-Val-Gln(Trt)-Gly-Ile-Ile-Asn(Trt)-Phe-Glu(OBu^t)-Gln(Trt)-Lys(*i*Noc)-Glu(OBu^t)-Ser(Bu^t) -Asn(Trt)-Gly-Pro-Val-Lys(*i*Noc)-Val-Trp(Boc)-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-NH-resin (1.3 g). A part of the resin (100 mg, 16 µmol) was treated with the TFA cocktail (TFA-triisopropylsilane-H_2O, 95 : 2.5 : 2.5, 2.0 mL) for 2 h at room temperature. The reaction

mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 50% MeCN aq. containing 6 M urea and 20% AcOH (4.0 mL). Diphenyldiselenide (100 mg, 0.32 mmol) and TCEP•HCl (200 mg, 0.70 mmol) were added to the reaction mixture and the solution was vortexed for 19 h at room temperature. After diphenyldiselenide was extracted with ether (x3), the mixture was purified by RP-HPLC to give the peptide **5** (1.2 mg, 280 nmol, 1.8%). ESI mass, found: m/z 1422.1, 1066.9, 853.5, calcd for $[M+3H]^{3+}$: 1421.9, $[M+4H]^{4+}$: 1066.7, $[M+5H]^{5+}$: 853.5. Amino acid analysis: $Asp_{2.92(3)}Thr_{0.81(1)}Ser_{0.91(1)}Glu_{4.03(4)}Pro_{3.20(2)}Gly_{5(5)}Ala_{1.81(2)}Val_{4.59(5)}Ile_{1.24(2)}$ $Leu_{0.94(1)}Phe_{0.89(1)}Lys_{3.93(4)}$.

H-[Lys(*i*Noc)^{36,70}, Cys(Acm)^{57}]-SOD-1 (34-72)-SC_6H_4-*m*-OH 6

Starting from Fmoc-Gly-(Et)Cys(Trt)-$[Lys(Boc)]_2$-Rink amide MBHA-resin (0.2 mmol), peptide chain was elongated manually by the same manner as described for the preparation of peptide **5** to obtain Fmoc-Lys(*i*Noc)-His(Trt)-Gly-(Et)Cys(Trt)-$[Lys(Boc)]_2$-NH-resin. The peptide chain was then elongated by microwave-assisted peptide synthesizer. The resulting resin was washed with MeOH (x3) and dried in vacuo to yield H-Ser(Bu^t)-Ile-Lys(*i*Noc)-Gly-Leu-Thr(Bu^t)-Glu(OBu^t)-Gly-Leu-His(Trt)-Gly-Phe-His(Trt)-Val-His(Trt)-Glu(OBu^t)-Phe-Gly-Asp(OBu^t)-Asn(Trt)-Thr(Bu^t)-Ala-Gly-Cys(Acm)-Thr(Bu^t)-Ser(Bu^t)-Ala-Gly-Pro-His(Trt)-Phe-Asn(Trt)-Pro-Leu-Ser(Bu^t)-Arg(Pbf)-Lys(*i*Noc)-His(Trt)-Gly-(Et)Cys(Trt)-$[Lys(Boc)]_2$-NH-resin (1.50 g). A part of the resin (310 mg, 41 μmol) was treated with the TFA cocktail (TFA-triisopropylsilane-3,6-dioxa-1,8-octanedithiol-H_2O, 92.5 : 2.5 : 2.5 : 2.5, 5.0 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 50% MeCN aq. containing 6 M urea and 5% AcOH (10 mL). 3-Hydroxybenzenethiol (500 μL) was added to the reaction mixture and the solution was vortexed for 15 h at 37 °C. 3-Hydroxybenzenethiol was extracted with ether (x3), and the mixture was purified by RP-HPLC to give the peptide **6** (13 mg, 2.9 μmol, 7.0%). ESI mass, found: m/z 1522.4, 1142.1, calcd for $[M+3H]^{3+}$: 1522.3, $[M+4H]^{4+}$: 1142.0. Amino acid analysis:$Asp_{2.93(3)}Thr_{2.55(3)}Ser_{2.20(3)}Glu_{1.98(2)}Pro_{2.63(2)}Gly_{7(7)}Ala_{1.97(2)}Val_{0.93(1)}Ile_{0.77(1)}Leu_{3.03(3)}Phe_{3.05(3)}Lys_{2.12(2)}His_{4.76(5)}Arg_{0.97(1)}$.

H-[Lys(*i*Noc)^{75,91}]-SOD-1 (73-108)-$SCH_2CH_2SO_3H$ 7

Starting from Fmoc-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-Rink amide MBHA-resin (0.2 mmol), peptide chain was elongated by microwave-assisted peptide synthesizer except for Asp^{92}-Gly^{93} and Lys^{75,91}, which were introduced as described for the preparation of peptide **5**. The resulting resin was washed with MeOH (x3) and dried in vacuo to yield H-Gly-Pro-Lys(*i*Noc)-Asp(OBu^t)-Glu(OBu^t)-Glu(OBu^t)-Arg(Pbf)-His(Trt)-Val-Gly-Asp(OBu^t)-Leu-Gly-Asn(Trt)-Val-Thr(Bu^t)-Ala-Asp(OBu^t)-Lys(*i*Noc)-Asp(OBu^t)-(Dmb)Gly-Val-Ala-Asp(OBu^t)-Val-Ser(Bu^t)-Ile-Glu(OBu^t)-Asp(OBu^t)-Ser(Bu^t)-Val-Ile-Ser(Bu^t)-Leu-Ser(Bu^t)-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-NH-resin (1.3 g). A part of the resin (200 mg, 31.0 μmol) was treated with the TFA cocktail (TFA-triisopropylsilane-H_2O, 95 : 2.5 : 2.5, 3.0 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 25% MeCN aq. containing 6 M urea and 5% AcOH (6.0 mL). Sodium 2-mercaptoethanethiosulfonate (MESNa) (300 mg) was added to the reaction mixture and the solution was vortexed for 20 h at 37 °C. The mixture was purified by RP-HPLC to give the peptide **7** (1.6 mg, 400 nmol, 1.3%). ESI mass, found: m/z 1359.8, 1020.6, calcd for $[M+3H]^{3+}$: 1359.8, $[M+4H]^{4+}$: 1020.1. Amino acid analysis: $Asp_{6.41(7)}Thr_{0.75(1)}Ser_{3.20(4)}Glu_{3.09(3)}Pro_{1.23(1)}Gly_{5(5)}Ala_{1.84(2)}Val_{4.34(5)}Ile_{1.76(2)}Leu_{1.97(2)}Lys_{2.09(2)}His_{0.87(1)}Arg_{0.93(1)}$.

H-[Lys(*i*Noc)^{122,128,136}, Cys(Acm)^{111,146}]-SOD-1 (109-153)-OH 8

Starting from Fmoc-Gln(Trt)-O-Wang resin LL (610 mg, 0.2 mmmol), peptide chain was elongated by microwave-assisted peptide synthesizer, except for the Lys^{122,128,136}, which were introduced using Fmoc-Lys(*i*Noc)-OH as described for the preparation of peptide **5**. After the peptide chain elongation, the resulting resin was washed with MeOH (x3) and dried in vacuo to yield H-Asp(OBu^t)-His(Trt)-Cys(Acm)-Ile-Ile-Gly-Arg(Pbf)-Thr(Bu^t)-Leu-Val-Val-His(Trt)-Glu(OBu^t)-Lys(*i*Noc)-Ala-Asp(OBu^t)-Asp(OBu^t)-Leu-Gly-Lys(*i*Noc)-Gly-Gly-Asn(Trt)-Glu(OBu^t)-Glu(O*t*Bu)-Ser(Bu^t)-Thr(Bu^t)-Lys(*i*Noc)-Thr(Bu^t)-Gly-Asn(Trt)-Ala-Gly-Ser(Bu^t)-Arg(Pbf)-Leu-Ala-Cys(Acm)-Gly-Val-Ile-Gly-Ile-Ala-Gln(Trt)-O-resin (1.75 g). A part of the resin (440 mg, 50.0 μmol) was treated with the TFA cocktail (TFA-triisopropylsilane-H_2O, 95 : 2.5 : 2.5, 9.0 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was purified by RP-HPLC to give the peptide **8** (25 mg,

4.8 μmol, 9.6%). ESI mass, found: m/z 1714.3, 1286.0, 1029.0, calcd for $[M+3H]^{3+}$: 1714.2, $[M+4H]^{4+}$: 1285.9, $[M+5H]^{5+}$: 1028.9. Amino acid analysis: $Asp_{4.88(5)}Thr_{2.48(3)}Ser_{1.46(2)}Glu_{4.75(4)}Gly_{8(8)}Ala_{4.03(4)}Val_{2.49(3)}$ $Ile_{3.28(4)}Leu_{3.01(3)}Lys_{2.80(3)}His_{1.84(2)}Arg_{2.00(2)}$.

Trial ligation of peptide 5 and 6

Peptide **5** (110 μg, 25 nmol) and **6** (110 μg, 25 nmol) were dissolved in DMSO (4 μL) containing 1w/v% DPDS. When DIEA (0.05 μL, 250 nmol) diluted with DMSO (1 μL) was added, most of peptide **5** was precipitated and the reaction did not proceed.

Synthesis of human superoxide dismutase-1 (five segments strategy)

Ac-[Lys(*i*Noc)^{3,9}, Cys(Acm)^{6}]-SOD-1 (1-16)-SeC$_6$H$_5$ 9

Starting from Fmoc-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-Rink amide MBHA-resin (0.1 mmol), peptide chain was elongated manually by the same manner as described for the preparation of peptide **5** to obtain Fmoc-Lys(*i*Noc)-Gly-Asp(OBu^t)-(Dmb)Gly-Pro-Val-Gln(Trt)-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-NH-resin. The peptide chain was elongated by microwave-assisted peptide synthesizer except for Lys^3 residue, which was introduced using Fmoc-Lys(*i*Noc)-OH as described for the preparation of peptide **5**. The terminal amino group of the resin was acetylated by 10% Ac_2O and 5% DIEA in NMP for 10 min at room temperature. The resulting resin was washed with MeOH (x3) and dried in vacuo to yield Ac-Ala-Thr(Bu^t)-Lys(*i*Noc)-Ala-Val-Cys(Acm)-Val-Leu-Lys(*i*Noc)-Gly-Asp(OBu^t)-(Dmb)Gly-Pro-Val-Gln(Trt)-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-NH-resin (540 mg, 0.1 mmol). A part of the resin (110 mg, 20 μmol) was treated with the TFA cocktail (TFA- triisopropylsilane-H_2O, 95 : 2.5 : 2.5, 2.0 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 50% MeCN aq. containing 6 M urea and 20% AcOH (4.0 mL). Diphenyldiselenide (100 mg) and TCEP•HCl (200 mg) were added to the reaction mixture and the solution was vortexed for 1 d at 37 °C. After diphenyldiselenide was extracted with ether (x3), the mixture was purified by RP-HPLC to give the peptide **9** (5.4 mg, 2.6 μmol, 13%). ESI mass, found: m/z 1033.7, calcd for $[M+2H]^{2+}$: 1033.6. Amino acid analysis: $Asp_{1.06(1)}Thr_{0.92(1)}Glu_{1,10(1)}Pro_{1.66(1)}Gly_{3(3)}Ala_{2.04(2)}$ $Val_{2.98(3)}Leu_{1.05(1)}Lys_{2.05(2)}$.

H-[Lys(*i*Noc)^{23,30}]-SOD-1 (17-33)-SC$_6$H$_4$-*m*-OH 10

Starting from Fmoc-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-Rink amide MBHA-resin (0.1 mmol), peptide was elongated manually by the same manner as described for the preparation of peptide **5** to obtain Fmoc-Lys(*i*Noc)-Val-Trp(Boc)-Gly-(Et)Cys(Trt)- [Lys(Boc)]$_2$-NH-resin. The peptide chain was then elongated by microwave-assisted peptide synthesizer except for Lys^{23}, which was introduced using Fmoc-Lys(*i*Noc)-OH as described for the preparation of peptide **5**. The resulting resin was washed with MeOH (x3) and dried in vacuo to yield H-Ile-Ile-Asn(Trt)-Phe-Glu(OBu^t)-Gln(Trt)-Lys(*i*Noc)-Glu(OBu^t)-Ser(Bu^t)-Asn(Trt)-Gly-Pro-Val-Lys(*i*Noc)-Val-Trp(Boc)-Gly- (Et)Cys(Trt)-[Lys(Boc)]$_2$-NH-resin (560 mg). A part of the resin (110 mg, 19 μmol) was treated with the TFA cocktail (TFA-triisopropylsilane-H_2O, 95 : 2.5 : 2.5, 2.0 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 50% MeCN aq. containing 6 M urea and 5% AcOH (2.0 mL). 3-Hydroxybenzenethiol (100 μL) was added to the reaction mixture and the solution was vortexed for 18 h at 37 °C. After 3-hydroxybenzenethiol was extracted with ether (x3), the mixture was purified by RP-HPLC to give the peptide **10** (10 mg, 4.3 μmol, 23%). ESI mass, found: m/z 1162.7, calcd for $[M+2H]^{2+}$: 1162.8. Amino acid analysis: $Asp_{1.91(2)}Ser_{0.82(1)}Glu_{2.90(3)}Pro_{1.91(1)}Gly_{2(2)}Val_{1.83(2)}$ $Ile_{1.06(2)}Lys_{2.05(2)}$.

H-[Lys(*i*Noc)^{36,70}, Cys(Acm)^{57}]-SOD-1 (34-72)-SCH$_2$CH$_2$SO$_3$H 11

A part of the resin in the synthesis of **6** (200 mg, 27 μmol) was treated with the TFA cocktail (TFA-triisopropylsilane-H_2O, 95 : 2.5 : 2.5, 4.0 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 25% MeCN aq. containing 6 M urea and 5% AcOH (8.0 mL). MESNa (400 mg) was added to the reaction mixture and the solution was vortexed for 18 h at 37 °C. The mixture was purified by RP-HPLC to give the peptide **11** (15 mg, 3.3 μmol, 12%). ESI mass, found: m/z 1527.5, 1146.1, calcd for $[M+3H]^{3+}$: 1527.7, $[M+4H]^{4+}$: 1146.0. Amino acid analysis: $Asp_{3.30(3)}Thr_{2.87(3)}Ser_{2.48(3)}Glu_{2.28(2)}Pro_{2.55(2)}Gly_{7(7)}Ala_{2.20(2)}Val_{1.02(1)}Ile_{0.96(1)}Leu_{3.36(3)}Phe_{3.19(3)}Lys_{2.18(2)}His_{5.24(5)}Arg_{1.16(1)}$.

Fmoc-[Lys(*i*Noc)^{75,91}]-SOD-1 (73-108)-SC$_6$H$_4$-*m*-OH 12

Starting from Fmoc-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-Rink amide MBHA-resin (0.2

mmol), peptide chain was elongated by the microwave-assisted peptide synthesizer, except for Asp^{92}-Gly^{93} and $Lys^{75,91}$, which were introduced as described for the preparation of peptide **5**. The resulting resin was washed with MeOH (x3) and dried in vacuo to yield Fmoc-Gly-Pro-Lys(*i*Noc)-Asp(OBu^t)-Glu(OBu^t)-Glu(OBu^t)-Arg(Pbf)-His(Trt)-Val-Gly-Asp(OBu^t)-Leu-Gly-Asn(Trt)-Val-Thr(Bu^t)-Ala-Asp(OBu^t)-Lys(*i*Noc)-Asp(OBu^t)-(Dmb)Gly-Val-Ala-Asp(OBu^t)-Val-Ser(Bu^t)-Ile-Glu(OBu^t)-Asp(OBu^t)-Ser(Bu^t)-Val-Ile-Ser(Bu^t)-Leu-Ser(Bu^t)-Gly-(Et)Cys(Trt)-[Lys(Boc)]$_2$-NH- resin (1.6 g). A part of the resin (310 mg, 39 μmol) was treated with the TFA cocktail (TFA-triisopropylsilane-H_2O, 95 : 2.5 : 2.5, 4.5 mL) for 2 h at room temperature. The reaction mixture was filtered, and the filtrate was concentrated by a N_2 stream and precipitated with ether. The precipitate was washed with ether (x3) and dried in vacuo. The residue was dissolved in 50% MeCN aq. containing 6 M urea and 5% AcOH (12 mL). 3-Hydroxybenzenethiol (600 μL) was added to the reaction mixture and the solution was vortexed for 18 h at 37 °C. After 3-hydroxybenzenethiol was extracted with ether (x3), the mixture was purified by RP-HPLC to give the peptide **12** (6.9 mg, 1.6 μmol, 4.1%). ESI mass, found: m/z 1428.4, calcd for $[M+3H]^{3+}$: 1428.5. Amino acid analysis: $Asp_{6.89(7)}Thr_{0.82(1)}Ser_{3.19(4)}Glu_{3.70(3)}Pro_{2.02(1)}Gly_{5(5)}Ala_{1.98(2)}Val_{4.70(5)}Ile_{1.84(2)}Leu_{2.10(2)}Lys_{2.06(2)}His_{0.93(1)}Arg_{0.98(1)}$.

H-[Lys(*i*Noc)^{75,91,122,128,136}, Cys(Acm)^{111,146}]-SOD-1 (73-153)-OH 13

Peptide **12** (5.6 mg, 1.3 μmol) and **8** (6.7 mg, 1.3 μmol) were dissolved in DMSO (260 μL) containing 3-hydroxy-1,2,3-benzotriazin-4(3H)-one (HOOBt) (6.4 mg, 39 μmol) and DIEA (4.6 μL, 26 μmol), and then the mixture was vortexed for 10 h at room temperature. To remove Fmoc group of the N terminus, piperidine (30 μL) was added to the mixture and voltexed for 30 min at room temperature. The mixture was precipitated with EtOAc/ether 1:1, and then the precipitate was washed twice with ether. The residue was purified by RP-HPLC to yield the peptide **13** (5.2 mg, 570 nmol, 43%). ESI mass, found: m/z 1815.7, 1513.4, 1297.7, 1135.9, calcd for $[M+5H]^{5+}$: 1815.8, $[M+6H]^{6+}$: 1513.3, $[M+7H]^{7+}$: 1297.3, $[M+8H]^{8+}$: 1135.2. Amino acid analysis: $Asp_{12.0(12)}Thr_{3.46(4)}Ser_{4.72(6)}Glu_{6.32(7)}Pro_{1.28(1)}Gly_{13(13)}Ala_{5.99(6)}Val_{7.12(8)}Ile_{4.94(6)}Leu_{5.12(5)}Lys_{5.12(5)}His_{2.91(3)}Arg_{2.99(3)}$.

Synthesis of the human superoxide dismutase-1

Ac-[Lys(*i*Noc)^{3,9,23,30,36,70,75,91,122,128,136}, Cys(Acm)^{6,57,111,146}]-SOD-1 (1-153)-OH 16

Peptide **9** (1.1 mg, 550 nmol) and **10** (1.2 mg, 500 nmol) were dissolved in DMSO (50 μL) containing 1w/v% DPDS and DIEA (0.87 μL, 5.0 μmol), and then the mixture was vortexed for 1 h at room temperature to give peptide **14** (ESI mass, found: m/z 1411.5, 1058.8, 847.5, calcd for $[M+3H]^{3+}$: 1411.6, $[M+4H]^{4+}$: 1058.9, $[M+5H]^{5+}$: 847.4). Without isolation of peptide **14**, peptide **11** (1.4 mg, 300 nmol) in DMSO (30 μL) containing HOOBt (2.5 mg, 15 μmol) and DIEA (0.87 μL, 5.0 μmol) was added, and the mixture was vortexed for 3 h at room temperature to give peptide **15** (ESI mass, found: m/z 1738.0, 1448.7, 1242.1, 1087.1, 966.1, calcd for $[M+5H]^{5+}$: 1738.1, $[M+6H]^{6+}$: 1448.6, $[M+7H]^{7+}$: 1241.8, $[M+8H]^{8+}$: 1086.7, $[M+9H]^{9+}$: 966.1). Without isolation of peptide **15**, peptide **13** (2.3 mg, 250 nmol) in DMSO (40 μL) containing HOOBt (1.5 mg, 9.0 μmol), DIEA (1.1 μL, 6.0 μmol) and a tiny portion of AgCl was added, and the mixture was vortexed for 12 h at room temperature. The mixture was precipitated with EtOAc/ether 1:1, and then the precipitate was washed twice with ether. The crude product was purified by gel filtration chromatography using G3000PW$_{XL}$ (7.5 x 300 mm, Tosoh, Tokyo) in 50% CH_3CN aq. containing 0.1% TFA at a flow rate of 0.7 mL / min to yield polypeptide **16** (2.6 mg, 150 nmol, 29%). ESI mass, found: m/z 1175.3, 1102.0, 1037.2, 979.5, 928.0, 882.1, 839.7, 801.8, 766.7, calcd for $[M+15H]^{15+}$: 1175.5, $[M+16H]^{16+}$: 1102.1, $[M+17H]^{17+}$: 1037.3, $[M+18H]^{18+}$: 979.7, $[M+19H]^{19+}$: 928.2, $[M+20H]^{20+}$: 881.9, $[M+21H]^{21+}$: 839.9, $[M+22H]^{22+}$: 801.8, $[M+23H]^{23+}$: 767.0. Amino acid analysis: $Asp_{17.3(18)}Thr_{6.67(8)}Ser_{7.84(10)}Glu_{12.1(13)}Pro_{4.79(5)}Gly_{25(25)}Ala_{9.60(10)}Val_{12.9(14)}Ile_{6.95(9)}Leu_{9.11(9)}Phe_{4.96(4)}Lys_{10.5(11)}His_{7.43(8)}Arg_{4.00(4)}$.

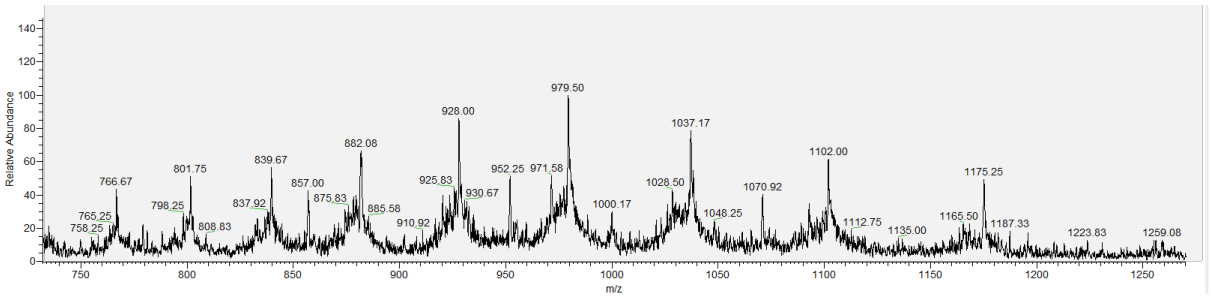

Figure 3-8. ESI mass spectrum of **16**.

Human SOD-1 (1-153)-OH 18

Polypeptide **16** (2.6 mg, 150 nmol) was dissolved in 50% AcOH aq. (0.3 mL) containing AgOAc (4.9 mg, 29.0 μmol), then the mixture was vortexed for 7 h at 37 °C in the

dark. ESI mass analysis showed the complete removal of Acm groups to give peptide **17**. ESI mass, found: m/z 1238.7, 1156.2, 1084.1, 1020.4, 963.8, 912.9, 867.5, 826.3, 788.6, 754.4, calcd for $[M+14H]^{14+}$: 1239.1, $[M+15H]^{15+}$: 1156.5, $[M+16H]^{16+}$: 1084.3, $[M+17H]^{17+}$: 1020.6, $[M+18H]^{18+}$: 963.9, $[M+19H]^{19+}$: 913.3, $[M+20H]^{20+}$: 867.6, $[M+21H]^{21+}$: 826.4, $[M+22H]^{22+}$: 788.9, $[M+23H]^{23+}$: 754.6. Without the isolation of peptide **17**, a solution containing 6 M Gdn•HCl aq. and 15 % (v/v) 3-mercaptopropionic acid (MPA) (1.0 mL) was added to the mixture, and then the formed precipitate was separated after centrifugation. The precipitate was repeatedly washed with same solution and to the combined supernatant (ca. 5.0 ml) was added powdered Zn (ca. 200 mg), which was activated by washing three times with 1 M HCl aq. followed by washing three times with H_2O. The reaction mixture was vigorously vortexed for 1 h at room temperature. After filtration, the mixture was purified by RP-HPLC using YMC-Pack Protein-RP (4.6 x 150 mm, YMC, Kyoto) at a flow rate of 1 mL / min to yield the deprotected polypeptide **18** (1.5 mg, 97 nmol, 67% over two steps). ESI mass, found: m/z 1321.3, 1219.8, 1132.7, 1057.3, 991.2, 933.0, 881.3, 834.8, 793.2, 755.4, 721.2, 689.8, calcd for $[M+12H]^{12+}$: 1321.6, $[M+13H]^{13+}$: 1220.0, $[M+14H]^{14+}$: 1132.9, $[M+15H]^{15+}$: 1057.4, $[M+16H]^{16+}$: 991.4, $[M+17H]^{17+}$: 933.2, $[M+18H]^{18+}$: 881.4, $[M+19H]^{19+}$: 835.0, $[M+20H]^{20+}$: 793.3, $[M+21H]^{21+}$: 755.6, $[M+22H]^{22+}$: 721.3, $[M+23H]^{23+}$: 690.0. Amino acid analysis: $Asp_{17.2(18)}Thr_{6.35(8)}Ser_{7.08(10)}Glu_{12.8(13)}Pro_{5.76(5)}Gly_{25(25)}Ala_{9.43(10)}Val_{12.8(14)}Ile_{7.20(9)}Leu_{8.88(9)}Phe_{4.89(4)}Lys_{10.2(11)}His_{7.61(8)}Arg_{3.85(4)}$.

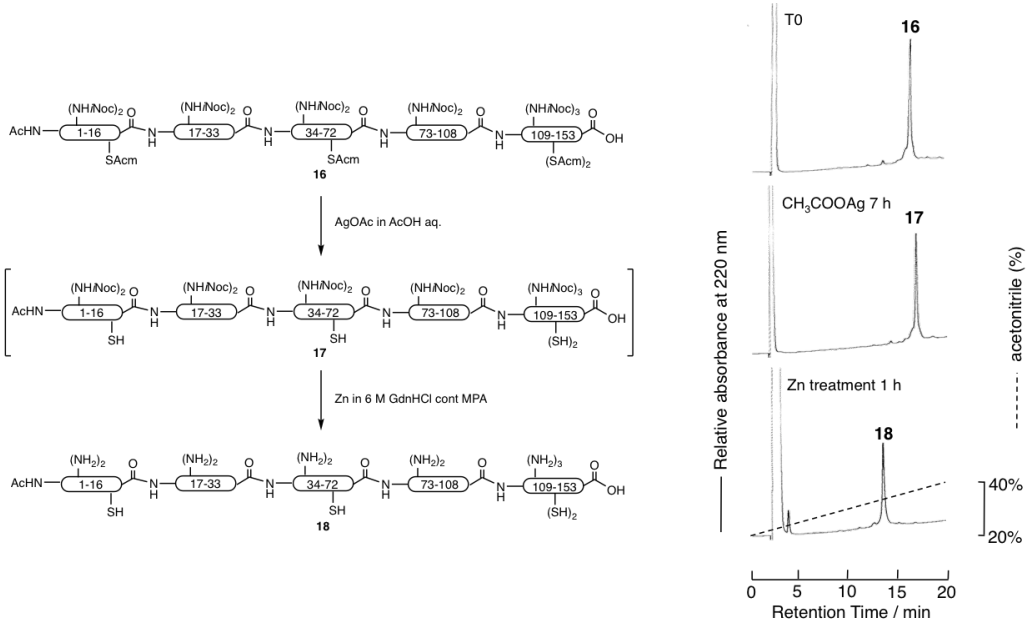

Figure 3-9. Deprotection of SOD-1 and its HPLC profiles. Elution conditions are the same as those of Figure 3-6.

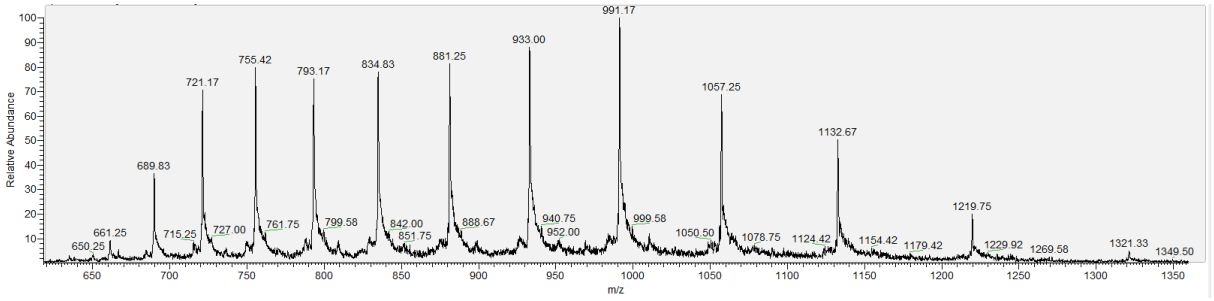

Figure 3-10. ESI mass spectrum of **18**.

Folded human SOD-1 (1-153)-OH 19

Polypeptide **18** (1.5 mg, 97 nmol) was dissolved in 6 M Gdn•HCl aq. (170 μL) containing 0.1 M trishydroxyaminomethane (Tris), 3 mM reduced and 0.5 mM oxidized glutathione, which was adjusted to pH 8.0 by 6 M HCl aq. A buffer (10 mL) containing 0.1 M Tris, 3 mM reduced and 0.5 mM oxidized glutathione, which was adjusted to pH 8.0 by 6 M HCl aq., was added to the mixture was stored for 12 h at 4 °C. The mixture was then acidified by addition of AcOH and purified by RP-HPLC using YMC-Pack Protein-RP (4.6 x 150 mm, YMC, Kyoto) at a flow rate of 1 mL / min to yield folded human SOD-1 **19** (0.93 mg, 59 nmol, 61%). ESI mass, found: m/z 1321.3, 1219.8, 1132.8, 1057.3, 991.3, 933.0, 881.3, calcd for $[M+12H]^{12+}$: 1321.4, $[M+13H]^{13+}$: 1219.8, $[M+14H]^{14+}$: 1132.8, $[M+15H]^{15+}$: 1057.3, $[M+16H]^{16+}$: 991.3, $[M+17H]^{17+}$: 933.0, $[M+18H]^{18+}$: 881.3. Amino acid analysis: $Asp_{17.9(18)}Thr_{6.68(8)}Ser_{7.12(10)}Glu_{12.1(13)}Pro_{4.74(5)}Gly_{25(25)}Ala_{9.90(10)}Val_{13.3(14)}Ile_{7.52(9)}Leu_{9.38(9)}Phe_{4.52(4)}Lys_{10.9(11)}His_{7.92(8)}Arg_{3.98(4)}$.

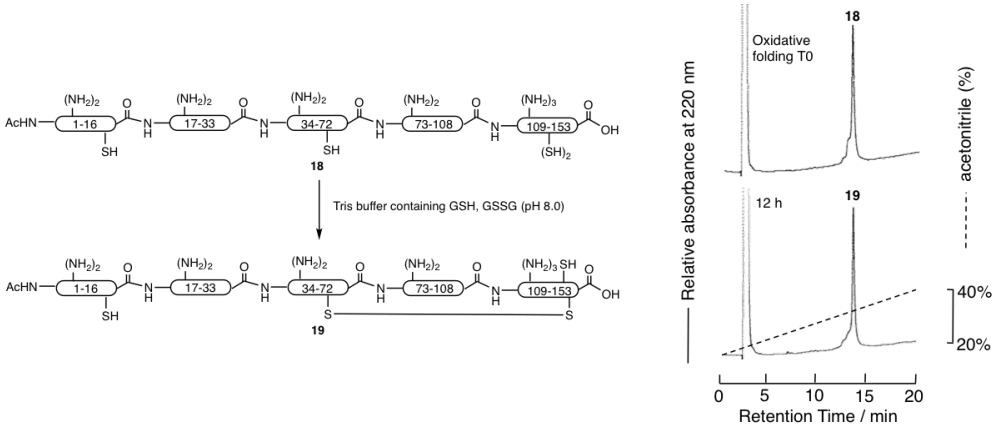

Figure 3-11. Oxidative folding of **18** and its HPLC profiles. The elution conditions are the same as those of Figure 3-6.

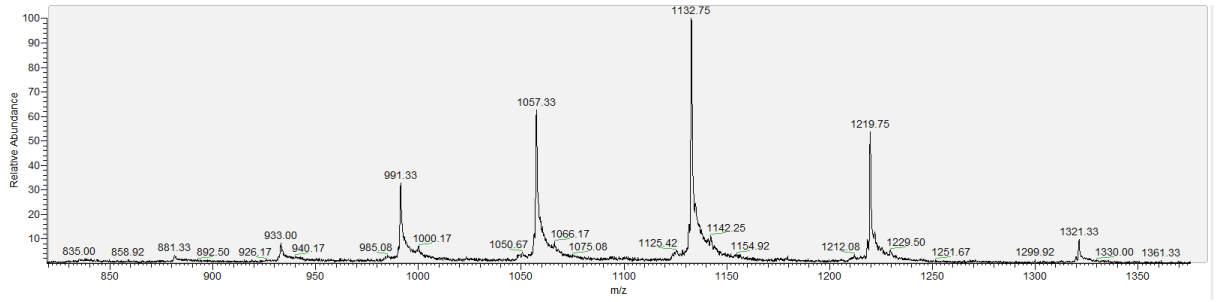

Figure 3-12. ESI mass spectrum of **19**.

Determination of disulfide pairing

Folded human SOD-1 **19** (7.9 μg, 0.5 nmol) was dissolved in 0.1 M ammonium acetate buffer (pH 5.0) at a concentration of 0.16 mg/mL and Asp-N (Asp-N, 1/40 weight of **19**) was added. The solution was kept for 6 h at 37 °C. The digested solution was acidified by addition of AcOH (10 μL) and the mixture was analyzed by ESI mass. The mass numbers corresponding to the theoretical values of the structures (Fragment A to C in Fig 3-12) were observed. Fragment A: found: m/z 1031.6, 516.3, calcd for $[M+H]^{+}$: 1031.6, $[M+2H]^{2+}$: 516.3. Fragment B: found: m/z 564.3, 423.5, 339.0, calcd for $[M+3H]^{3+}$: 564.3, $[M+4H]^{4+}$: 423.5, $[M+5H]^{5+}$: 339.0. Fragment C: found: m/z 1074.3, 895.5, 767.7, 671.8, calcd for $[M+5H]^{5+}$: 1073.7, $[M+6H]^{6+}$: 894.9, $[M+7H]^{7+}$: 767.2, $[M+8H]^{8+}$: 671.4.

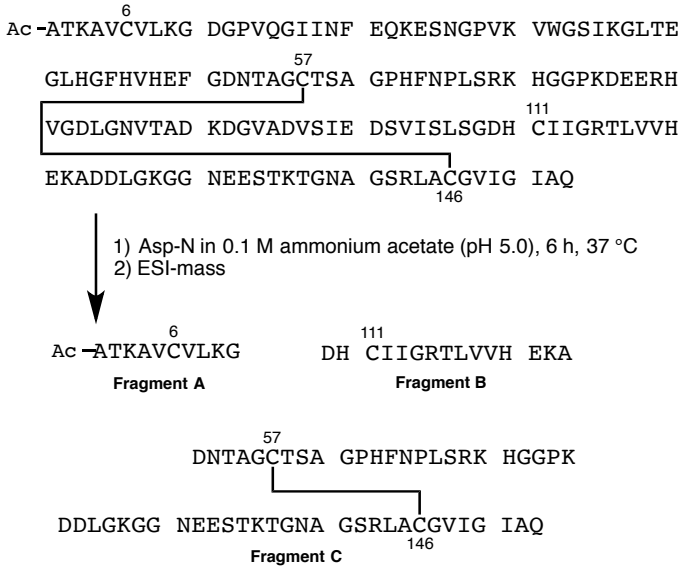

Figure 3-13. Identification of Asp-N digested fragments.

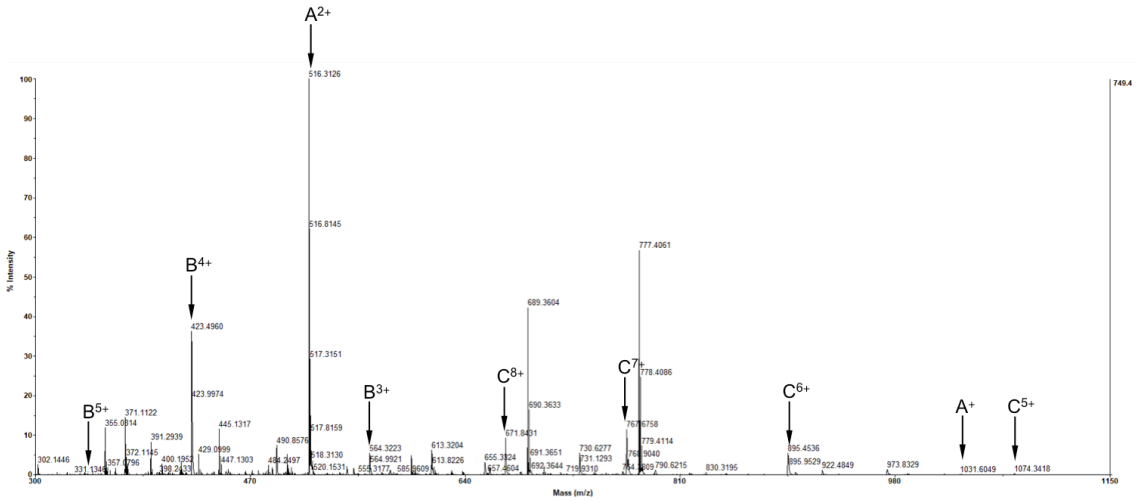

Figure 3-14. ESI mass spectrum of Asp-N digested fragments.

CD spectrum measurement

For the preparation of apo dimer, folded human SOD-1 **19** (48 μg, 3.0 nmol) was dissolved in 20 mM phosphate buffer (250 μL, pH 7.0) and incubated for 2 h at 37 °C. For the preparation of holo dimer, folded human SOD-1 **19** (3 nmol) was also dissolved in 20 mM phosphate buffer (250 μL, pH 7.0) containing 140 μM $ZnCl_2$ and $CuCl_2$, and incubated for 2 h at 37 °C. The CD spectra were measured between 195 nm and 250 nm at 37 °C using a 0.1 cm light path length cell.

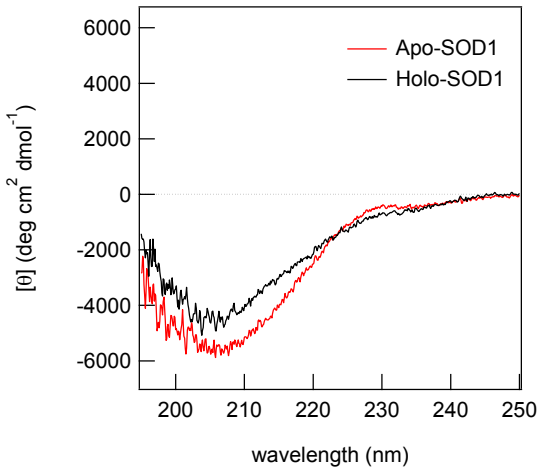

Figure 3-15. CD spectrum of apo and holo human SOD-1.

Reconstitution

Folded human SOD-1 **19** (150 μg, 9.5 nmol) was dissolved in 10 mM ammonium acetate buffer (90 μL, pH 3.0). The solution was dialyzed against 10 mM ammonium acetate buffer (pH 6.3) for 3 h at room temperature. A 0.1 M $ZnCl_2$ aq. (2.0 μL, 200 nmol) was added to the dialyzed solution and incubated for 1 h at 37 °C. A 0.1 M $CuCl_2$ aq. (2.0 μL, 200 nmol) was also added to the mixture then incubated for 2 h at 37 °C. A part of metal ions was dialyzed against same buffer for 2 h at room temperature to obtain SOD-1 **20**. The protein concentration of dialyzed solution was determined by amino acid analysis to be 0.77 mg/mL. For ESI mass measurement, the excess metal ions were further removed by gel filtration chromatography as shown in Figure 3-16. ESI mass found: m/z 3195.3, 2904.8, 2662.9 calcd for $[M+2Cu+2Zn+2H]^{10+}$: 3194.9, $[M+2Cu+2Zn+3H]^{11+}$: 2904.6, $[M+2Cu+2Zn+4H]^{12+}$: 2662.6. Amino acid analysis: $Asp_{17.3(18)}Thr_{6.41(8)}Ser_{7.31(10)}Glu_{12.7(13)}$ $Pro_{6.47(5)}Gly_{25(25)}Ala_{10.1(10)}Val_{12.6(14)}Ile_{7.20(9)}Leu_{9.02(9)}Phe_{4.80(4)}Lys_{10.4(11)}His_{7.59(8)}Arg_{3.92(4)}$.

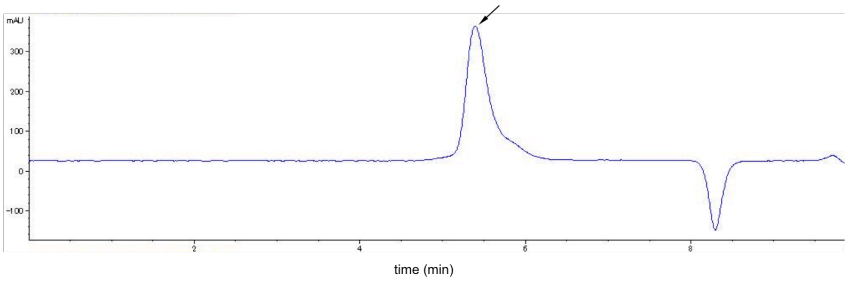

Figure 3-16. Size-exclusion chromatography elution profile of reconstituted SOD-1 **20**. Elution conditions: column, Agilent Bio SEC-3 (3 μm, 100 Å, Agilent, Tokyo) packed UHPLC PEEK column (2.1 x 250 mm, GL Sciences, Tokyo) at the flow rate of 80 μL / min; eluent, 100 mM ammonium acetate containing 5% CH_3CN (pH 7. 0). Detection: 214 nm.

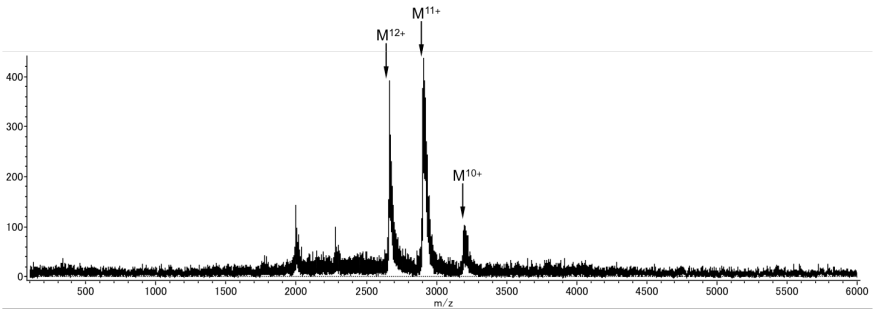

Figure 3-17. ESI mass spectrum of reconstituted SOD-1. The ESI spectrum was recorded by JMS-T100LC (JEPL, Tokyo).

Dismutase activity measurement

The dismutase activity measurement was carried out as described by McCord and Fridovich,[9] using the SOD activity detection kit. The analysis was performed using 96 well plate. Briefly, the solution of SOD-1 or the same volume of distilled water was mixed with xanthine solution and the solution of xanthine oxidase was added. The resultant solutions (S and B*l*, respectively) were shaken for 1 min and kept for 28 min at 37 ºC. SDS solution was added to terminate the reaction and the resultant solution was mixed for 5 min. As a control, the same procedure was performed in the absence of xanthine oxidase at the same time (S-B*l*, B*l*-B*l*, respectively). The absorbance at 590 nm of S, B*l*, S-B*l*, B*l*-B*l* (E_S, E_{Bl}, $E_{S\text{-}Bl}$, $E_{Bl\text{-}Bl}$, respectively) was measured and the SOD activity was calculated by the following equation:

$$\text{SOD activity (Inhibition rate)}=[(E_{Bl} - E_{Bl\text{-}Bl})-(E_S - E_{S\text{-}Bl})]/(E_{Bl} - E_{Bl\text{-}Bl}) \times 100$$

Triplicate experiments were performed simultaneously to obtain each data point.

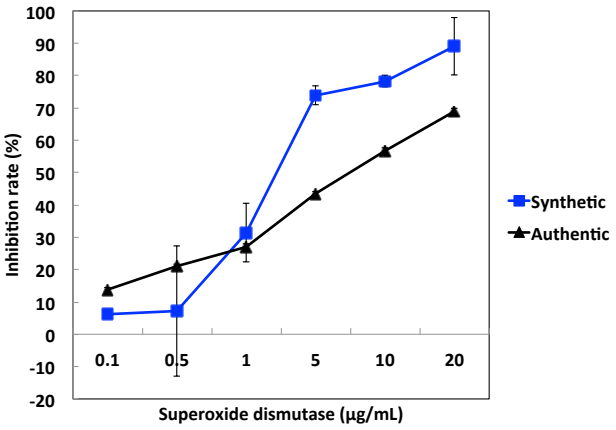

Figure 3-18. Comparison of dismutase activity between synthetic and authentic SOD-1.

3-4. References

[1] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779; b) P.E. Dawson, S.B.H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923 – 960.

[2] a) T. Durek, P. F. Alewood, *Angew. Chem. Int. Ed.* **2011**, *50*, 12042–12045; *Angew. Chem.* **2011**, *123*, 12248–12251; b) L. Raibaut, H. Drobecq, O. Melnyk, *Org. Lett.* **2015**, *17*, 3636–3639; c) N. J. Mitchell, L. R. Malins, X. Liu, R. E. Thompson, B. Chan, L. Radom, R. J. Payne, *J. Am. Chem. Soc.* **2015**, *137*, 14011–14014; d) L. Raibaut, M. Cargoet, N. Ollivier, Y.M.

Chang, H. Drobecq, E. Boll, R. Desmet, J.-C. M. Monbaliu, O. Melnyk, *Chem. Sci.* **2016**, *7*, 2657 – 2665; e) C. C. Hanna, S. S. Kulkarni, E. E. Watson, B. Premdjee, R. J. Payne, *Chem. Commun.* **2017**, *53*, 5424 – 5427.

[3] H.-D. Jakubke, *Chem. Ber.* **1964**, *97*, 2816 – 2828.

[4] a) H. Hojo, S. Aimoto, *Bull. Chem. Soc. Jpn.* **1991**, *64*, 111 – 117; b) S. Aimoto, *Biopolymers* **1999**, *51*, 247–265; c) H. Hojo, Y. Murasawa, H. Katayama, T. Ohira, Y. Nakahara, Y. Nakahara, *Org. Biomol. Chem.* **2008**, *6*, 1808–1813; d) Y. Asahina, S. Kamitori, T. Takao, N. Nishi, H. Hojo, *Angew. Chem. Int. Ed.* **2013**, *52*, 9733 – 9737; *Angew. Chem.* **2013**, *125*, 9915 – 9919.

[5] a) D. Bang, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2004**, *43*, 2534 – 2538; *Angew. Chem.* **2004**, *116*, 2588 – 2592; b) D. Bang, B. L. Pentelute, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2006**, *45*, 3985 – 3988; *Angew. Chem.* **2006**, *118*, 4089 – 4092; c) Z. Tan, S. Shang, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2010**, *49*, 9500 – 9503; *Angew. Chem.* **2010**, *122*, 9690 – 9693; d) K. Sato, A. Shigenaga, K. Tsuji, S. Tsuda, Y. Sumikawa, K. Sakamoto, A. Otaka, *ChemBioChem* **2011**, *12*, 1840–1844; e) R.E. Thompson, X. Liu, N. Alonso-Garcia, P. J. B. Pereira, K. A. Jolliffe, R. J. Payne, *J. Am. Chem. Soc.* **2014**, *136*, 8161 – 8164; f) M. Pan, Y. He, M. Wen, F.-M. Wu, D.-M. Sun, S.-J. Li, L.-H. Zhang, Y.-M. Li, C.-L. Tian, *Chem. Commun.* **2014**, *50*, 5837–5839; g) E. Boll, J.-P. Ebran, H. Drobecq, O. E. Mahdi, L. Raibaut, N. Ollivier, O. Melnyk, *Org. Lett.* **2015**, *17*, 130–133; h) Y. Asahina, T. Kawakami, H. Hojo, *Chem. Commun.* **2017**, *53*, 2114 – 2117.

[6] a) D. J. Boerema, V. A. Tereshko, S. B. H. Kent, *Pept. Sci.* **2008**, *90*, 278 – 286; b) S. Tang, Y.-Y. Si, Z.-P. Wang, K.-R. Mei, X. Chen, J.-Y. Cheng, J.-S. Zheng, L. Liu, *Angew. Chem. Int. Ed.* **2015**, *54*, 5713 – 5717; *Angew. Chem.* **2015**, *127*, 5805 – 5809; c) K. Aihara, K. Yamaoka, N. Naruse, T. Inokuma, A. Shigenaga, A. Otaka, *Org. Lett.* **2016**, 18, 596 – 599.

[7] H. Hojo, Y. Onuma, Y. Akimoto, Y. Nakahara, Y. Nakahara, *Tetrahedron Lett.* **2007**, *48*, 25 – 28.

[8] A. Ghassemian, X. Vila-Farres, P. F. Alewood, T. Durek, *Bioorg. Med. Chem.* **2013**, *21*, 3473 – 3478.

[9] J. M. McCord, I. Fridovich, *J. Biol. Chem.* **1969**, *244*, 6049 – 6055.

[10] D. F. Veber, W. J. Paleveda, Jr., Y. C. Lee, R. Hirschmann, *J. Org. Chem.* **1977**, *42*, 3286 – 3288.

[11] a) Y.Asahina, M. Kanda, A. Suzuki, H. Katayama, Y. Nakahara, H. Hojo, *Org. Biomol. Chem.* **2013**, *11*, 7199–7207; b) Y. Asahina, S. Komiya, A. Ohagi, R. Fujimoto, H. Tamagaki,

K. Nakagawa, T. Sato, S. Akira, T. Takao, A. Ishii, Y. Nakahara, H. Hojo, *Angew. Chem. Int. Ed.* **2015**, *54*, 8226 – 8230; *Angew. Chem.* **2015**, *127*, 8344 – 8348.

[12] Y. Yamazaki, T. Takao, *Anal. Chem.* **2008**, *80*, 8246 – 8252.

[13] M. Ida, M. Ando, M. Adachi, A. Tanaka, K. Machida, K. Hongo, T. Mizobata, M. Y. Yamanaka, Y. Watanabe, K. Nakashima, Y. Kawata, *J. Biochem.* **2016**, *159*, 247 – 260.

4. Summary

4-1. Summary

In chapter 1, the brief history of the research on selenium as an esstrace element was introduced. The discovery of selenocysteine as well as its biological incorporation to proteins, as well as the biological and chemical synthesis of Sec-containing protein were also discussed. Finally, based on these discussion, the motivation to perform this thesis work was described.

In chapter 2, the chemical synthesis of diselenide derivative of BPIns (Se-Ins) was discussed. Sec-substituted A and B chain of BPIns was successfully synthesized by Fmoc method optimized for Sec-containing peptide without β-elimination of the selenium atom, and the subsequent oxidative folding was accomplished in a reasonable yield. The structure of Se-Ins determined by the X-ray crystal analysis revealed that it is quite similar to that of the native BPIns. The Se-Ins exhibited the comparable biological activity with that of the native BPIns in the phosphorylation assay of Akt and GSK3β using HeLa cells. In contrast, the degradation rate by the IDE demonstrated that the Se-Ins has eight times higher resistance than that of native BPIns due to the stability of diselenide bond in the structure. Thus, this study demonstrated that the Sec-substitution increases the protein stability and results in the development of new long-lasting insulin analogue.

In chapter 3, a one-pot four segment ligation was realized by the thioester method utilizing a peptide aryl selenoester. The preparation of peptide aryl selenoester was successfully accomplished by the Fmoc method using the NAC device. Subsequent model ligation revealed that the aminolysis reaction between the aryl selenoester and the aryl thioester proceeds with a perfect selectivity in the presence of DIEA and DPDS in DMSO due to the high reactivity of selenoester. The utility of the one-pot four segment ligation was demonstrated by assembling the entire sequence of human SOD-1 composed of 153 amino acid residues without any deprotection steps during the ligation. The synthetic SOD-1 exhibited the comparable dismutase activity with the commercially available recombinant SOD-1, which proved the success of the synthesis. This method would be generally applicable for the chemical protein synthesis. In addition, it extends the scope of protein accessible by the total chemical synthesis due to its purification- and deprotection-free nature throughout the ligation reaction.

The utility of selenium chemistry to the synthesis of Sec-containing peptide as well as the ligation method demonstrated in this thesis would promote the functional and structural

studies of Sec-containing peptide and protein in future.

4-2. Lists of publication

1) K. Arai[+], T. Toshiki[+], M. Okumura[+], S. Watanabe[+], Y. Amagai, Y. Asahina, L. Molder, H. Hojo, K. Inaba, M. Iwaoka, "Preparation of Selenoinsulin as a Long-Lasting Insulin Analogue" *Angew. Chem. Int. Ed.* **2017**, *56*, 5522–5526; *Angew. Chem.* **2017**, *129*, 5614–5618.
[+] These authors contributed equally to this work.

2) T. Toshiki, T. Andoh, T. Takao, H, Hojo, " One-Pot Four-Segment Ligation Using Seleno- and Thioesters: Synthesis of Superoxide Dismutase" *Angew. Chem. Int. Ed.* **2017**, *56*, 15708–15711; *Angew. Chem.* **2017**, *129*, 15914–15917.

4-3. Related papers

1) T. Toshiki, Y. Urabe, Y. Asahina, H. Hojo, T. Nomura, K. Dedachi, K. Arai, M. Iwaoka, "Model Study Using Designed Selenopeptides on the Importance of the Catalytic Triad for the Antioxidative Functions of Glutathione Peroxidase", *J. Phys. Chem. B.* **2014**, *118*, 492–500.

2) P. Gunasekaran[+], S.-R. Lee[+], S.-M. Jeong, J.-W. Kwon, T. Takei, Y. Asahina, G. Bang, S. Kim, M. Ahn, E. K. Ryu, H. N. Kim, K.-Y. Nam, S. Y. Shin, H. Hojo, S. Namgoong, N.-H. Kim, J. K. Bang, "Pyrrole-Based Macrocyclic Small-Molecule Inhibitors That Target Oocyte Maturation", *ChemMedChem* **2017**, *12*, 580–589.
[+] These authors contributed equally to this work.