Supplementary Material

Enhancing thermal stability of single-chain Fv fragment by in vivo global fluorination of the proline residues

Selvakumar Edwardraja,[a] Sokalingam Sriram,[a] Raghunathan Govindan,[a] Nediljko Budisa[b],and Sun-Gu Lee*[a]

[a] Biomolecular Engineering Lab
Department of Chemical Engineering
Pusan National University
30, Jangjeon-dong, Geumjeong-gu
Busan – 609 735
Korea (Republic of)

[b] Department of Biocatalysis
Institute of Chemistry
Technical University of Berlin
Strasse des 17. Juni 124
Berlin D-10623 Germany

*corresponding author (sungulee@pusan.ac.kr)

Fig. S1 The relative frequency profile of prolines at proline distributed sites in the immunoglobulin heavy chain and light chain variable domains. The framework and CDR regions are indicated with gray and black bars respectively. The conserved proline sites are indicated with the dotted line bars. The data was obtained from the link http://www.biochemistry.utoronto.ca/steipe/research/canonical.html, where the precompiled frequency profiles of each amino acid residue positions for immunoglobulin domains are available.

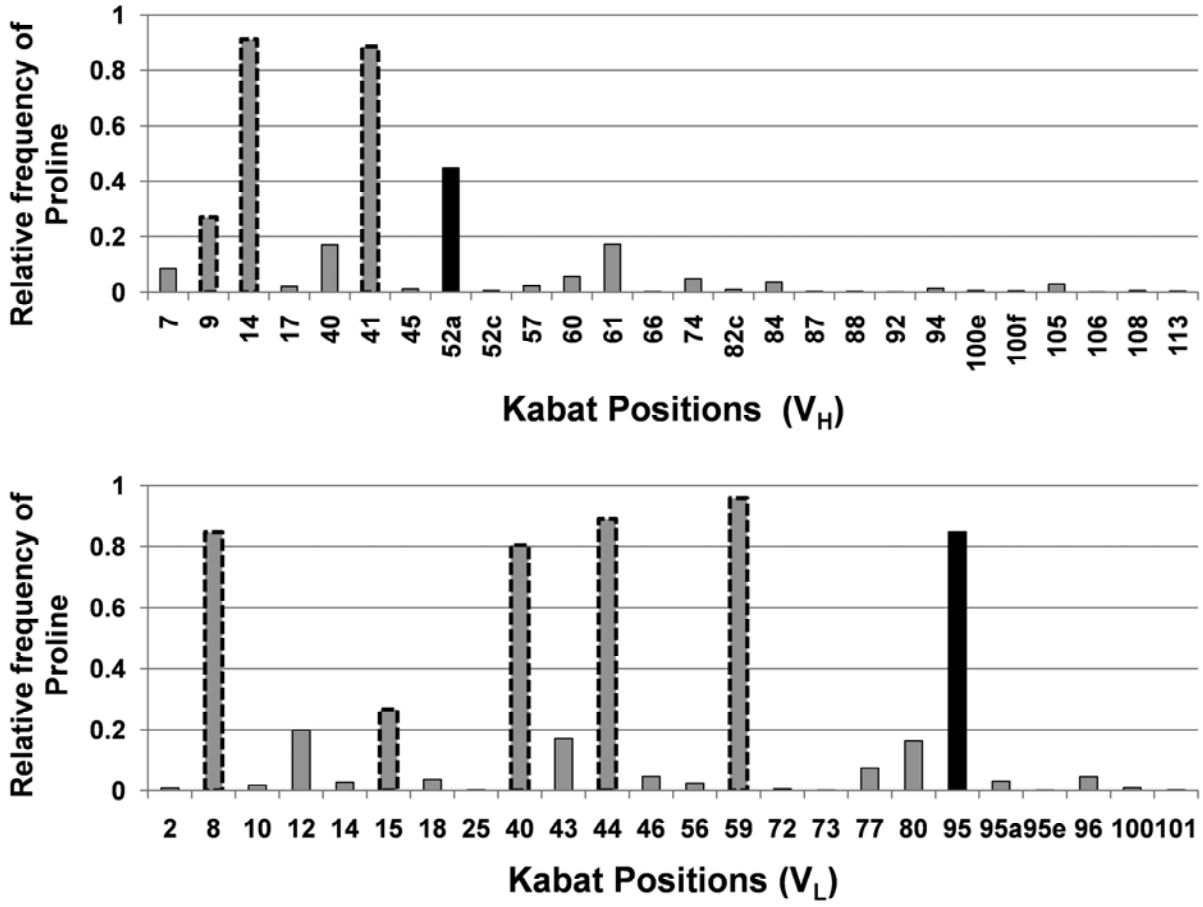

Fig. S2 DNA sequence and protein sequence of hu-MscFv. T7-tag and 6xHis tag are highlighted in the N-terminus and C-terminus of the protein sequence respectively. Proline distributions are highlighted in the protein sequences.

>pQE60/T7-hu-MscFv/834 bases

ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAATTCCAGGAACAGCTG
GTTGAATCTGGTGGCGGTCTGGTTCAGCCAGGCGGCTCCCTCCGTCTCTCTTGCGCTGCG
TCCGGCTTCGACTTTTCTTCTCACTGGATCTACTGGGTTCGTCAGGCGCCAGGTAAAGGT
CTGGAATGGGTTTCCACGATTTACACGGGCAGCGATTCCACCTACTATGCGACGTGGGCG
AAAGGTCGCTTCACGATCAGCAAGGACAACAGCAAGAACACGGTCTACCTCCAGATGAAC
TCTCTGCGTGCGGAAGATACCGCAGTTTACTATTGCGCGCGTGATCTGGGTGGCTCCTCT
TCTACTAGCTACATCTCTGACCTCTGGGGCCAAGGCACTCTGGTTACCGTGTCTTCTGGC
GGTGGTGGCAGCGGTGGCGGCGGTTCCGGCGGTGGTGGTTCCGAACTCGTTCTCACCCAG
TCTCCTGCGACGCTGTCTCTGAGCCCAGGTGAACGTGCCACTCTGTCCTGTACGCTCTCT
TCTGCGCACAAAACCTACTCTATCGCGTGGTATCAGCAGAAACCGGGTCAGGCACCGCGT
TACCTGATCCAGCTGAAATCTGACGGTTCTTACACCAAAGGTACCGGTGTGCCGGCTCGT
TTTTCTGGTAGCAGCTCTGGTGCGGACCGTACCCTCACTATCAGCTCTCTGGAACCGGAA
GACTTCGCGGTATACTACTGCTCTACCGACTACGCGACCGGTTACTACGTTTTCGGTCAG
GGCACGAAGGTTGAGATCAAACGTCACCACCATCATCACCAT***TAATAA***AAGCTT

>pQE60-T7-hu-MscFv (274 residues)

MASMTGGQQMGRGSEFQEQLVESGGGLVQ**P**GGSLRLSCAASGFDFSSHWIYWVRQA**P**GKG
LEWVSTIYTGSDSTYYATWAKGRFTISKDNSKNTVYLQMNSLRAEDTAVYYCARDLGGSS
STSYISDLWGQGTLVTVSSGGGGSGGGGSGGGGSELVLTQS**P**ATLSLS**P**GERATLSCTLS
SAHKTYSIAWYQQK**P**GQA**P**RYLIQLKSDGSYTKGTGV**P**ARFSGSSSGADRTLTISSLE**P**E
DFAVYYCSTDYATGYYVFGQGTKVEIKR**HHHHHH**

Fig. S3 SDS-PAGE analysis of total cell fractions scFvs of the BL21(DE3)pLysS (KC1325) producing hu-MscFv with natural proline or its analogues. The over expressed proteins are indicated by arrow (lane M: protein ladder; lane NP: positive control (hu-MscFv with all 20 natural amino acids); lane RP: hu-MscFv with 19 natural amino acids plus 4R-FPro in substitution of natural prolines; lane SP: hu-MscFv with 19 natural amino acids plus 4S-FPro in substitution of natural prolines.

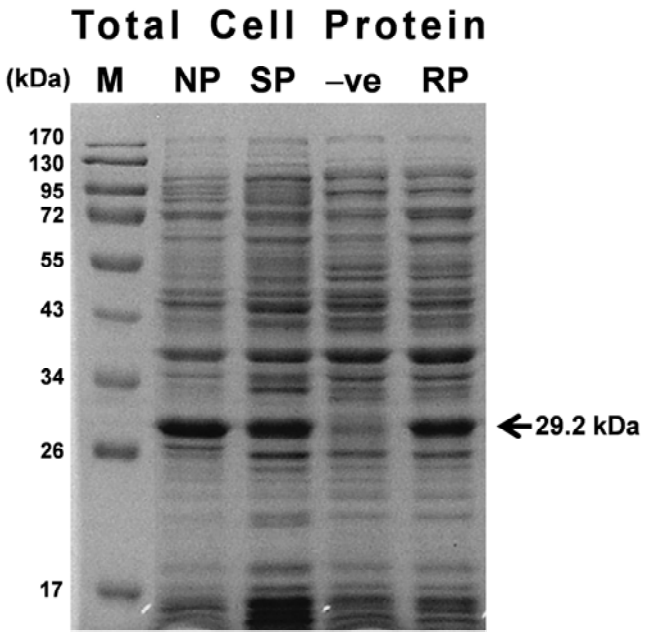

Fig. S4 SDS-PAGE analysis of purified scFvs of the BL21(DE3)pLysS (KC1325) producing hu-MscFv with proline or 4*R*-FPro. The purified proteins are indicated by arrow (lane M: protein ladder; lane NP: hu-MscFv with all 20 natural amino acids; lane RP: hu-MscFv with 19 natural amino acids plus 4*R*-FPro in substitution of natural prolines.

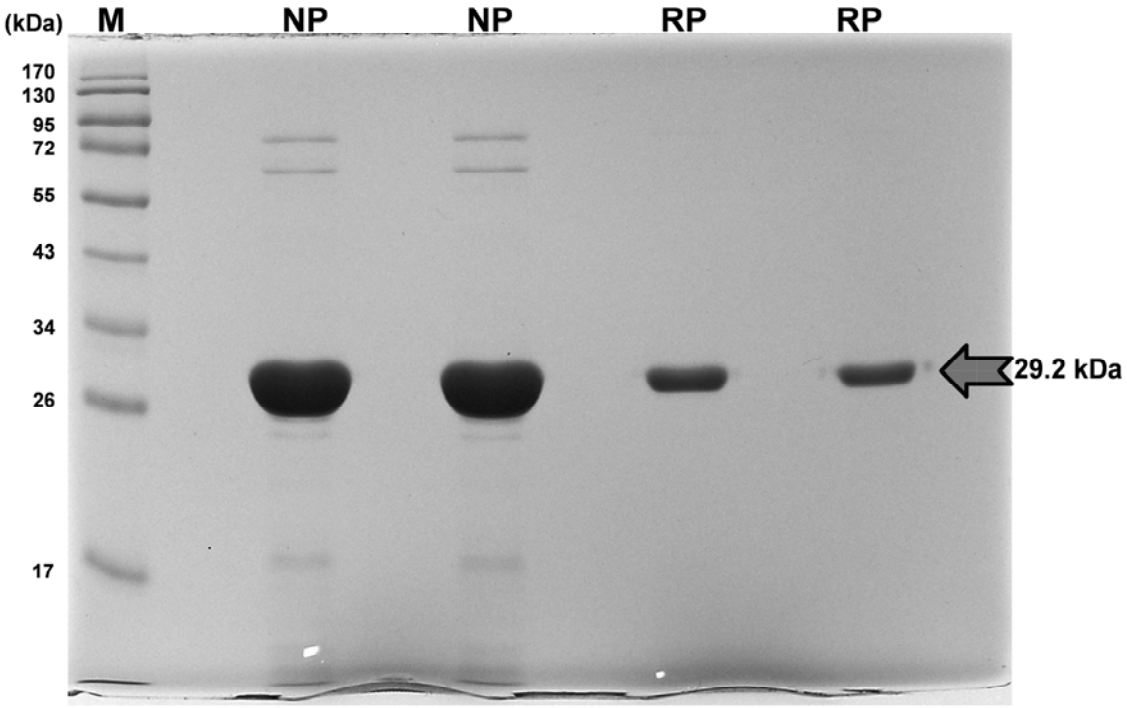

Fig. S5 LC-ESI-MS/MS analysis of hu-MscFv after trypsin digestion. (**a**) LC-ESI-MS/MS profile of the tryptic fragments of hu-MscFv which includes the ion peaks of the peptide, L.VLTQSPATL.S, containing proline, with observed mass of 929.5 Da. (**b**) LC-ESI-MS/MS profile of the tryptic fragments of hu-MscFv which include the ion peaks of L.VLTQSP@ATL.S containing 4*R*-FPro (P@) instead of proline, with observed mass of 947.5 Da.

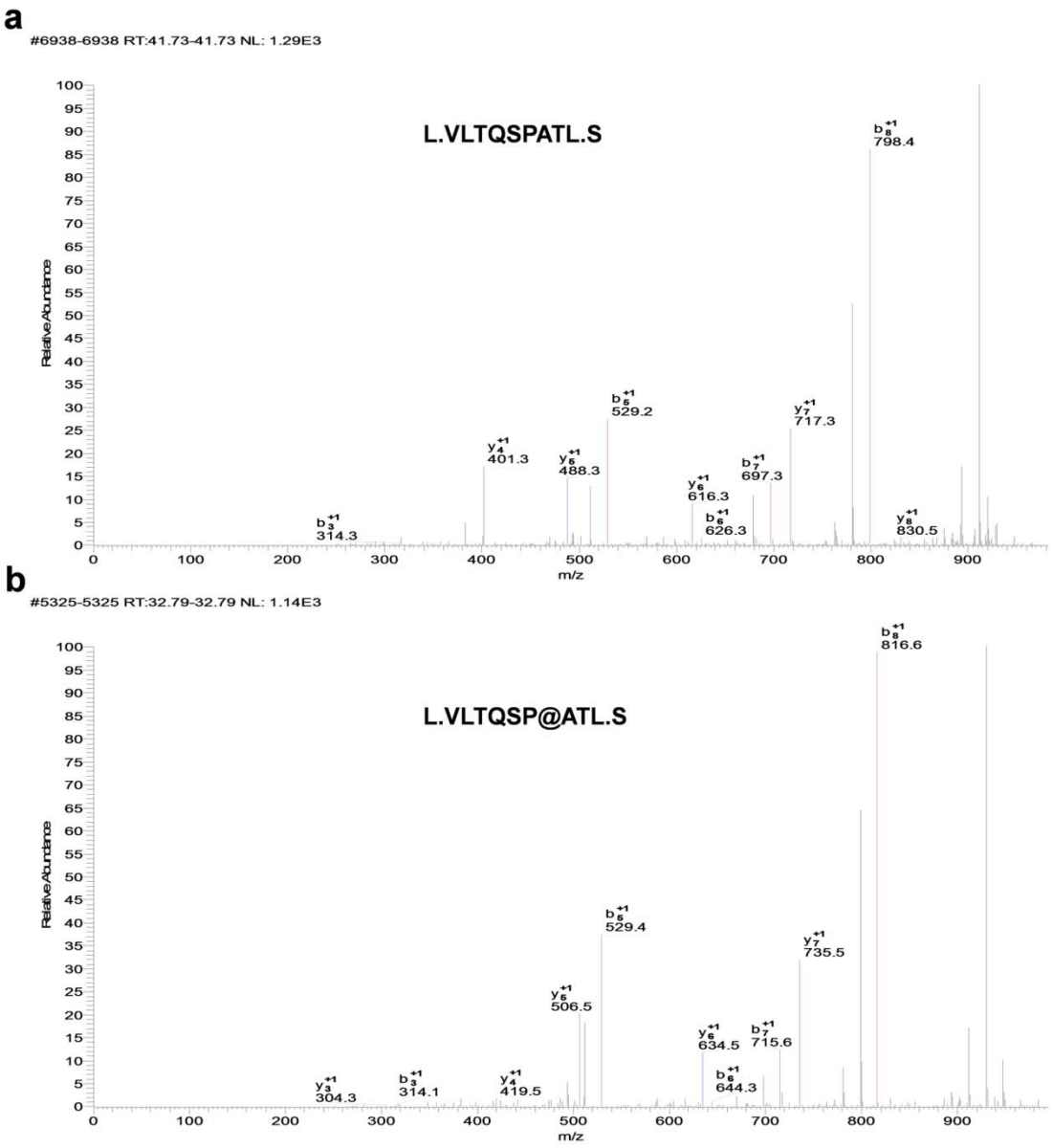

Fig. S6 LC-ESI-MS/MS analysis of hu-MscFv after trypsin digestion. (**a**) LC-ESI-MS/MS profile of the tryptic fragments of hu-MscFv which includes the ion peaks of the peptide, W.YQQKPGQAPRYL.I, containing proline, with observed mass of 1448.7 Da. (**b**) LC-ESI-MS/MS profile of the tryptic fragments of hu-MscFv which include the ion peaks of W.YQQKP@GQAP@RYL.I containing 4R-FPro (P@) instead of proline, with observed mass of 1484.7 Da.

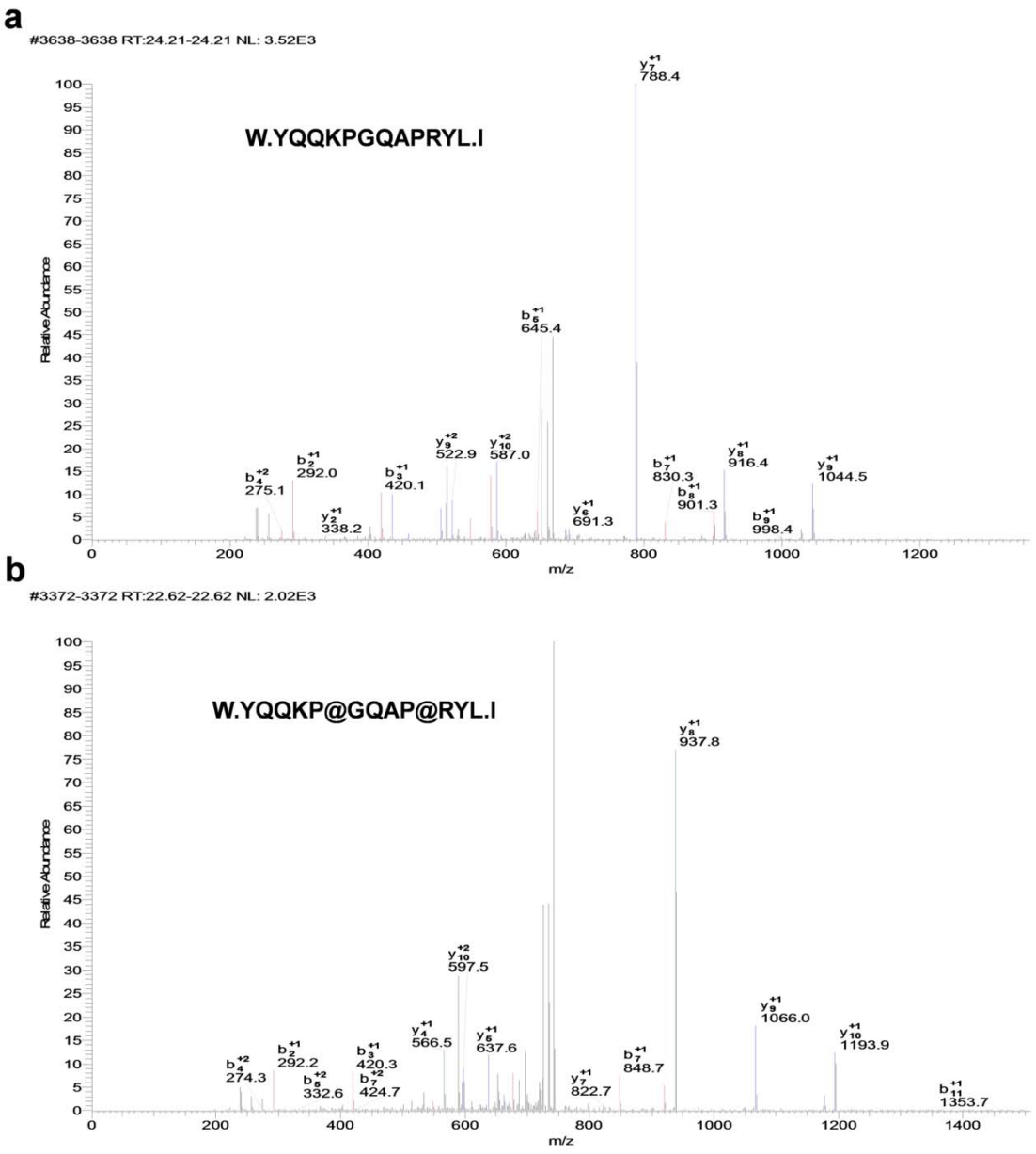

Fig. S7 LC-ESI-MS/MS analysis of hu-MscFv after trypsin digestion. (**a**) LC-ESI-MS/MS profile of the tryptic fragments of hu-MscFv which includes the ion peaks of the peptide, Y.TKGTGVPARF.S, containing proline, with observed mass of 1033.5 Da. (**b**) LC-ESI-MS/MS profile of the tryptic fragments of hu-MscFv which include the ion peaks of Y.TKGTGVP@ARF.S containing 4R-FPro (P@) instead of proline, with observed mass of 1051.5 Da.

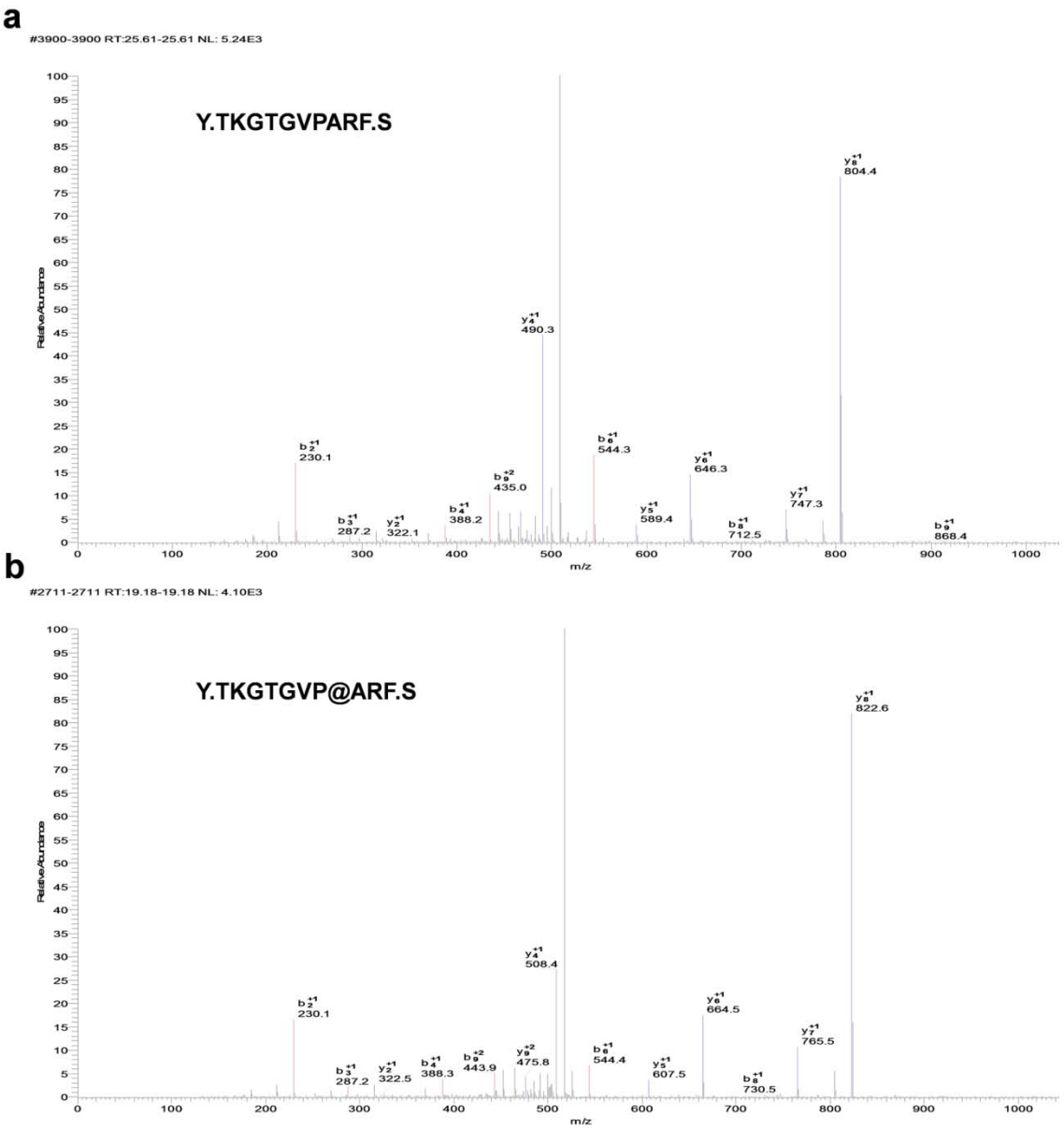

Fig. S8 Stereo image of the modelled structure of 4R-FP-hu-MscFv. All proline residues were replaced with 4R-FPro. Fluoroprolines (H14, H41, L8, L15, L40, L44, L59, L80) are indicated with labels (Kabat numbering) and fluorine atoms are highlighted with cyan color. Both V_H and V_L domains are blue to red that indicate the N-terminus to C-terminus respectively. All fluorinated proline residues except VL(4R)FPro8 exhibit trans-peptide conformation.