J Mol Evol (1997) 44:542–551

JOURNAL OF MOLECULAR EVOLUTION

Molecular Phylogenetic Characterization of *Streptomyces* Protease Inhibitor Family

Seiichi Taguchi,[1] Shuichi Kojima,[2] Mahito Terabe,[2] Yoshinori Kumazawa,[3] Hiroshi Kohriyama,[2] Masayuki Suzuki,[1] Kin-ichiro Miura,[2] Haruo Momose[1]

[1] Department of Biological Science and Technology, Science University of Tokyo, 2641 Yamazaki, Noda-shi, Chiba 278, Japan
[2] Institute for Biomolecular Science, Gakushuin University, Mejiro, Tokyo 171, Japan
[3] Department of Earth and Planetary Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-01, Japan

Received: 23 August 1996 / Accepted: 20 November 1996

Abstract. We previously found that proteinaceous protease inhibitors homologous to *Streptomyces* subtilisin inhibitor (SSI) are widely produced by various *Streptomyces* species, and we designated them ''SSI-like proteins'' (Taguchi S, Kikuchi H, Suzuki M, Kojima S, Terabe M, Miura K, Nakase T, Momose H [1993] *Appl Environ Microbiol* 59:4338–4341). In this study, SSI-like proteins from five strains of the genus *Streptoverticillium* were purified and sequenced, and molecular phylogenetic trees were constructed on the basis of the determined amino acid sequences together with those determined previously for *Streptomyces* species. The phylogenetic trees showed that SSI-like proteins from *Streptoverticillium* species are phylogenetically included in *Streptomyces* SSI-like proteins but form a monophyletic group as a distinct lineage within the *Streptomyces* proteins. This provides an alternative phylogenetic framework to the previous one based on partial small ribosomal RNA sequences, and it may indicate that the phylogenetic affiliation of the genus *Streptoverticillium* should be revised. The phylogenetic trees also suggested that SSI-like proteins possessing arginine or methionine at the P1 site, the major reactive center site toward target proteases, arose multiple times on independent lineages from ancestral proteins possessing lysine at the P1 site. Most of the codon changes at the P1 site inferred to have occurred during the evolution of SSI-like proteins are consistent with those inferred from the extremely high G + C content of *Streptomyces* genomes. The inferred minimum number of amino acid replacements at the P1 site was nearly equal to the average number for all the variable sites. It thus appears that positive Darwinian selection, which has been postulated to account for accelerated rates of amino acid replacement at the major reaction center site of mammalian protease inhibitors, may not have dictated the evolution of the bacterial SSI-like proteins.

Key words: Protease inhibitor — Amino acid sequence alignment — *Streptomyces* — *Streptoverticillium* — Molecular phylogeny — Reactive center site — Amino acid replacement — Codon change — G + C content — Positive Darwinian selection

Introduction

Streptomyces subtilisin inhibitor (SSI) is a proteinaceous protease inhibitor (Murao and Sato 1972) whose structure–function relationships have been extensively studied in combination with a variety of physicochemical methods and the generation of site-specific mutants (Hiromi et al. 1985; Kojima et al. 1990, 1991, 1993, 1994a,b). Finding that homologous proteins of SSI are widely distributed in streptomycetes, particularly in two

Abbreviations: MP, maximum parsimony; NJ, neighbor-joining; SIL, *Streptomyces* subtilisin inhibitor-like; SSI, *Streptomyces* subtilisin inhibitor

Correspondence to: S. Taguchi; e-mail: staguchi@rs.noda.sut.ac.jp

Table 1. Bacterial strains and SSI-like proteins[a]

Strain	SSI-like protein	Reference
Streptomyces	SIL-protein	
S. albogriseolus S-3253 ISP5003	SSI	Murao and Sato 1972
S. antifibrinolyticus JCM7351	PSN	Sugino et al. 1978
S. griseoincarnatus ISP5274	API-2c′	Suzuki et al. 1981
S. lividans 66 ISP5434	SLPI (STI1)	Ueda et al. 1992
S. longisporus ISP5166	STI2	Strickler et al. 1992
S. cacaoi ISP5057	SIL1	Kojima et al. 1994b
S. parvulus ISP5048	SIL2	Taguchi et al. 1994
S. coelicolor Müller ISP5233	SIL3	Taguchi et al. 1994
S. lavendulae KCC-S985	SIL4	Taguchi et al. 1994
S. virginiae ISP5094	SIL8	Terabe et al. 1994b
S. thermotolerans ISP5227	SIL10	Terabe et al. 1994a
S. hygroscopicus ISP5578	SIL12	Terabe et al. 1996
S. galbus ISP5089	SIL13	Terabe et al. 1994a
S. azureus ISP5106	SIL14	Terabe et al. 1994a
S. bikiniensis JCM4011	SIL15	Terabe et al. 1995
Streptoverticillium	SIL-V protein	
Sv. flavopersicus ISP5093	SIL-V1	This study
Sv. orinoci ISP5571	SIL-V2	This study
Sv. eurocidicus ISP5604	SIL-V3	This study
Sv. netropsis ISP5259	SIL-V4	This study
Sv. luteoverticillatus ISP5038	SIL-V5	This study
Sv. cinnamoneum subsp. *cinnamoneum*	SAC-I	Tanabe et al. 1994

[a] ISP: International *Streptomyces* Project; JCM: Japan Collection of Microorganisms, Riken; KCC: KCC Culture Collection of Actinomycetes, Kaken Pharmaceutical Co. Ltd

genera, *Streptomyces* and *Streptoverticillium* (traditional classification), we have recently designated them ''SSI-like proteins'' (Taguchi et al. 1992, 1993a,b), here abbreviated to ''SIL'' proteins for those in *Streptomyces* and ''SIL-V'' proteins for those in *Streptoverticillium.* To date, plasminostreptin (PSN) from *Streptomyces antifibrinolyticus* (Sugino et al. 1978), alkaline protease inhibitor (API-2c′) from *Streptomyces griseoincarnatus* (Suzuki et al. 1981), *Streptomyces lividans* protease inhibitor (SLPI) or *Streptomyces* trypsin inhibitor 1 (STI1) from *Streptomyces lividans* (Ueda et al. 1992; Strickler et al. 1992), STI2 from *Streptomyces longisporus* (Strickler et al. 1992), and SAC-I from *Streptoverticillium cinnamoneum* (Tanabe et al. 1994) have been isolated independently by us and other groups as SSI-like proteins, and their structure—function relationships have been well characterized (Kojima et al. 1994b; Taguchi et al. 1994; Terabe et al. 1994a,b, 1995, 1996; Ueda et al. 1992). To our knowledge, no other bacterial protease inhibitor family has been found.

The discovery of these ''natural mutants'' of SSI prompted us to investigate the evolutionary process of inhibitor proteins in order to gain an insight into the physiological significance of protease–protease inhibitor interactions in nature (Taguchi 1995). In the present study, complete amino acid sequences of five SIL-V proteins newly isolated from *Streptoverticillium* strains were determined, and representative SSI-like proteins, including these SIL-V proteins, were phylogenetically characterized. The phylogenetic trees constructed are used to address the following two subjects: (1) the phylogeny of the producer strains on the basis of the molecular tree of the SSI-like proteins, and (2) the tempo and mode of amino acid replacement at the P1 site, the center of the reactive site which primarily determines the inhibition specificity of SSI-like proteins toward proteases (Hiromi et al. 1985).

Materials and Methods

Bacterial Strains and Protease Inhibitor Samples. All the strains and their protease inhibitors investigated in this study are listed in Table 1. The five *Streptoverticillium* strains were provided by Dr. T. Nakase of the Institute of Physical and Chemical Research (RIKEN), Wako-shi, Japan.

Purification of SIL-V Proteins. The five SIL-V proteins were purified to homogeneity from the culture supernatants of the *Streptoverticillium* strains essentially as described previously (Taguchi et al. 1994).

S-Pyridylethylation of Cysteine Residues, Enzymatic and Chemical Digestions, and Sequence Analysis. Cystine residues of purified SIL-V proteins were reduced with dithiothreitol and *S*-pyridylethylated with 4-vinylpyridine as described previously (Taguchi et al. 1994). Digestion of modified SIL-V proteins and their peptides by proteases or BrCN was carried out as described previously (Taguchi et al. 1994; Ueda et al. 1992). Amino acid sequences were determined using an Applied Biosystems model 476A protein sequencer.

Complete Sequence Determination of SSI-like Proteins. (1) SIL-V1: Digestion of *S*-pyridylethylated SIL-V1 by arginyl endopeptidase pro-

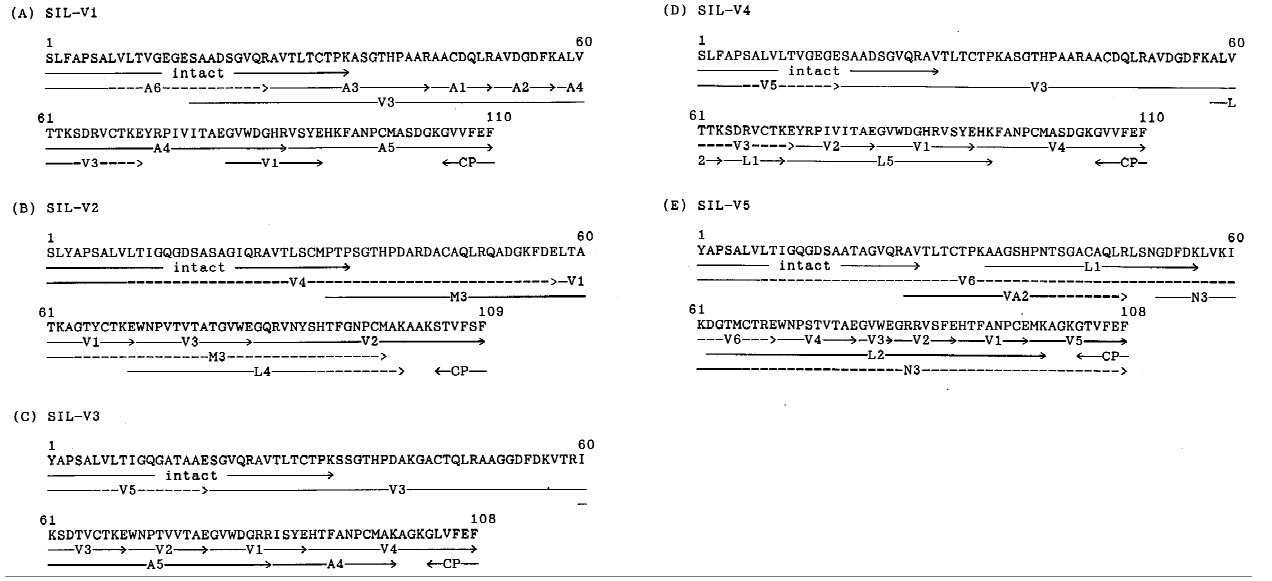

Fig. 1. Complete amino acid sequences and sequencing strategy for SIL-V1 **(A),** SIL-V2 **(B),** SIL-V3 **(C),** SIL-V4 **(D),** and SIL-V5 **(E).** *Long arrows* delineate amino acid sequences identified by the sequencer and *dashed lines* indicate the remaining regions. Amino acids released by carboxypeptidase A are indicated by *CP*.

duced six peptides, whose sequences were determined. The partial sequence of intact SIL-V1 indicated that peptide A6 was an amino-terminal fragment and that peptide A3 followed it. Peptide A5, possessing no Arg residue, was a carboxy-terminal fragment. The connectivities of A3-A1-A2-A4 and A4-A5 were confirmed by sequence analysis of the *S. aureus* V8 protease-digested peptides, V3 and V1, respectively, as shown in Fig. 1A.

(2) SIL-V2: Four peptides isolated from the *S. aureus* V8 protease digests of *S*-pyridylethylated SIL-V2 were subjected to sequence analysis. By comparison with the partial sequence of intact SIL-V2, peptide V4 was found to be an amino-terminal fragment. Peptide V2, with no Glu residue, was positioned at a carboxy terminus. The sequence of peptide M3 obtained by BrCN digestion and that of peptide L4 obtained from the lysylendopeptidase digests of *S*-pyridylethylated SIL-V2 confirmed the connectivities of V4-V1 and V1-V3-V2, respectively. The remaining sequence of the carboxy-terminal region in peptide V4 was clarified by the analysis of peptide M3. The total sequence of SIL-V2 is shown in Fig. 1B.

(3) SIL-V3: The amino acid sequences of five peptides obtained by *S. aureus* V8 protease digestion of *S*-pyridylethylated SIL-V3 were determined. The partial sequence of intact SIL-V3 confirmed the order V5–V3 in the amino-terminal region. Peptide V4, having Phe as a carboxy terminus, was a carboxy-terminal fragment. The sequences of the peptides obtained by arginyl endopeptidase digestion, A5 and A4, confirmed the connectivities of V3-V2-V1 and V1-V4, respectively, as shown in Fig. 1C.

(4) SIL-V4: The sequences of five peptides isolated from the *S. aureus* V8 protease digests of *S*-pyridylethylated SIL-V4 were analyzed. The order, V5–V3, in the amino-terminal region was confirmed from the partial sequence of intact SIL-V4. Peptide V4 was positioned at a carboxy terminus because it had Phe as a carboxy-terminal residue. The amino acid sequences of peptides L2 and L1 isolated from the lysylendopeptidase digests clarified the remaining sequence of the carboxy-terminal region in peptide V3, and the sequence of peptide L5 confirmed the connectivities of V3-V2-V1-V4, as shown in Fig. 1D.

(5)SIL-V5: Six peptides were isolated from the *S. aureus* V8 protease digests of *S*-pyridylethylated SIL-V5 and their sequences were analyzed. Peptide V6, whose sequence was identical to that of intact SIL-V5, was positioned at an amino terminus. Peptide V5, possessing Phe as a carboxy-terminal residue, was a carboxy-terminal fragment. Sequence analysis of peptide L1 and L2 isolated from the lysylendopeptidase digests, peptide N3 obtained by asparginylendopeptidase digestion, and peptide VA2 isolated from the arginylendopeptidase digests of peptide V6 clarified the remaining sequence of the carboxy-terminal region in peptide V6, as shown in Fig. 1E. Peptide L2 also confirmed the connectivities of V6-V4-V3-V2-V1-V5.

The complete amino acid sequences of the five SIL-V proteins were thus determined. The revealed carboxy-terminal sequences were consistent with the results of carboxypeptidase A treatment of intact SIL-V proteins, as shown in Fig. 1. All the amino acid sequences determined in the present study will appear in the EMBL amino acid sequence data base with accession numbers P80596–P80600.

Data Analysis. SSI-like protein sequences were aligned with Clustal W (Gibson et al. 1994). All phylogenetic analyses were done with PHYLIP version 3.5 (Felsenstein 1993). Distance trees were constructed by the neighbor-joining (NJ) method (Saitou and Nei 1987) with NEIGHBOR of the PHYLIP package after pairwise sequence differences (Table 2) were calculated with PROTDIST using the Dayhoff PAM001 matrix. Confidence values for internal lineages were assessed from 300 bootstrap replications of the original sequence data. Maximum parsimony (MP) analyses were done with PROTPARS of the package and the majority rule consensus tree was obtained with CONSENSE.

Results

Determination of Complete Amino Acid Sequences of SIL-V Proteins

Figure 1 shows the complete primary structures of the five SIL-V proteins determined by the sequencing strategy described in Materials and Methods. Like other SIL proteins as well as SSI, each SIL-V protein is composed of approximately 100 amino acids. The amino acid sequence of SIL-V1 turned out to be identical to that of SIL-V4.

Table 2. Pairwise sequence differences among SSI-like proteins[a]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. SSI																				
2. API-2c′	0.10																			
3. SIL14	0.31	0.32																		
4. SIL10	0.38	0.39	0.32																	
5. STI2	0.44	0.42	0.28	0.25																
6. SIL4	0.46	0.47	0.38	0.33	0.30															
7. PSN	0.46	0.46	0.33	0.17	0.28	0.38														
8. SIL13	0.43	0.43	0.33	0.34	0.28	0.48	0.25													
9. SLPI	0.44	0.49	0.38	0.34	0.36	0.43	0.32	0.29												
10. SIL3	0.44	0.47	0.32	0.35	0.38	0.44	0.32	0.31	0.01											
11. SIL2	0.50	0.46	0.31	0.29	0.37	0.46	0.31	0.30	0.15	0.14										
12. SAC-I	0.81	0.85	0.67	0.65	0.80	0.66	0.63	0.65	0.62	0.63	0.53									
13. SIL-V5	0.75	0.80	0.70	0.65	0.83	0.66	0.67	0.67	0.58	0.58	0.54	0.16								
14. SIL-V3	0.81	0.82	0.67	0.64	0.81	0.69	0.68	0.66	0.62	0.64	0.55	0.29	0.32							
15. SIL-V2	0.88	0.86	0.67	0.68	0.78	0.65	0.64	0.70	0.68	0.67	0.65	0.41	0.45	0.42						
16. SIL-V1(V4)	0.82	0.83	0.63	0.62	0.81	0.72	0.74	0.65	0.56	0.59	0.59	0.53	0.50	0.34	0.56					
17. SIL12	0.75	0.82	0.68	0.64	0.67	0.71	0.68	0.62	0.59	0.60	0.65	0.72	0.64	0.70	0.67	0.65				
18. SIL8	1.04	1.02	0.97	0.93	0.98	0.95	0.97	0.86	0.95	0.97	0.88	0.99	0.92	0.94	0.97	1.00	0.90			
19. SIL15	0.88	0.81	0.82	0.85	0.85	0.84	0.83	0.81	0.89	0.88	0.84	1.10	1.01	1.08	0.94	1.07	0.97	0.58		
20. SIL1	0.99	1.01	0.96	1.03	1.01	0.92	0.96	0.86	0.81	0.81	0.83	0.93	0.95	0.94	0.96	1.00	0.90	1.10	1.06	

[a] Pairwise sequence differences among aligned SSI-like proteins (see Fig. 2) were calculated with PROTDIST of the PHYLIP package using the Dayhoff PAM matrix, an empirical model that scales probabilities of change from one amino acid to another. The values shown are scaled in units representing the expected fraction of amino acids that have changed

Sequence Comparison of SSI-Like Proteins

The sequences of the five SIL-V proteins from *Streptoverticillium*—i.e., SIL-V1(V4), V2, V3, and V5—together with that for SAC-I, were aligned with the 15 sequences of SIL proteins from *Streptomyces* already determined using Clustal W (Fig. 2). The alignment varied in a few local regions when alignment parameters (e.g., gap penalties) and matrices were changed. However, these small variations in alignment proved to have little influence on the phylogenetic framework constructed from the aligned sequence data. In Fig. 2, 27 amino acid residues out of the total of 112 are seen to be completely conserved among the 20 SSI-like proteins. Another eight residues are semiconserved within categories of chemically similar amino acids. The existence of these conservation sites and the scarcity of gap sites support the legitimacy of this alignment, which is also consistent with the tertiary structure of SSI as deduced from crystallographic studies (Mitsui et al. 1979a,b). In fact, the 27 conserved residues, marked with asterisks in Fig. 2, correspond well with those that may be required for structure maintenance and the inhibitory action of the SSI and SIL proteins (Kojima et al. 1993, 1994a; Taguchi et al. 1994). For example, residues Val13, Arg29, Trp86, Cys71, Cys101, Asn99, and Phe113 have been shown or are considered to play roles in folding the proper structural coordination of residues around the reaction center or residues for dimer formation, and substitution at these residues had a drastic result—the conversion of the inhibitor molecule to a substrate of subtilisin BPN′ (Hiromi et al. 1985; Kojima et al. 1993, 1994a; Tamura et al. 1991).

Molecular Phylogeny of SSI-Like Proteins

Table 2 provides the pairwise differences among SSI-like proteins calculated with PROTDIST using the Dayhoff PAM001 matrix. The pairwise differences range from 0.01 to 1.10 with an average of 0.64, and there is a clear tendency for pairwise differences involving one of three SIL proteins (SIL1, SIL8, and SIL15) as a partner to be larger than those for pairs within the other 17 SSI-like proteins. This suggests that these three SIL proteins are phylogenetically distant from the other 17 SSI-like proteins and that the phylogenetic root of all the SSI-like proteins is likely to lie on a lineage leading to one or another of these three SIL proteins.

Figure 3 shows an NJ tree constructed from the pairwise difference data in Table 2. No other SSI-like protein sequences exist that clearly serve as an outgroup to these proteins. A phylogenetic tree constructed by the UPGMA (Unweighed Pair-Group Method with Arithmetric Mean) method, which assumes a constancy of evolutionary rates on each lineage, placed the root on the terminal lineage leading to SIL-1 (data not shown). The root of the NJ tree is thus tentatively placed on the terminal lineage leading to SIL1. Phylogenetic relationships supported by more than 95% bootstrap probability values are usually considered to be statistically significant. By this criterion, sister-group relationships for the following SSI-like pro-

```
                7   11          21          31          41          51          61
SSI             YAPS ALVLTVGKGV SATTAAPERA VTLTCAPGPS GTHPAAGSAC ADLAAV--GGDL NALT
API-2c'         YAPS ALVLTVGKGV SAATVTPERA VTLTCAPGPS GTHPAADSAC ADLAAV--GGDL DALT
SIL14           YAPS ALVLTVGEGE SAAAATPERA VTLTCAPRPS GTHPVAGSAC AELRGV--GGDV HALT
SIL10           YAPS ALVLTVGHGE SAIAATPERA VTLTCAPKAA GTHPAAGAAC AELRGV--GGDF DALT
STI2            YAPS ALVLTVGHGT SAAAATPLRA VTLNCAPTAS GTHPAPALAC ADLRGV--GGDI DALK
SIL4            YAPS ALVLTIGHGG AAATATPERA VTLTCAPTSS GTHPAASAAC AELRGV--GGDF AALK
PSN             YAPS ALVLTMGHGN SAATVNPERA VTLNCAPTAS GTHPAALQAC AELRGA--GGDF DALT
SIL13           YAPS ALVLTMGHGE SAAAVSPARA VTLNCAPSAS GTHPAPALAC AELRAA--GGDL DALA
SLPI            YAPS ALVLTVGHGE SAATAAPLRA VTLTCAPTAS GTHPAAAAAC AELRAA--HGDP SALA
SIL3            YAPS ALVLTVGHGE SAATAAPLRA VTLTCAPTAS GTHPAADAAC AELRAA--HGDP SALA
SIL2            YAPS ALVLTIGQGE SAAATSPLRA VTLTCAPKAT GTHPAADAAC AELRRA--GGDF DALS
SAC-I           YAPS ALVLTIGQGD SAAAAGIQRA VTLTCMPKAD GTHPNTRGAC AQLRLA--GGDF EKVT
SIL-V5          YAPS ALVLTIGQGD SAATAGVQRA VTLTCTPKAA GSHPNTSGAC AQLRLS--NGDF DKLV
SIL-V3          YAPS ALVLTIGQGA TAAESGVQRA VTLTCTPKSS GTHPDAKGAC TQLRAA--GGDF DKVT
SIL-V2          YAPS ALVLTIGQGD -SASAGIQRA VTLSCMPTPS GTHPDARDAC AQLRQA--DGKF DELT
SIL-V1(V4)      FAPS ALVLTVGEGE SAADSGVQRA VTLTCTPKAS GTHPAARAAC DQLRAV--DGDF KALV
SIL12           YPAS ALVLTVGHGA DAATAEVQRA VTLSCRPTPT GTHPAPAQAC AELHSV--GGAL GLLR
SIL8            YAPS AMVFSVAQGD DVAAPTVVRA TTVSCAPGAR GTHPDPKAAC AALKS--TGGAF DRLL
SIL15           YAPS AVVLSIGKGD ASGPVTVLRA TTLSCAPVPG GTHPAPEAAC AELKAGFAGGGF GGLL
SIL1            YAPS AVVISKTQGA -SADAPAQRA VTLRCLPVG- GDHPAPEKAC AALREA--GGDP AALP
                .  * *.*.  *        ** *. * *   * **   **  *    *    .

                    71P1       81          91          101         111
SSI             RGED---VM CPMVYDPVLL TVDGVWQGKR VSYERVFSNE CEMNAHGSSV FAF
API-2c'         RSEG---VM CPMIYDPVLL TVDGVWQGER VSYERVFSNE CEMNAHGSSV LAF
SIL14           ATDG---VM CTKQYDPVVV TVDGVWQGRR VSYERTFSNE CVKNAYGSGV FAF
SIL10           ARDG---VM CTKQYDPVVV TVEGVWQGKR VSYERTFSND CMKNAYGTGV FSF
STI2            ARDG---VI CNKLYDPVVV TVDGVWQGKR VSYERTFGNE CVKNSYGTSL FAF
SIL4            ARDD---VW CNKLYDPVVV TAQGVWQGQR VSYERTFGNS CERDAVGGSL FAF
PSN             VRGD---VA CTKQFDPVVV TVDGVWQGKR VSYERTFANE CVKNSYGMTV FTF
SIL13           GPAD---TV CTKQYAPVVI TVDGVWQGKR VSYERTFANE CVKNASGSSV FAF
SLPI            AEDS---VM CTREYAPVVV TVDGVWQGRR LSYERTFANE CVKNAGSASV FTF
SIL3            ADDA---VM CTREYAPVVV TVDGVWQGRR LSYERTFANE CVKNAGSASV FTF
SIL2            AADG---VM CTREYAPVVV TVDGVWQGRR LSYERTFANE CVKNAGSASV FTF
SAC-I           KIKE-G-TA CTREWNPSVV TAEGVWEGRR VSFERTFANP CELKAGKGTV FEF
SIL-V5          KIKD-G-TM CTREWNPSTV TAEGVWEGRR VSFEHTFANP CEMKAGKGTV FEF
SIL-V3          RIKS-D-TV CTKEWNPTVV TAEGVWDGRR ISYEHTFANP CMAKAGKGLV FEF
SIL-V2          ATKA-G-TY CTKEWNPVTV TATGVWEGQR VNYSHTFGNP CMAKAAKSTV FSF
SIL-V1(V4)      TTKS-D-RV CTKEYRPIVI TAEGVWDGHR VSYEHKFANP CMASDGKGVV FEF
SIL12           TGAEPG-RM CTKEWRPITV TAEGVWDGRR VSYEHTFANN CFKNAAPTTV FEF
SIL8            SEPNPD-RA CPMHYAPVTV SAVGVWEGRR VAWDHTFANS CTMAATLDGN AVF
SIL15           ASPDPD-RA CPQHFDPVTV TLDGVWEGAR TSWQHTFSNA CVMGTTLDGG EAF
SIL1            RYVEDTGRV CTREYRPVTV SVQGVWDGRR IDHAQTFSNS CELEKQTASV YAF
                        *    * . .  *** * *    . * *  *           *
```

Fig. 2. Alignment of SSI-like protein sequences. The alignment was done with the Clustal W alignment software. Sequence sources of the SSI-like proteins used in this study are presented in Table 1. Highly variable residues with extensive length variations located at the N-terminal are excluded from the alignment; the alignment thus begins from the seventh residue of the SSI protein. *Bars* represent gaps. An *asterisk* denotes the complete conservation of an amino acid at the corresponding residue, while a *small dot* means that the corresponding residue is semiconserved within a category of chemically similar amino acids.

teins are statistically supported: (1) SSI and API-2c′; (2) SLPI and SIL3; (3) SLPI, SIL3, and SIL2; (4) SAC-I and SIL-V5; and (5) SIL8 and SIL15. Two other clusterings are also supported by strong bootstrap values: (6) SSI, API-2c′, SIL14, SIL10, STI2, SIL4, PSN, SIL13, SLPI, SIL3, and SIL2; and (7) SAC-I, SIL-V5, SIL-V3, SIL-V2, and SIL-V1(V4). However, since the bootstrap values for the other branches are far less than the significant level, the phylogenetic relationships corresponding to these branches are considered not to be reliably resolved in this study.

We also constructed an MP tree using the same dataset (Fig. 4). The MP tree obtained is essentially consistent with the NJ tree with respect to branching patterns and the confidence levels of individual branches. Minor discrepancies between the two trees exist in the placement of (1) SIL12, (2) SIL-V2 and SIL-V3, and (3) SIL14 and SIL10, but none of these relationships is statistically significant in either tree. It is therefore concluded that the MP and NJ trees consistently establish the phylogenetic relationships of (1)–(7) described above.

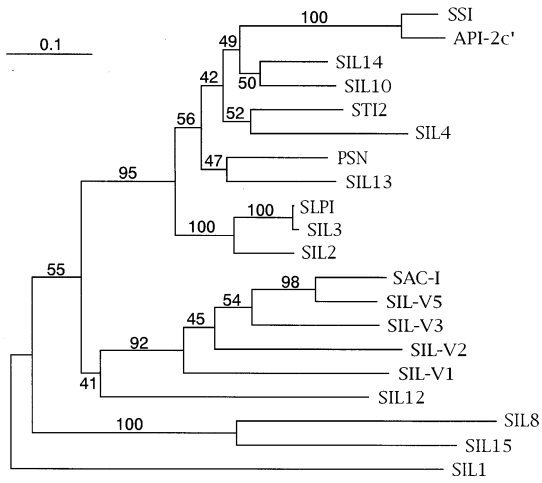

Fig. 3. Neighbor-joining tree for SSI-like proteins. The tree was constructed from the data presented in Table 2 as described in Materials and Methods. Bootstrap probability values from 300 replications are shown at the corresponding branches. The *scale bar* represents a difference corresponding to 0.1 (10%). The root is tentatively placed at an arbitrary position on the terminal lineage leading to SIL1 (see text for the reasoning).

Discussion

Phylogenetic Utility of SSI-Like Proteins

Molecular phylogenetic studies of bacteria have been based mostly on small rRNA sequences (Olsen et al. 1994; Woese 1987). Ribosomal RNAs are ubiquitous to all organisms, and the fact that the evolutionary rate of rRNA sequences varies greatly within the molecule makes rRNA sequences applicable to the systematic study of both distantly and closely related species. Notwithstanding the advantages of small rRNA sequences as a useful molecular phylogenetic tool, however, the necessity of adopting other molecules in addition to rRNA sequences has been pointed out (see, e.g., Eisen 1995; Hasegawa and Hashimoto 1993, for reasons and discussion). Trees based on rRNA sequences may be more directly affected by an unequal base frequency of the genome than those based on protein sequences. In this context, the employment of protein sequences as a phylogenetic tool may be especially important for analyzing microorganisms with extremely high G + C contents, as is the case with streptomycetes (Wright and Bibb 1992). In our case, the amino acid frequencies of SSI-like proteins were free from the bias of streptomycetes genome

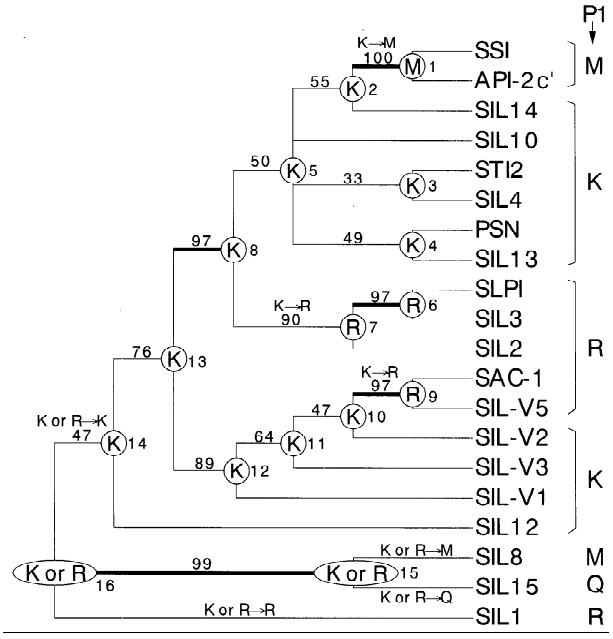

Fig. 4. Maximum parsimony tree for SSI-like proteins. The tree was constructed as described in Materials and Methods. Four equally parsimonious trees were obtained and the 50% majority rule consensus tree among them is shown. Since the root is likely to lie on a lineage leading to SIL1 or to SIL8–SIL15 (see text), the root is tentatively placed at node 16. *Numbers* are given to individual nodes. *Numbers on each branch* represent the corresponding bootstrap probability values from 300 replications, and branches for which the corresponding relationships are supported by >95% bootstrap probability are a *thick line.* The amino acid residue at the P1 site is shown for each SSI-like protein. Amino acid residues unambiguously inferred as an ancestral state by MP analysis are shown at each node. Phylogenetic positions for the occurrence of plausible amino acid replacements at the P1 site are shown on internal lineages. *''K or R''* at nodes 15 and 16 means that there are two equally parsimonious possibilities for the assignment of the character state at these nodes. If lysine was present at nodes 15 and 16, changes for K → R, K → M, and K → Q are inferred to have occurred for node 16 → SIL1, node 15 → SIL8, and node 15 → SIL15, respectively. If arginine was present at these nodes, changes for R → K, R → M, and R → Q are expected for node 16 → node 14, node 15 → SIL8, and node 15 → SIL15, respectively.

G + C content (data not shown), as in the case described previously (Hashimoto et al. 1994).

The SSI-like proteins used in this study are widely distributed in the genera *Streptomyces* and *Streptoverticillium.* Genes for three SIL proteins—SSI, STI1, and STI2 (Obata et al. 1989; Strickler et al. 1992)—were found to be encoded as a single copy in the genome, which discounts the possibility that a paralogous comparison is being made among duplicated genes. Although the SSI-like proteins produce a limited amount of sequence information (approximately 110 amino acids), they can be obtained in large amounts by cultivation, and it is relatively easy to determine their primary structures by automated Edman degradation of peptides generated by enzymatic and chemical digestions. The extremely high G + C content makes it difficult to sequence genes of *Streptomyces,* but determination of amino acid sequences is free from this constraint. For these various reasons. SSI-like proteins were expected to provide an alternative phylogenetic framework for *Streptomyces* to the previous one based on partial small rRNA sequences (Wellington et al. 1992).

Phylogenetic Affiliation of Streptoverticillium

Members of the genus *Streptoverticillium* have traditionally been classified separately from those of the genus *Streptomyces,* mostly due to the characteristic morphological feature of the former of forming spore whorls (Locci and Schofield 1989). However, recent work based on partial 16S rRNA sequences (about 500 nucleotides) (Wellington et al. 1992; Williams et al. 1983; Witt and Stackebrandt 1990) and partial amino acid sequences corresponding to the amino-terminal regions of *Escherichia coli* L30 protein homologs (about 20 amino acids) (Ochi and Hiranuma 1994) from several species of *Streptoverticillium* suggests that members of the genus *Streptoverticillium* are phylogenetically included in those of the genus *Streptomyces* in such a way that members of the two genera are not separated but mixed in the manner of a mosaic. Our trees based on SSI-like proteins (Figs. 3 and 4) are consistent with these findings in that *Streptoverticillium* species are phylogenetically included in *Streptomyces,* but differ in that *Streptoverticillium* forms a monophyletic group as a clade within *Streptomyces.* This relationship is supported by high bootstrap values—i.e., 92% for the NJ tree (Fig. 3) and 89% for the MP tree (Fig. 4). We thus consider that a monophyletic relationship among the *Streptoverticillium* species is very likely, which points to the importance of the formation of spore whorls as a shared derived character in members of this genus.

The closer relationship among the *Streptoverticillium* species was also supported by the direct sequence comparison among the SIL-V proteins. As shown in Fig. 2, SIL-V proteins, including also SAC-I, exhibit several sequence features characteristic of these proteins. For instance, a one-residue insertion in the flexible loop region of SSI (Thr64 to Met70 of SSI) was observed only for SIL-V proteins, whereas a two-residue insertion was observed for SIL8, -12, and -15, and a three-residue insertion was seen for SIL1. There are four residues (Gln52, Lys67, Pro100, and Lys107) which are completely conserved among SIL-V proteins but are never observed in SSI and other SIL proteins, although their contribution to the structure and function of inhibitor proteins remains uncertain. Taken together, these results support the monophyly of the species which produce SIL-V proteins.

The complete identity of primary structure between SIL-V1 and SIL-V4 accords with the classification of their producer strains (*Sv. flavopersicus* and *Sv. netropsis*) in the same category (I.-1.) by Goodfellow (1989).

The close phylogenetic relationship among SLPI, SIL3 (see Figs. 3 and 4), and SILA3 (Taguchi et al. 1996) points to a close relationship among their producer strains, *S. lividans* 66, *S. coelicolor* Müller, and *S. coelicolor* A3(2), and is in agreement with previous studies on genomic mapping, DNA-DNA hybridization, and chemical taxonomy (Hatano et al. 1994). Thus, the molecular phylogeny based on SSI-like proteins is mostly consistent with previous knowledge concerning the phylogeny of streptomycetes.

Molecular Evolution at the P1 Reactive Site

Although the structure–function relationships between SSI and proteases have been studied extensively, the physiological roles of protease–inhibitor interactions in the life cycle of *Streptomyces* have not been well elucidated. On the basis of an analysis of mutant strains devoid of SSI or API-2c′, it has been suggested that SSI-like proteins may play an important role(s) in physiological and/or morphological regulation in *Streptomyces* (Taguchi et al. 1995b,c). With the aim of gaining further insight into the physiological significance of SSI-like proteins, we examined the tempo and mode of evolution at the P1 site. The center of the reactive site, P1, of serine protease inhibitors generally corresponds to the major specificity determinant toward proteases (Laskowski and Kato 1980).

As shown in Fig. 2, four amino acids (Lys, Arg, Met, and Gln) are observed as P1 site residues of the SSI-like proteins. By MP analyses of the SSI-like protein sequences, the occurrence of the different P1 site residues could be assigned to specific lineages on the phylogenetic tree (Fig. 4). For instance, SSI and API-2c′ possessing Met at the P1 site are likely to have arisen from SSI-like proteins possessing Lys at the P1 site in the evolutionary process of node 2 → node 1. SIL8, which also possesses Met at the P1 site, is phylogenetically distinct from SSI or API-2c′, and may thus have arisen from an SSI-like protein possessing Lys or Arg at the P1 site in the evolutionary process of node 15 → SIL8. Among SSI-like proteins possessing Arg at the P1 site, SLPI, SIL3, and SIL2 are clearly distinct from SAC-I and SIL-V5, suggesting that these two subgroups independently evolved from SSI-like proteins possessing Lys at the P1 site (in the evolutionary processes of node 8 → node 7 and node 10 → node 9, respectively). It is also obvious that the Arg residue at the P1 site of SIL1 has an independent origin. Hence, these proteins are likely to have acquired the same amino acid residue at the P1 site by parallel amino acid substitutions. SIL15, possessing Gln at the P1 site, may have evolved from an ancestral protein possessing Lys or Arg at the P1 site in the evolutionary process of node 15 → SIL15.

We tried to infer the mode of codon changes corresponding to these amino acid replacements at the P1 site.

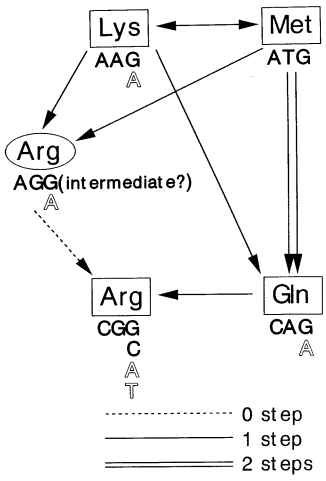

Fig. 5. Possible codon change pathways at the P1 site of SSI-like proteins. Four amino acids—lysine, arginine, methionine, and glutamine—occur at the P1 site of the SSI-like proteins (see Figs. 2 and 4). The most parsimonious codon change pathway(s) is sought between the most likely codons corresponding to these amino acids in the light of the high G + C content of genome; the number of steps needed for each pathway is indicated by the *different kinds of lines.* The most plausible direction of the codon change pathway(s) in light of the high G + C content is indicated by an *arrow.* Since synonymous codon changes occur much more rapidly than nonsynonymous ones, no step is assigned to such synonymous pathways as AGG → CGG in PROTPARS of the PHYLIP package.

A notable feature of *Streptomyces* is a high G + C composition (>70 mol%) in its genomic DNA (Wright and Bibb 1992), presumably driven by ''GC pressure'' causing a mutation from A:T pairs to G:C pairs at high frequency (Sueoka 1962; Muto and Osawa 1987). Since the G + C content at the third position of the codons reaches approximately 93%, codon usage is also extremely biased (Wright and Bibb 1992). The codon usage compiled from 64 *Streptomyces* genes (Wright and Bibb 1992) is: Lys (AAA: 4.9% and *AAG:* 95.1%), Arg (CGT: 6.3%, *CGC:* 46.2%, CGA: 3.9%, *CGG:* 36.5%, AGA: 1.3%, and AGG: 5.7%), Met (*ATG:* 100%), and Gln (CAA: 6.7% and *CAG:* 93.3%) (major codons are italicized). The codon sequences for the P1 site residues of SSI, SLPI (STI1), and STI2 were identified as ATG (Met), CGG (Arg), and AAG (Lys), respectively (Ikenaka et al. 1974; Obata et al. 1989; Strickler et al. 1992), which is consistent with biased codon usage.

Figure 5 shows plausible codon change pathways among Lys, Arg, Met, and Gln and the minimum number of base substitution steps needed for the individual pathways. All changes from Lys to Met in node 2 → node 1, and from Lys to Arg in node 8 → node 7 and in node 10 → node 9 require a single step in a direction that is consistent with the GC pressure (Fig. 5). The P1 site residue of an ancestral SSI-like protein at node 16 (Fig. 4) was inferred to be either Lys or Arg by MP analysis. If Lys were present at the P1 site of node 16, all of the changes for Lys → Arg in node 16 → SIL1, Lys → Met in node 15 → SIL8, and Lys → Gln in node 15 → SIL15

Table 3. Inferred minimum number of amino acid replacements at each residue of SSI-like proteins[a]

	0[b]	1	2	3	4	5	6	7	8	9
0[c]		1	1	1	0	0	2	0	2	1
10	7.3	3	7.5	0	10	11	4	4	11	7.3
20	9	5	7	0	0	2	0	1	5.3	0
30	7	0	10.3	7.3	10	0	3	0	0	4
40	4	13	12.3	0	0	2	5	0	4	10.5
50	8.5	3	4	6	0	5	9	11	7	2
60	10.8	11	15.3	8	13	5	7	3	5	12
70	0	3	6	6.8	6	8.8	0	5	7	3
80	2	5	9	0	0	0	4	0	8	0
90	5	5	4	5	3	3	0	5.5	0	9
100	0	10	9	7	6	11	6	10	9	4
110	5	8	0							

[a] The minimum numbers of amino acid replacements inferred to have occurred during the evolution of SSI-like proteins were obtained with PROTPARS of the PHYLIP package. The values shown are the averages of those obtained from four equally parsimonious trees. The numbering of residues in this table, which refer to the conventional numbering system for SSI (see Fig. 2), starts with the conventional residue 7 and counts gap sites. Thus, the P1 site (Met73 for SSI) corresponds to residue No. 72 in this table

[b] Columns: digits of corresponding residue numbers

[c] Rows: tens or hundreds of corresponding residue numbers

could be achieved by only one base substitution step that is consistent with GC pressure. On the other hand, if Arg were present at the P1 site of node 16, all of the changes for Arg → Lys in node 16 → node 14, Arg → Met in node 15 → SIL8, and Arg → Gln in node 15 → SIL15 would be against the GC pressure. It is therefore most likely that an ancestral SSI-like protein at node 16 possessed Lys at the P1 site, under the assumption that the root of all the investigated SSI-like proteins comes to any point on the lineages leading to either SIL1 or SIL8–SIL15, as described above. This proposition, however, needs to be evaluated in more detail by collecting additional SSI-like protein sequences.

Table 3 shows the inferred minimum number of amino acid replacements at each residue of the listed SSI-like proteins during the course of their evolution. As pointed out earlier, no amino acid replacement appears to have occurred at 27 of the 112 sites, but replacements may have occurred at the remaining 85 sites at different frequencies (ranging from 1 to 15.3, with an average of 6.3). At the P1 site (residue 72 corresponding to Met73 of SSI), a minimum of six substitutions are assigned. This almost equals the average number of all the variable sites, suggesting that the rate of amino acid replacement at the P1 site is neither accelerated nor decelerated as compared to those of the other variable sites. The primary contact region of SSI with subtilisin BPN′ was assigned in the region Gly66 to Tyr75 (Mitsui et al. 1979a) and a minimum substitution of amino acid at its region averaged 5.7. This suggests that not quite high-speed amino acid replacement has occurred at the primary contact region including the P1 reactive site. None of other highly substituted sites have so far been reported to play crucial role in inhibitory action of SSI (Taguchi 1995).

Evolutionary Implications

What do these findings imply with respect to the evolutionary constraints operating on SSI-like proteins and with respect to the physiological role of these proteins? As explained above, the structure–function relationships of naturally occurring SSI-like proteins, as well as of mutagenized SSIs, have been well characterized and point to the P1 residue playing a predominant role in determining interactions with target proteases (Taguchi et al. 1995b,c). SIL proteins possessing such basic amino acids as Arg or Lys at the P1 site have been found to inhibit trypsin in vitro (Kojima et al. 1994b; Taguchi et al. 1994; Terabe et al. 1995, 1996), while mutagenesis studies of SSI have shown that replacement of the P1 site residue (Met73) by either Lys or Tyr results in the acquisition of inhibitory activity toward trypsin or chymotrypsin, respectively, without substantial loss of inhibitory activity toward the extrinsic protease, subtilisin (Kojima et al. 1990, 1991). All five SIL-V proteins purified in this study exhibited effective inhibition toward trypsin (Taguchi S, unpublished data), which is consistent with a Lys or Arg residue occupying their P1 site.

It is widely known that the major reaction center site of mammalian protease inhibitors has undergone amino acid substitution at considerably higher rates than the other sites of the inhibitors in some rodent lineages on which genes for the inhibitors are multiplied by gene duplication (Hill and Hastie 1987; Suzuki et al. 1991; Rheaume et al. 1994; Goodwin et al. 1996; references therein). It has been hypothesized that the accelerated substitution rates at the reaction center site may be due to some sort of diversifying or positive Darwinian selection that drives high rates of amino acid divergence, and that protease-secreting parasites may be the selective agent responsible for it (Hill and Hastie 1987). Interestingly,

accelerated substitution rates at the reaction center site can be seen in between species comparisons, as well as among paralogous inhibitors within a species (Creighton and Darby 1990).

The present study shows that amino acid replacements at the P1 site of SSI-like proteins have been neither accelerated nor decelerated as compared to the other variable sites (Table 3), and that all unambiguously inferrable pathways of amino acid substitutions (i.e., Lys → Met in node 2 → node 1, Lys → Arg in node 8 → node 7 and in node 10 → node 9) are not contrary to the reasonable pathways inferred in light of the extremely high G + C content. It thus appears that strong positive Darwinian selection at the P1 site postulated to have operated on mammalian protease inhibitors may not have dictated the evolution of the bacterial SSI-like proteins.

We have recently isolated three proteases from an SSI-deficient *S. albogriseolus* mutant strain as candidates for intrinsic target proteases of SSI (Taguchi et al. 1995b,c). One of the target proteases, termed SAM-P20, was found to be strongly inhibited by SSI both in vivo and in vitro (Taguchi et al. 1995b,c), and homologs of the SAM-P20 gene were inferred to be distributed widely in *Streptomyces* (Taguchi et al. 1995a). On the basis of these data, we hypothesize that the major physiological role of the SSI-like inhibitor proteins is the activity regulation of intrinsic protease(s) rather than the defense against extrinsic protease(s). In this sense, it is possible that broad inhibitory specificities of SSI toward target proteases can exert advantage in interacting as one molecule with multiple target proteases with various substrate specificities. Examination of the evolutionary mode at the active sites of the intrinsic proteases from the streptomycetes would provide a key to understanding the coevolutionary scheme between the inhibitors and cognate protease(s).

Acknowledgments. This work was supported by grants-in-aid (No. 07760099 to S.T. and No. 07780518 to S.K.) from the Ministry of Education, Science, Sports and Culture of Japan and by grants from the Nissan Science Foundation and Takano Life Science Research Foundation (to S.T.) and from the Asahi Glass Foundation and Hamaguchi Foundation (to S.K.).

References

Creighton TE, Darby NJ (1990) Functional evolutionary divergence of proteolytic enzymes and their inhibitors. Trends Biochem Sci 14: 319–324

Eisen JA (1995) The RecA protein as a model molecule for molecular systematic studies of bacteria: comparison of trees of RecAs and 16S rRNAs from the same species. J Mol Evol 41:1105–1123

Felsenstein J (1993) PHYLIP: phylogeny inference package, version 3.5c. University of Washington, Seattle

Gibson T, Higgins D, Thompson J (1994) Clustal W, version 1.4. EMBL, Heidelberg, Germany

Goodfellow M (1989) Supragenetic classification of actinomycetes. In: Williams ST, Sharpe ME, Holt JG (eds) Bergey's manual of systematic bacteriology, vol 4. The Williams & Wilkins, Baltimore, pp 2333–2339

Goodwin RL, Baumann H, Berger FG (1996) Patterns of divergence during evolution of α1-proteinase inhibitors in mammals. Mol Biol Evol 13:346–358

Hasegawa M, Hashimoto T (1993) Ribosomal RNA tree misleading? Nature 361:23

Hashimoto T, Nakamura Y, Nakamura F, Shirakura T, Adachi J, Goto N, Okamoto K, Hasegawa M (1994) Protein phylogeny gives a robust estimation for early divergences of eukaryotes: phylogenetic place of a mitochondria-lacking protozoan. *Giardia lamblia.* Mol Biol Evol 11:65–71

Hatano K, Tamura T, Nishii T (1994) Taxonomic status of *Streptomyces coelicolor* A3(2) and *Streptomyces lividans* 66. Actinomycetologica 8:47–50

Hill RE, Hastie ND (1987) Accelerated evolution in the reactive centre regions of serine protease inhibitors. Nature 326:96–99

Hiromi K, Akasaka K, Mitsui Y, Tonomura B, Murao S (1985) Protein protease inhibitor—the case of *Streptomyces* subtilisin inhibitor (SSI). Elsevier/North Holland, Amsterdam

Ikenaka T, Odani S, Sakai M, Nabeshima Y, Sato S, Murao S (1974) Amino acid sequence of an alkaline proteinase inhibitor (*Streptomyces* subtilisin inhibitor) from *Streptomyces* albogriseolus S-3253. J Biochem 76:1191–1209

Kojima S, Obata S, Kumagai I, Miura K (1990) Alteration of the specificity of the *Streptomyces* subtilisin inhibitor by gene engineering. Biotechnology 8:449–452

Kojima S, Nishiyama Y, Kumagai I, Miura K (1991) Inhibition of subtilisin BPN′ by reaction site P1 mutants of *Streptomyces* subtilisin inhibitor. J Biochem 109:377–382

Kojima S, Kumagai I, Miura K (1993) Requirement for a disulfide bridge near the reactive site of protease inhibitor SSI (*Streptomyces* subtilisin inhibitor) for its inhibitory action. J Mol Biol 230:395–399

Kojima S, Fujimura K, Kumagai I, Miura K (1994a) Contribution of salt bridge in the protease inhibitor SSI (*Streptomyces* subtilisin inhibitor) to its inhibitory action. FEBS Lett 337:195–199

Kojima S, Terabe M, Taguchi S, Momose H, Miura K (1994b) Primary structure and inhibitory properties of a proteinase inhibitor produced by *Streptomyces cacaoi.* Biochim Biophys Acta 1207:120–125

Laskowski M Jr, Kato I (1980) Protein inhibitors of proteinases. Annu Rev Biochem 49:593–626

Locci R, Schofield GM (1989) Genus *Streptoverticillium.* In: Williams ST, Sharpe ME, Holt JG (eds) Bergey's manual of systematic bacteriology, vol 4. The Williams & Wilkins, Baltimore, pp 2492–2504

Mitsui Y, Satow Y, Watanabe Y, Hirono S, Iitaka Y (1979a) Crystal structure of *Streptomyces* subtilisin inhibitor and its complex with subtilisin BPN′. Nature 277:447–452

Mitsui Y, Satow Y, Watanabe Y, Iitaka Y (1979b) Crystal structure of a bacterial protein proteinase inhibitor (*Streptomyces* subtilisin inhibitor) at 2.6 Å resolution. J Mol Biol 131:697–724

Murao S, Sato S (1972) S-SI, a new alkaline protease inhibitor from *Streptomyces albogriseolus* S-3253. Agric Biol Chem 36:160–163

Muto A, Osawa S (1987) The guanine and cytosine content of genomic DNA and bacterial evolution. Proc Natl Acad Sci USA 84:166–169

Obata S, Taguchi S, Kumagai I, Miura K (1989) Molecular cloning and nucleotide sequence determination of gene encoding *Streptomyces* subtilisin inhibitor (SSI). J Biochem 105:367–371

Ochi K, Hiranuma H (1994) A taxonomic review of the genera *Kitasatosporia* and *Streptoverticillium* by analysis of ribosomal protein AT-L30. Int J Syst Bacteriol 44:285–292

Olsen GJ, Woese CR, Overbeek R (1994) The winds of (evolutionary) change: breathing new life into microbiology. J Bacteriol 176:1–6

Rheaume C, Goodwin RL, Latimer JJ, Baumann H, Berger FG (1994) Evolution of murine α_1-proteinase inhibitors: gene amplification and reactive center divergence. J Mol Evol 38:121–131

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Strickler JE, Berka TR, Gorniak J, Fornwald J, Keys R, Rowland JJ, Rosenberg M, Taylor DP (1992) Two novel *Streptomyces* protein protease inhibitors. Purification, activity, cloning and expression. J Biol Chem 267:3236–3241

Sueoka N (1962) On the genetic basis of variation and heterogeneity of DNA base composition. Proc Natl Acad Sci USA 48:582–592

Sugino H, Kakinuma A, Iwanaga S (1978) Plasminostreptin, a protein protease inhibitor produced by *Streptomyces antifibrinolyticus.* 3. Elucidation of the primary structure. J Biol Chem 253:1546–1555

Suzuki K, Uyeda M, Shibata M (1981) Partial amino acid sequence of an alkaline protease inhibitor, API-2 (b and c). Agric Biol Chem 45:629–634

Suzuki Y, Yoshida K, Honda E, Sinohara H (1991) Molecular cloning and sequence analysis of cDNA coding for guinea pig α1-anti proteinases S and F and contrapsin. J Biol Chem 266:928–932

Taguchi S (1995) Communication between protease and protease inhibitor in the *Streptomyces* world. Actinomycetologica 9:216–227

Taguchi S, Kojima S, Kumagai I, Ogawara H, Miura K, Momose H (1992) Isolation and partial characterization of SSI-like protease inhibitors from *Streptomyces.* FEMS Microbiol Lett 99:293–297

Taguchi S, Kikuchi H, Kojima S, Kumagai I, Nakase T, Miura K, Momose H (1993a) High frequency of SSI-like protease inhibitors among *Streptomyces.* Biosci Biotech Biochem 57:522–524

Taguchi S, Kikuchi H, Suzuki M, Kojima S, Terabe M, Miura K, Nakase T, Momose H (1993b) *Streptomyces* subtilisin inhibitor-like proteins are distributed widely in streptomycetes. Appl Environ Microbiol 59:4338–4341

Taguchi S, Kojima S, Terabe M, Miura K, Momose H (1994) Comparative studies on the primary structures and inhibitory properties of subtilisin-trypsin inhibitors from *Streptomyces.* Eur J Biochem 220:911–918

Taguchi S, Endo T, Naoi Y, Momose H (1995a) Molecular cloning and sequence analysis of a gene encoding an extracellular serine protease from *Streptomyces lividans* 66. Biosci Biotech Biochem 59: 1386–1388

Taguchi S, Odaka A, Watanabe Y, Momose H (1995b) Molecular characterization of a gene encoding extracellular serine protease isolated from a subtilisin inhibitor-deficient mutant of *Streptomyces albogriseolus* S-3253. Appl Environ Microbiol 61:180–186

Taguchi S, Suzuki M, Kojima S, Miura K, Momose H (1995c) *Streptomyces* serine protease (SAM-P20): recombinant production, characterization, and interaction with endogenous protease inhibitor. J Bacteriol 177:6638–6643

Taguchi S, Kojima S, Miura K, Momose H (1996) Taxonomic characterization of closely related *Streptomyces* spp. based on the amino acid sequence analysis of protease inhibitor proteins. FEMS Microbiol Lett 135:169–173

Tamura A, Kanaori K, Kojima S, Kumagai I, Miura K, Akasaka K (1991) Mechanisms of temporary inhibition in *Streptomyces* subtilisin inhibitor induced by an amino acid substitution, tryptophan 86 replaced by histidine. Biochemistry 30:5275–5286

Tanabe M, Kawahara K, Asano T, Kato K, Kakinuma A (1994) Primary structure and reactive site of *Streptoverticillium* anticoagulant (SAC), a novel protein inhibitor of blood coagulation produced by *Streptoverticillium cinnamoneum* subsp. *cinnamoneum.* J Biochem 115:752–761

Terabe M, Kojima S, Taguchi S, Momose H, Miura K (1994a) Three novel subtilisin-trypsin inhibitors from *Streptomyces:* primary structures and inhibitory properties. J Biochem 116:1156–1163

Terabe M, Kojima S, Taguchi S, Momose H, Miura K (1994b) Primary structure and inhibitory properties of a subtilisin-chymotrypsin inhibitor from *Streptomyces virginiae.* Eur J Biochem 226:627–632

Terabe M, Kojima S, Taguchi S, Momose H, Miura K (1995) A subtilisin inhibitor produced by *Streptomyces bikiniensis* possesses a glutamine residue at reactive site P1. J Biochem 117:609–613

Terabe M, Kojima S, Taguchi S, Momose H, Miura K (1996) New subtilisin-trypsin inhibitors produced by *Streptomyces:* primary structures and their relationship to other proteinase inhibitors from *Streptomyces.* Biochim Biophys Acta 1293:233–240

Ueda Y, Kojima S, Tsumoto K, Takeda S, Miura K, Kumagai I (1992) A protease inhibitor produced by *Streptomyces lividans* 66 exhibits inhibitory activities toward both subtilisin BPN′ and trypsin. J Biochem 112:204–211

Wellington EM, Stackbrandt E, Sanders D, Wolstrup J, Jorgensen NO (1992) Taxonomic status of *Kitasatosporia* and proposed unification with *Streptomyces* on the basis of phenotypic and 16S rRNA analysis and emendation of *Streptomyces* Waksman and Henrici 1943, 339^{AL}. Int J Syst Bacteriol 42:156–160

Williams ST, Goodfellow M, Alderson G, Wellington EMH, Sneath PHA, Sackin MJ (1983) Numerical classification of *Streptomyces* and related genera. J Gen Microbiol 129:1743–1813

Witt D, Stackbrandt E (1990) Unification of the genera *Streptoverticillium* and *Streptomyces*, and amendation of *Streptomyces* Waksman and Henrici 1943, 339^{AL}. Syst Appl Microbiol 13:361–371

Woese CR (1987) Bacterial evolution. Microbiol Rev 51:221–271

Wright F, Bibb MJ (1992) Codon usage in the G + C-rich *Streptomyces* genome. Gene 113:55–65