

Cysteine-to-serine shuffling using a yeast expression system improves protein secretion: case of a nonglycosylated mutant of miraculin, a taste-modifying protein

Keisuke Ito^{a,b}, Taishi Sugawara^a, Ayako Koizumi^a, Ken-ichiro Nakajima^a, Akiko Shimizu-Ibuka^a, Mitsunori Shiroishi^c, Hidetsugu Asada^c, Takami Yurugi-Kobayashi^c, Tatsuro Shimamura^c, Tomiko Asakura^a, Takumi Misaka^a, So Iwata^{c,d,e}, Takuya Kobayashi^{c,d*}, Keiko Abe^{a*}

a Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan.

b Department of Food and Nutritional Sciences, Graduate School of Nutritional and Environmental Sciences, University of Shizuoka, Yada 52-1, Suruga-ku, Shizuoka 422-8526, Japan.

c Iwata Human Receptor Crystallography Project, ERATO, JST, Konoe-cho, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan.

d Department of Medical Chemistry, Kyoto University Faculty of Medicine, Konoe-cho, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan.

e Division of Molecular Biosciences, Membrane Protein Crystallography Group, Imperial College London, London SW7 2AZ, UK.

* Corresponding author: Takuya Kobayashi

Laboratory of Cell Biology, 3rd Flr., Bldg. A, Graduate School of Medicine, Kyoto University, Yoshida-konoe-cho, Sakyo-ku, Kyoto, 606-8501, Japan

E-mail: t-coba@mfour.med.kyoto-u.ac.jp

Tel: +81-75-753-4386

Fax: +81-75-753-4660

* Corresponding author: Keiko Abe

Laboratory of Biological Function, 3rd Flr., Bldg. 7B, Graduate School of Agricultural and Life Sciences, The University of Tokyo 1-1-1, Yayoi, Bunkyo-ku, Tokyo, 113-8657, Japan

E-mail: aka7308@mail.ecc.u-tokyo.ac.jp

Tel: +81-3-5841-5129

Fax: +81-3-5841-8006

Abstract

Intentional mutagenesis studies should consider the size of the library and the time required for expression screening. Here, we proposed a cysteine-to-serine shuffling mutation strategy (CS shuffling) using a yeast expression system. This strategy of site-directed shuffling mutagenesis of cysteine-to-serine residues aims to identify the cysteine residues that cause protein misfolding in heterologous expression. In the case of a nonglycosylated mutant of the taste-modifying protein miraculin (MCL), which was used here as a model protein, 25% of all constructs obtained from CS shuffling expressed MCL mutant, and serine mutations were found at Cys47 or Cys92, which are involved in the formation of the disulfide bond. This indicates that these residues had the potential to provoke protein misfolding via incorrect disulfide bonding. The CS shuffling can be performed using a small library and within one week, and is an effective screening strategy of soluble protein expression.

Keywords

Protein secretion; mutagenesis; screening method; disulfide bond; cysteine; yeast expression system

Introduction

Soluble protein expression is an important first step during various types of protein studies, such as structural dissection, protein–protein interaction analysis, and antibody development. More effective strategies are needed for improved recombinant-protein expression. Random mutagenesis is commonly used to improve the protein characteristics including expression level (Uchiyama et al. 2000; Koshorreck et al. 2009). This approach, however, is associated with a time-consuming mutant screening and is a trial-and-error process, as the size of the screening libraries used for mutagenesis is, in general, very large. To circumvent these problems, it is necessary to identify a more adequate strategy for mutagenesis.

Cysteine residues and disulfide bonds play important roles in the folding and structural stability of proteins. Replacement of the disulfide-bond-forming cysteine residues with other amino acids leads to destabilization of the conformation (Zavodsky et al. 2001). Whereas the replacement of a cysteine residue is not easy, each disulfide bond has a distinct effect upon the properties of the protein (Kawamura et al. 2008). In the case of heterologous expression, cysteine residues sometimes interfere with the process of protein folding by forming aberrant disulfide bonds. In the case of the *Escherichia coli* expression system, for example, the expression of many eukaryotic-type proteins often results in the

formation of insoluble aggregates as inclusion bodies. In several cases, coexpression with protein disulfide isomerase increases the production of an active protein in its soluble form (Liu et al. 2005; Yuan et al. 2004). This indicates that misfolding via the formation of incorrect disulfide bonds leads to insoluble aggregates. Replacement of the cysteine residues that form the incorrect disulfide bond would be necessary to restore the foldability and stability of each recombinant protein.

The yeast expression system is advantageous as a high-throughput analytical system (Newstead et al. 2007; Sugawara et al. 2009). We previously reported an advanced system that was useful for the expression of proteins bearing various mutations (Ito et al. 2008). This system is based on the cloning of multiple PCR fragments in a single step using homologous recombination. In addition to the high-speed performance, a quality-control system in the yeast secretory pathway prevents the release of misfolded or incompletely folded proteins (Hagihara and Kim 2002; Sakoh-Nakatogawa et al. 2009). Thus, yeast cells have the ability to select folded proteins. These characteristics of the yeast expression system are suitable for mutagenesis screening of cysteine residues.

In this study, we observed that cysteine residues had dual roles of stabilization and misfolding in heterologous expression, and we proposed a cysteine-to-serine shuffling mutation strategy (CS

shuffling) (Fig. 1), based on our previous results (Ito et al. 2008). This strategy aimed to identify the cysteine residues that provoke protein misfolding in the heterologous expression system.

Materials and methods

CS shuffling mutation

The taste-modifying protein miraculin (MCL), a model protein in this study, has a unique property of modifying the taste of sourness into sweetness (Kurihara and Beidler 1969); however, structural information for this protein is not available. The yield of the nonglycosylated MCL (ngMCL) mutant having N42Q/N186Q mutations generated for structural analysis in our *Aspergillus oryzae*-based expression system was poor compared with that of wild-type MCL (Ito et al. 2007). MCL has seven cysteine residues in its monomeric subunit (Fig. 1a), which form three disulfide bonds at the Cys47-Cys92, Cys148-Cys159, and Cys152-Cys155 pairs of residues. Cys138 forms a disulfide bond between different subunits. To identify the optimum cysteine residue to be used for mutation, a CS

shuffling strategy was used (Fig. 1b). The cysteine residues of ngMCL were mutated to serine randomly.

If the replaced cysteine residue was essential for structural foldability, the mutant protein should not pass through the quality-control system in yeast. On the other hand, proteins that were mutated at cysteine residues that cause the misfolding were secreted and detected using on-plate immunodetection. To obtain

PCR fragments corresponding to MCL, the following gene-specific primers were used:

5'-GGTGGTGGTGATTATAAAGATGATGATGATAAAGATTCTGCTCCAAATCCAGTTTGGACAT

T-3', 5'-CCACAAGGTACTTTCGTTTSTCCACCTAGAGTTGTCCAA-3',

5'-TTGGACAACCTCTAGGTGGASAACGAAAGTACCTTGTTG-3',

5'-TCTCTGCTTTCATGCCATSTAGATGGACTTCTTCCAC-3',

5'-GTGGAAGAAGTCCATCTASATGGCATGAAAGCAGAGA-3',

5'-TTTTACAAGTTGGTTTTCTSTCCAACTGTTTSTGGTTCTTSTAAAGTTAAATSTGGTGATGTTG

GTATTAC-3',

5'-GTAAATACCAACATCACCASATTTAACTTTASAAAGAACCASAACAGTTGGASAGAAAAACCA

CTTGTA AAA-3',

and

5'-AAATTGACCTTGAAAATATAAATTTTCCCCCTATTAGAAGTAAACAGTTTTTTGG-3'. Primers

contained a gene-specific region (bold), a homologous region (italic), and a mutation site (underlined).

Cysteine is encoded by TGT or TGC, and serine is encoded by TCT, TCC, TCA, or TCG. Therefore, the second nucleotide coding cysteine (TGT) in ngMCL was degenerated to TST (TCT (serine) or TGT (cysteine)). Four PCR fragments, the MCL signal sequence, and the *Sma*I-linearized pRS426_GAL1 vector were cotransformed into the *S. cerevisiae* strain FGY217 (Ito et al. 2008).

Liquid expression

For soluble expression of MCL, transformants were collected and grown in 10 mL of -Ura selection medium (0.2% yeast synthetic drop-out medium without Ura, 0.7% yeast nitrogen base without amino acids, and 2% glucose) at 30 °C for 24 h. At an OD₆₀₀ of 7.0, cells were harvested by centrifugation, resuspended in an equal volume of expression medium (1% casamino acid, 1.5% yeast nitrogen base without amino acids, and 2% galactose, pH 4.0), and grown at 20 °C for 24 h. The culture supernatant was collected and analyzed by immunodetection using an anti-MCL antibody (Ito et al. 2007). The expression yields were estimated by comparison with the detection level of native MCL (Ito et al. 2007).

On-plate detection of secreted MCL

After transformation, cells were placed on -Ura expression plates (2% agarose, 0.2% yeast synthetic drop-out medium without Ura, 0.7% yeast nitrogen base without amino acids, 0.1% glucose, and 2% galactose) and were covered with a PVDF membrane. Colonies were grown at 30 °C for 24 h and at 20 °C for 72 h. The membrane was analyzed by immunodetection using an anti-MCL antibody and an alkaline phosphatase-linked anti-rabbit IgG antibody (Ito et al. 2007).

Results and discussion

Expression screening using CS shuffling

A small amount of wild-type MCL was secreted, whereas the ngMCL was not secreted (Fig. 2a).

Approximately 2×10^3 colonies of CS-shuffled mutant library per experiment were obtained, and 5×10^2 colonies were identified as ngMCL-expressing clones using on-plate immunodetection (Fig. 2b). Ten strongly positive clones were inoculated into liquid media, and dimeric ngMCL was secreted by these clones (Fig. 2c). The mutations were confirmed by sequencing (Fig. 2d). All clones that were selected had

serine mutations at Cys47 or Cys92, which corresponds to the first disulfide bond of this protein, suggesting that these residues had the potential to lead to protein misfolding via incorrect disulfide bonding in yeast and had a relatively small contribution to structural stability. In contrast, the Cys148 and Cys159 residues, which form the second disulfide bond, were not replaced with serine. This indicates that the disulfide bond strongly contributes to the structural stability of ngMCL and should not be replaced by other amino acids. No significant tendency to a secretable clone was observed at the Cys152 and Cys155 residues, which form the third disulfide bond. These cysteine residues may contribute neither to the stability nor to the misfolding in the yeast expression system.

The contribution of the various cysteine residues to the foldability, stability, and activity of the protein may be independent, and it is very difficult to predict their functional roles. By the CS shuffling strategy, the ngMCL with additional mutation of C47S/C92S was successfully obtained as a dimeric form with a yield of 1.5 mg/L of culture (Fig. 2c). A reduced size of the mutation library improves screening efficiency. In the case of ngMCL, which comprises 191 amino acids, more than 20^{191} random mutagenesis reactions are necessary for screening. This value can be reduced to only 2^6 by focusing on the six cysteine residues. The CS shuffling strategy can be applied for the screening of secretable mutants and will decrease both the mutants library size and the screening time.

Acknowledgments

This study was supported in part by the ERATO Iwata Human Receptor Crystallography Project (JST) (to S.I.), the Targeted Proteins Research Program (to S.I. and T.M.), a Grant-in-Aid for Scientific Research (B) (20370035) (to T.K.), a Grant-in-Aid for Challenging Exploratory Research (22659059) (to T.K.), a grant from the Research and Development Program for New Bioindustry Initiatives (to K.A.), Grants-in-Aid for Scientific Research (S) (to K.A.), and Grants-in-Aid for JSPS Fellows (to K.I.) in Japan.

References

- Hagihara Y, Kim PS (2002) Toward development of a screen to identify randomly encoded, foldable sequences. *Proc Natl Acad Sci U S A* 99:6619–6624
- Ito K, Asakura T, Morita Y, Nakajima K, Koizumi A, Shimizu-Ibuka A, Masuda K, Ishiguro M, Terada T, Maruyama J, Kitamoto K, Misaka T, Abe K (2007) Microbial production of sensory-active miraculin.

Biochem Biophys Res Commun 360:407-411

Ito K, Sugawara T, Shiroishi M, Tokuda N, Kurokawa A, Misaka T, Makyio H, Yurugi-Kobayashi T, Shimamura T, Nomura N, Murata T, Abe K, Iwata S, Kobayashi T (2008) Advanced method for high-throughput expression of mutated eukaryotic membrane proteins in *Saccharomyces cerevisiae*.

Biochem Biophys Res Commun 371:841-845

Kawamura S, Ohkuma M, Chijiwa Y, Kohno D, Nakagawa H, Hirakawa H, Kuhara S, Torikata T (2008) Role of disulfide bonds in goose-type lysozyme. FEBS J 275:2818-2830

Koschorreck K, Schmid RD, Urlacher VD (2009) Improving the functional expression of a *Bacillus licheniformis* laccase by random and site-directed mutagenesis. BMC Biotechnol 9:12

Kurihara K, Beidler LM (1969) Mechanism of the Action of Taste-modifying Protein, Nature 222:1176-1179

Liu Y, Zhao TJ, Yan YB, Zhou HM (2005) Increase of soluble expression in *Escherichia coli* cytoplasm by a protein disulfide isomerase gene fusion system. Protein Expr Purif 44:155-161

Newstead S, Kim H, von Heijne G, Iwata S, Drew D (2007) High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 104:13936-13941

Sakoh-Nakatogawa M, Nishikawa S, Endo T (2009) Roles of protein-disulfide isomerase-mediated disulfide bond formation of yeast Mnl1p in endoplasmic reticulum-associated degradation. *J Biol Chem* 284:11815-11825

Sugawara T, Ito K, Shiroishi M, Tokuda N, Asada H, Yurugi-Kobayashi T, Shimamura T, Misaka T, Nomura N, Murata T, Abe K, Iwata S, Kobayashi T (2009) Fluorescence-based optimization of human bitter taste receptor expression in *Saccharomyces cerevisiae*. *Biochem Biophys Res Commun* 382:704-710

Uchiyama H, Inaoka T, Ohkuma-Soyejima T, Togame H, Shibanaka Y, Yoshimoto T, Kokubo T (2000) Directed evolution to improve the thermostability of prolyl endopeptidase. *J Biochem* 128:441-447

Yuan S, Duan H, Liu C, Liu X, Liu T, Tao H, Zhang Z (2004) The role of thioredoxin and disulfide isomerase in the expression of the snake venom thrombin-like enzyme calobin in *Escherichia coli* BL21 (DE3). *Protein Expr Purif* 38:51-60

Zamyatnin AA (1972) Protein volume in solution. *Prog Biophys Mol Biol* 24 :107-123

Zavodszky M, Chen CW, Huang JK, Zolkiewski M, Wen L, Krishnamoorthi R (2001) Disulfide bond effects on protein stability: Designed variants of *Cucurbita maxima* trypsin inhibitor-V. *Protein Sci* 10:149–160

Figure Legends

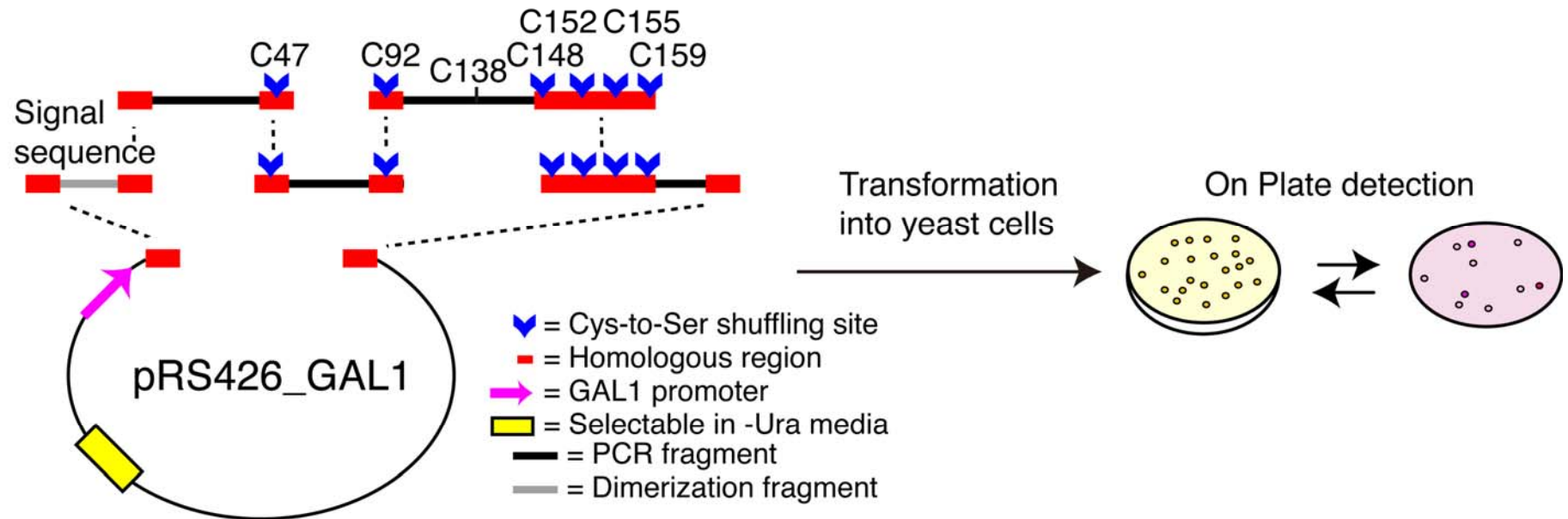
Fig. 1. Cysteine residues of MCL and the CS shuffling mutation strategy. (a) The primary structure of MCL (BAH84844). Cysteine residues are shown in white over a black background. (b) Illustration of the CS shuffling mutation strategy.

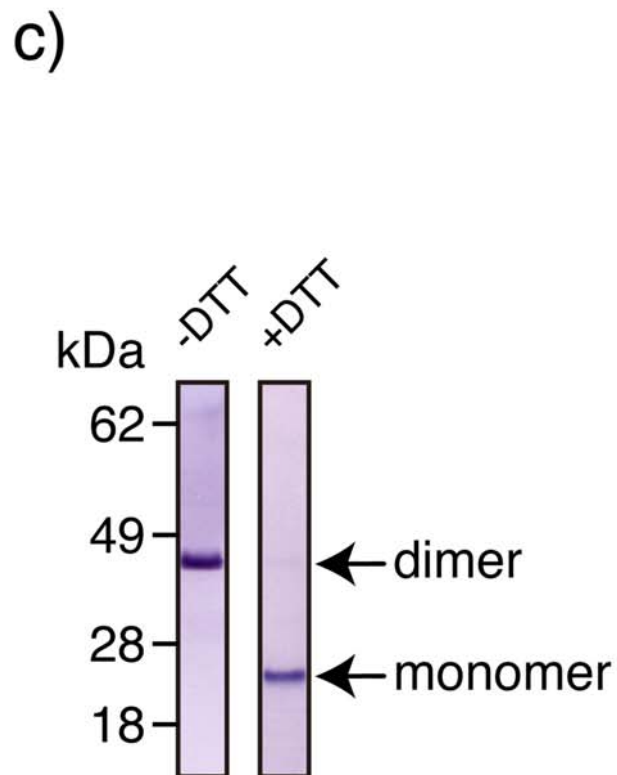
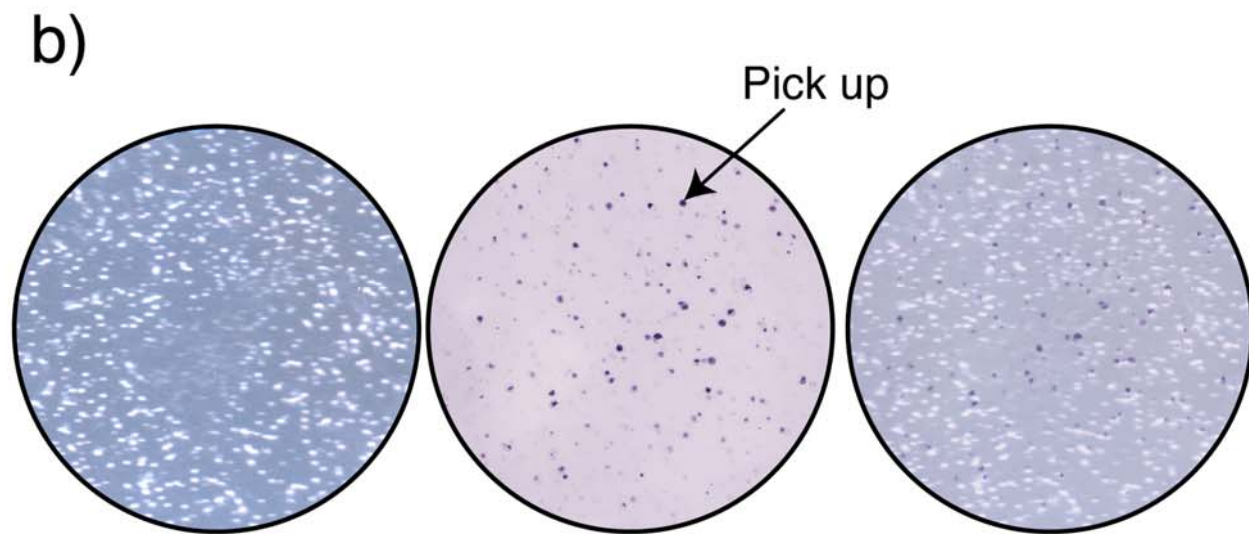
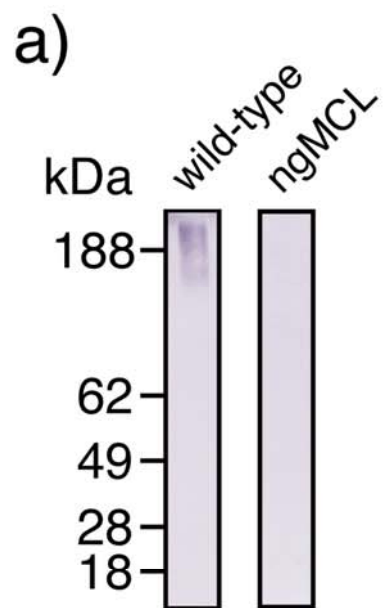
Fig. 2. Application of the CS shuffling strategy for the secretion screening of ngMCL. (a) Secretion of wild-type and ngMCL using a yeast expression system. (b) On-plate immunodetection of CS-shuffled ngMCL mutants. Colonies of transformants were analyzed by on-plate immunodetection, and merged images are shown. (c) Liquid expression of the ngMCL mutant. The culture supernatant was analyzed by immunoblotting in the absence or presence of dithiothreitol (DTT). A representative result is shown. (d) Sequence analysis of ngMCL secretable mutants. The plasmids of ten clones detected using on-plate detection were sequenced randomly. Amino acids at the CS shuffling mutation sites are shown.

a)

1-DSAPNPVLDI	DGEKLRTGTN	YYIVPVLRDH	GGGLTVSATT
41-PNGTFV C PPR	VVQTRKEVDH	DRPLAFFPEN	PKEDVVRVST
81-DLNINFSAFM	P C RWTSSTVW	RLDKYDESTG	QYFVTIGGVK
121-GNPGPETISS	WFKIEEF C GS	GFYKLVF C PT	V C GS C KVK C G
161-DVGIIYIDQKG	RRRLALSDKP	FAFEFNKTVY	F

b)





d)

		Amino acid position					
		47	92	148	152	155	159
Clone No.	1	Ser	Cys	Cys	Cys	Cys	Cys
	2	Ser	Ser	Cys	Ser	Ser	Cys
	3	Ser	Ser	Cys	Cys	Cys	Cys
	4	Ser	Cys	Cys	Cys	Cys	Cys
	5	Ser	Ser	Cys	Cys	Cys	Cys
	6	Ser	Cys	Cys	Cys	Cys	Cys
	7	Ser	Ser	Cys	Ser	Ser	Cys
	8	Cys	Ser	Cys	Cys	Cys	Cys
	9	Ser	Ser	Cys	Ser	Ser	Cys
	10	Ser	Cys	Cys	Cys	Ser	Cys