Intact-cell MALDI-TOF mass spectrometry analysis of peptaibol formation by the genus *Trichoderma/Hypocrea*: can molecular phylogeny of species predict peptaibol structures?

Torsten Neuhoof,¹ Ralf Dieckmann,^{1†} Irina S. Druzhinina,²
Christian P. Kubicek² and Hans von Döhren¹

Correspondence
Hans von Döhren
Doehren@chem.tu-berlin.de

Christian Kubicek
ckubicek@mail.zserv.tuwien.ac.at

¹TU Berlin, Institut für Chemie, FG Biochemie und Molekulare Biologie, Franklinstr. 29, 10587 Berlin, Germany

²Forschungsbereich Gentechnik und Angewandte Biochemie, Institut für Verfahrenstechnik, Umwelttechnik und Technische Biowissenschaften, TU Wien, Getreidemarkt 9-166, 1060 Wien, Austria

Peptaibols are characteristic linear α -aminoisobutyrate-containing peptides produced by certain Ascomycetes, especially of the genus *Hypocrea/Trichoderma* [*Hypocrea* and *Trichoderma* are the names for the teleo- and anamorph forms of the same taxon; where known to occur in nature, the teleomorph is used to name the species. To aid the inexperienced reader, both names (the less well known one in parentheses) are given at the first mention of each species.] Here we have investigated whether phylogenetic relationships within *Trichoderma* permit a prediction of the peptaibol production profiles. To this end, representative strains from a third (28) of the known species of *Trichoderma*, identified by the sequences of diagnostic genes and covering most clades of the established multilocus phylogeny of *Trichoderma/Hypocrea*, were investigated by intact-cell MALDI-TOF mass spectrometry. Peptaibols were detected in all strains, and some strains were found to produce up to five peptide families of different sizes. Comparison of the data with phylogenies derived from rRNA spacer regions (ITS1 and 2) and RNA polymerase subunit B (*rpb2*) gene sequences did not show a strict correlation with the types and sequences of the peptaibols produced, but the production of some groups of peptaibols appears to be found only in some clades or sections of the genus, which could be used for more targeted screening of novel compounds of this type. In an analysis of peptaibol structures, we have defined conserved key positions and have further identified and compared sequences of the corresponding adenylate domains within non-ribosomal peptide synthetases producing trichovirins, paracelsins and atroviridins. These phylogenies are not concordant with those of their producers *Hypocrea virens*, *Hypocrea jecorina* and *Hypocrea atroviridis* as obtained from ITS1 and 2, and *rpb2*, respectively, and therefore hint at a complex history of peptaibol diversity.

Received 5 February 2007
Revised 14 May 2007
Accepted 22 June 2007

INTRODUCTION

The mitosporic fungal genus *Trichoderma/Hypocrea* is well known for the production of peptaibols or peptaibiotics (Szekeres *et al.*, 2005; Whitmore & Wallace, 2004). These linear, non-ribosomally formed peptides of 5–20 residues are characteristically highly abundant in α -aminoisobutyric acid (Aib), and contain an *N*-acyl (usually acetyl) terminus and a C-terminal alcohol, such as phenylalaninol or

leucinol (Benedetti *et al.*, 1982; Brückner & Graf, 1983; Brückner *et al.*, 1984). A database of peptaibol structures (www.cryst.bbk.ac.uk/peptaibol/home.shtml) currently lists more than 300 structures, grouped into nine distinct subfamilies, which can be further differentiated according to their amino acid number and substitution pattern, and terminal processing (Chugh & Wallace, 2001). Peptaibols/peptaibiotics are also formed by other fungi of the Hypocreales, but reports of their production by basidiomycetes seem to have been due to infections of the basidiomycete fruiting bodies by these fungi which are frequently mycoparasites (Degenkolb *et al.*, 2003; Neuhoof *et al.*, 2007b). Within the genus *Trichoderma/Hypocrea*, so far subfamilies 1, 4, 5 and 9 have been described (Szekeres

[†]Present address: Anagnostec GmbH, Im Biotechnologiepark TGZ II, 14943 Luckenwalde, Germany.

Abbreviations: Aib, α -aminoisobutyric acid; ICMS, intact-cell MALDI-TOF mass spectrometry; NRPS, non-ribosomal peptide synthetase; SF_n, subfamily *n*.

et al., 2005). Subfamily 1 (SF1) comprises about half of the known structures and combines peptides ranging from 18 to 20 aa in length. All these peptides have partial sequence identities or similarities. Subfamily 4 (SF4) combines peptides of 11 or 14 aa, also sharing sequence similarities, but have no sequence relationship to SF1. Subfamilies 5 and 9 (SF5 and 9) have only a few members and comprise peptides with 11 or 6 and 7 residues, again with no similarities to the other subfamilies.

This particular biodiversity of peptaibols raises the question whether there is a correlation between peptaibols, which are often named after a putative producer species (harzianins, longibrachins, koningins, etc.), and the taxonomic status of the producer strains. The problem of the taxonomy of *Trichoderma* is difficult, and use of phenotypic species identification alone (as in most of these studies) is highly prone to error (Druzhinina *et al.*, 2006). In addition, some peptaibols had already been described before 1969, a time when *Trichoderma* was believed to consist only of a single species, *Trichoderma viride* (Rifai, 1969), and many more up to 1991, when *Trichoderma* was believed to consist only of nine species aggregates (Bissett, 1984, 1991a–c, 1992). Consequently, the species identity of most of the strains producing known peptaibols up to that time is highly uncertain. Only recently have strains with defined species identity been investigated for the peptaibols they produce, which has led to the identification of several new compounds (Wiest *et al.*, 2002; Degenkolb *et al.*, 2006a, b; Vizcaino *et al.*, 2006). In the latter study, peptaibols produced by five defined species of *Trichoderma* (*T. pubescens*, *T. strigosum*, *T. spirale*, *T. erinaceum* and *T. stromaticum*) were compared, and the authors concluded that the type of peptaibols produced correlates only poorly with the taxonomy of the species. This claim must be accepted with caution, however, as Degenkolb *et al.* (2006b) only investigated five out of more than 100 accepted species (cf. www.isth.info), and moreover included three species (*T. pubescens*, *T. erinaceum* and *T. strigosum*) which are phylogenetically closely related.

We have investigated peptaibol production by intact-cell MALDI-TOF mass spectrometry (ICMS) in 28 *Trichoderma/Hypocrea* species that represent a wide range of the diversity within this genus. ICMS has the advantages of being very rapid, using small samples (subcolony amounts) and requiring only minimal sample preparation. ICMS spectra in the range below 2 kDa identify characteristic metabolites, e.g. of bacilli (Williams *et al.*, 2002; Pabel *et al.*, 2003) and cyanobacteria (Erhard *et al.*, 1997; Welker *et al.*, 2002, 2004a, b), and we show here that highly characteristic peptaibol spectra can readily be obtained from microgram amounts of mycelia, thus providing characteristic metabolite fingerprints. A comparison of these fingerprints with molecular phylogenies of the species documents some correlation between types of peptaibol formed and phylogeny.

We have further classified the structures of 18- to 20-residue peptaibols and defined conserved residues corresponding to

conserved domains of the respective non-ribosomal peptide synthetases (NRPSs). Sequences of conserved adenylate domains of synthetases forming trichovirins, paracelsins and atroviridins have been identified from strains representative of different sections of the genus. Substrate binding site analysis permits the prediction of substrate profiles and domain position. Analysis of the nucleotide sequences of the Aib-, Pro- and Gln-activating domains displays similar Kimura distances of the NRPS genes and housekeeping genes.

METHODS

Reagents. 2,5-Dihydroxybenzoic acid (DHB) from Anagnostec was used as matrix for MALDI-TOF experiments. Trifluoroacetic acid (TFA), ethanol, acetonitrile and methanol from Merck were used as solvents.

Microbial strains and cultivation. The *Trichoderma/Hypocrea* strains used in this study are listed in Table 1. All of them have been identified previously at the species level, and the corresponding references are given in Table 1. All strains were cultivated on malt extract agar (3%) at 26 °C.

Extraction and preparation of mycelia for MALDI-TOF analysis.

A few micrograms of fungal mycelia was suspended in acetonitrile/methanol/water (1:1:1, by vol.) and 1 µl of the suspension was directly spotted onto target wells of a 100-position sample plate and immediately mixed with 1 µl matrix solution [10 mg DHB ml⁻¹ in acetonitrile/methanol/water (1:1:1, by vol.) and 0.3% trifluoroacetic acid]. The sample matrix mixture was allowed to air dry prior to analysis. Alternatively, freeze-dried mycelium obtained from shaken cultures or fungi grown on plates was homogenized in 60% ethanol and centrifuged. One microlitre of the protein solution was spotted onto a MALDI target plate and mixed with matrix.

Microbial characterization by MALDI-TOF MS analysis.

MS measurements were performed on a VOYAGER DE-PRO-TOF mass spectrometer from Applied Biosystems. Mass spectra were acquired in linear delayed extraction mode using an acceleration voltage of 20 kV and a low mass gate of 1500 Da. For desorption of the components, a nitrogen laser beam ($\lambda=337$ nm) was focused on the template. The laser power was set to just above the threshold of ionization. Spectra for individual specimens were compiled, averaging results from at least 100 shots taken across the width of the specimen for *m/z* values of 2000–20 000.

Measurements were performed in delayed extraction mode, allowing the determination of monoisotopic mass values. A low mass gate of 800 Da improved the measurement by filtering out the most intensive matrix ions. The mass spectrometer was used in positive ion detection and reflector mode. In addition, an identification took place by a comparison between calculated and detected mass data.

In silico retrieval and identification of conserved adenylate domains.

A full-length copy of *T. atroviride* *tex1* was identified in its genome sequence by a TBLASTN search with the *T. virens* *tex1* gene as probe. The amino acids activated by individual A (amino-acid-activating)-domains were identified by a BLAST search with the corresponding domains of *H. virens* and *H. jecorina* (*T. reesei*; H. von Döhren & C. P. Kubicek, unpublished). The paracelsin synthetase gene has been identified in the *H. jecorina* genome sequence (*T. reesei* v2.0) and is partially annotated at <http://genome.jgi-psf.org/cgi-bin/dispGeneModel?db=Trire2&id=123610>; A-domains have been identified by correlation with known paracelsin structures (H. von

Table 1. List of *Trichoderma/Hypocrea* strains used throughout this work

No.	Name	Anamorph or previous name	Strain no.*	Identity according to:
1	<i>H. atroviridis</i>		ATCC 74058 (P1)	Kullnig <i>et al.</i> (2001)
2	<i>H. atroviridis</i>		ATCC 36042,	Kuhls <i>et al.</i> (1996)
3	<i>H. aureoviridis</i>		CBS 254.63	Kullnig-Gradinger <i>et al.</i> (2002)
4	<i>H. citrina</i>		CBS 977.69	Kullnig-Gradinger <i>et al.</i> (2002)
5	<i>H. gelatinosa</i>		ATCC 7476	Druzhinina <i>et al.</i> (2005)
6	<i>H. hunua</i>		CBS 238.63	Kullnig-Gradinger <i>et al.</i> (2002)
7	<i>H. jecorina</i>	<i>T. reesei</i>	QM 6a	Kuhls <i>et al.</i> (1996)
8	<i>H. lixii</i>	<i>T. harzianum</i>	CBS 226.95	Kullnig-Gradinger <i>et al.</i> (2002)
9	<i>H. lixii</i>	<i>T. inhamatum</i>	CBS 273.78	Druzhinina <i>et al.</i> (2005)
10	<i>H. minutispora</i>		CBS 342.93	Lu <i>et al.</i> (2004)
11	<i>H. pachybasioides</i>	<i>T. croceum</i> (= <i>T. polysporum</i>)	DAOM 167068	Lu <i>et al.</i> (2004)
12	<i>H. semiorbis</i>		GJS 73-48	Kullnig-Gradinger <i>et al.</i> (2002)
13	<i>H. semiorbis</i>		CBS 244.63	Kullnig-Gradinger <i>et al.</i> (2002)
14	<i>H. strictipilosa</i>		CBS 347.93	Druzhinina <i>et al.</i> (2005)
15	<i>H. strictipilosa</i>	<i>T. fasciculatum</i>	CBS 118.72	Kullnig-Gradinger <i>et al.</i> (2002)
16	<i>H. tawa</i>		CBS 246.63	Kullnig-Gradinger <i>et al.</i> (2002)
17	<i>H. virens</i>	<i>T. flavofuscum</i>	CBS 248.59	Druzhinina <i>et al.</i> (2005)
18	<i>T. asperellum</i>		CBS 358.97	Druzhinina <i>et al.</i> (2005)
19	<i>T. brevicompactum</i>		CBS 109720	Kraus <i>et al.</i> (2004)
20	<i>H. schweinitzii</i>	<i>T. citrinoviride</i>	IMI 232088	Kuhls <i>et al.</i> (1996)
21	<i>T. pleuroticola</i>	Previously <i>T. fulvidum</i> , ' <i>T. cf. aureoviride</i> '	DAOM 175924	Druzhinina <i>et al.</i> (2005)
22	<i>T. hamatum</i>		CBS 393.92	Kullnig-Gradinger <i>et al.</i> (2002)
23	<i>T. koningii</i>		CBS 460.96	Kullnig-Gradinger <i>et al.</i> (2002)
24	<i>T. longibrachiatum</i>		CBS 816.68	Kuhls <i>et al.</i> (1996)
25	<i>T. longipile</i>		CBS 340.93	Kullnig-Gradinger <i>et al.</i> (2002)
26	<i>T. oblongisporum</i>		CBS 343.93	Druzhinina <i>et al.</i> (2005)
27	<i>T. pubescens</i>		CBS 119776	Kullnig-Gradinger <i>et al.</i> (2002)
28	<i>T. pubescens</i>		DAOM 162162	Kullnig-Gradinger <i>et al.</i> (2002)
29	<i>T. spirale</i>		CBS 346.93	Kullnig-Gradinger <i>et al.</i> (2002)
30	<i>T. strigosum</i>		CBS 348.93	Kullnig-Gradinger <i>et al.</i> (2002)
31	<i>T. stromaticum</i>		CBS 101875	Druzhinina <i>et al.</i> (2005)
32	<i>T. tomentosum</i>		CBS 349.93	Kullnig-Gradinger <i>et al.</i> (2002)
33	<i>T. viride</i>		CPK 419	C. P. Kubicek (unpublished data)
34	<i>T. viridescens</i>		ATCC 28020	Druzhinina <i>et al.</i> (2005)

*ATCC, American Type Culture Collection, Manassas, VA, USA; CBS, Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands; CPK, fungal strain collection at the Institute of Chemical Engineering, TU Wien, Austria; DAOM, Eastern Cereal and Oilseed Research, Ottawa, Canada; GJS, collection of Gary J. Samuels, United States Department of Agriculture, Beltsville, MD, USA; QM, Quartermaster Collection of the US Army, Natick, MA, USA.

Döhren & others, unpublished). Characteristic A-domains were identified by positional conservation and non-ribosomal code analysis (Rausch *et al.*, 2005).

Molecular phylogenetic methods. Sequences of rRNA spacer regions ITS1 and 2, and *rpb2* were taken from the multilocus database (www.isth.info), and were aligned using Genedoc 2.6 (Nicholas & McClain, 1987) and visually corrected. Phylogenetic trees were constructed by the neighbour-joining method (Saitou & Nei, 1987) with 1000 bootstrap replicates, using the computer program MEGA 3.1 (Kumar *et al.*, 2004).

For phylogeny of the peptaibol sequences, the amino acid sequences were first transcribed into a binary matrix in which the presence of a given amino acid at a given position was coded as 1 and its absence as 0. One and zero were subsequently replaced by the nucleotides T and A, and the matrix transcribed into a nucleotide alignment appropriate for analysis in MEGA 3.1 as above.

Tests for evolutionary mechanisms. To test the fit of the sequences to the model of neutral evolution, the *D* test statistic proposed by Tajima & Nei (1984) was computed with the DnaSP program (Rozas *et al.*, 2003). The extent of nucleotide divergence was estimated by using the uncorrected *p* distance (Nei & Kumar, 2000). The proportions of synonymous (p_s) and non-synonymous (p_n) differences per site were calculated by the modified Nei-Gojobori method implemented in DnaSP (Zhang *et al.*, 1998).

RESULTS

MS analysis of low-molecular-mass peptides

Our methodology allowed the analysis of the proteomic fingerprint within two mass ranges, 2000–20 000 Da and 1000–2000 Da, at the same time using the same sample

target. Both types of spectra are useful for the identification of specific strains (Neuhoﬀ *et al.*, 2007a, b). Peptaibol masses are predominantly in the 1000 to 2000 Da range. All peptaibols from the 32 *Trichoderma/Hypocrea* strains were detected as sodium $[M + Na]^+$ or potassium $[M + K]^+$ ions directly from microgram amounts of unfractionated cells. Peptaibols are easily released from fungal mycelia grown on plates. Although the sensitivity of detection has not been tested, peptaibols might be detected in single mycelial segments, as peptides have been detected in microcolonies and single filaments of cyanobacteria comprising about 100 bacterial cells (Fastner *et al.*, 2001; Welker *et al.*, 2004a, b, 2006).

Verification of the method by application to *H. lixii*

To document the quality of our method, a representative spectrum is shown for *H. lixii* (*T. harzianum* CBS 226.95; Fig. 1). Three sets of mass peaks were observed, corresponding to three peptaibols of SF1, SF4 and a third yet unknown subfamily.

Set 1, which exhibited the most abundant mass peak (m/z 1214) with most set members deviating in 14 Da ($-CH_2-$) steps, is similar to trichovorins/trichorozins, which belong to the 11-residue peptaibols of SF4. This set of mass peaks has been defined as group 6 in our evaluation (Table 2). An identical set of mass peaks has recently been reported for *H. lixii* CECT 2413 (Vizcaino *et al.*, 2006), for which the species identity to the *H. lixii* anamorphic state has also been established (Hermosa *et al.*, 2004). We have detected sets with identical masses in 16 other strains used in this study (see Table 2). As we only have molecular masses of these compounds, we cannot fully identify the respective peptaibols, however.

Set 2, with the most abundant mass at m/z 1572, represents a so far unknown peptaibol group, group 14 in Table 2 (see below). The nature of these presumably 16-residue peptaibols remains to be shown. This group of peptides has not been detected in strain *H. lixii* CECT 2413 (E. Monte, personal communication).

Set 3 in Fig. 1, characterized by the most abundant mass peak at m/z 1800, is a representative of peptaibol group 19 (Table 2), with similar masses to the well known SF1 18-residue trichokindins and trichorzins of the MA type

(containing Val in the terminal position, differing from Leu in the HA type, and Trp or Phe in the PA type), isolated from a putative *H. lixii* from soil in Nara, Japan (Iida *et al.*, 1994) and strain M-922835 (Goulard *et al.*, 1995), respectively. We thus speculate that these two isolates may indeed be *H. lixii* as well.

Two additional sets of peaks with very low intensity can be seen between sets 2 and 3, and preceding set 1; they may belong to the 10- and 17-residue compounds, possibly derived by proteolytic cleavage, but they have not been investigated further.

Peptaibol formation by different species of the genus *Trichoderma*

Having assessed that our method works well and provides data comparable to peptaibol production data by others, we then analysed the production profiles of 34 strains representing 28 species (Table 2): in total, we could identify 26 groups of peptaibols by mass data. The masses resemble H^+ , Na^+ or K^+ adducts of peptaibols, and all spectra have been compared for overlapping profiles. To identify known peptaibols on the basis of their mass data and to recognize possible new structures, we compared them with the monoisotopic masses of the relevant peptaibols from the Birkbeck database and complemented it with recently published data (Degenkolb *et al.*, 2006a, b; Krause *et al.*, 2006; Krause, 2006), thereby combining analogues of single isolates into groups. Thus all alamethicins and all paracelsins are included in groups 26 and 25, respectively, with other SF1 members.

As is shown in Table 2, some groups combine several sets of compounds with identical masses. Thus group 1 combines the trikoningins and trichogins, 11-residue peptaibols of the SF5 type, both with masses of 1052, 1066 and 1082 Da. These differ in mass from 11-residue trichovorins and trichorozins of the SF4 type, defined as group 6.

Members of the 20-residue trichosporins, paracelsins, saturnisporins, suzukacillins, trichocellins, trichokonins, longibrachins and trichoareocins all share masses 1922, 1936, 1950 and 1964 Da, and are thus combined into our group 25. The rationale of combining peptaibiotics by molecular masses originates from a comparison of their

Fig. 1. ICMS view of the 1000–2000 Da region of mycelia of *H. lixii* CBS 226.95. Three major sets of peptaibol peaks with highest mass peaks at 1213.7 (1), 1571.9 (2) and 1799.9 Da (3) are observed. Minor peak groups below 1 and between 2 and 3 have not been investigated. The peaks represent Na^+ and K^+ adducts of the molecular ions.

Table 2. Groups of peptaibols in *Trichoderma/Hypocrea*

Group no.	Prominent masses of Na ⁺ /K ⁺ adducts (maxima underlined)	Subfamily/type*	Example compounds	Detected in strains:†
1	1061, 1075, <u>1089</u> , <u>1105</u>	SF5	Trikoningin/trichogin/lipostrigocin type	22, 23, 28, 30
2	<u>1122</u> , 1137, <u>1151</u>	SFX1	Unknown	1, 2, 32,
3	1127, <u>1141</u> , 1155	SFX2	Unknown	33
4	1189, <u>1193</u> , <u>1207</u> , 1221, 1229, 1245	SFX3	Unknown	11, 13, 19
5	<u>1190</u> , 1192, 1206	SFX4	Unknown	22
6	1184, <u>1198</u> , <u>1199</u> , <u>1212</u> , <u>1213</u> , <u>1214</u> , <u>1226</u> , <u>1227</u>	SF4	Trichovorin/trichorozin type	2, 3, 5, 6, 7, 8, 9, 12, 14, 15, 17, 20, 21, 23, 24, 25, 26, 29, 32, 34
7	1345, 1347	SFX5	Unknown	7
8	1387, 1401, <u>1403</u> , 1417	SF4	Harzianin/trichovirin I – type I	11, 22
9	1439, 1453, <u>1467</u> , <u>1468</u> , 1482, 1483, 1497, 1498, 1512, 1513	SF4	Harzianin/trichovirin I – type II	34
10	1445, <u>1459</u> , 1473, 1487, 1489, 1503	SFX6	Unknown	1, 21, 23
11	1465, 1466, 1467, <u>1481</u> , <u>1482</u> , 1495, <u>1496</u> , <u>1497</u> , <u>1498</u> , 1511	SF4	Harzianin/trichovirin I – type III	5, 12, 15, 17, 25, 26, 29, 32
12	1465, 1479, 1493	SFX7	Unknown	6
13	1540, 1562, 1564, 1565, 1573, 1584	SFX8	Unknown	1, 4, 12, 15, 22, 23, 25, 28
14	1543, 1548, 1556, 1570, 1572, 1583, 1586, 1600	SFX9	Unknown	2, 7, 8, 26
15	1568, 1569, 1576, 1577, 1582, 1583, 1590, 1591, 1604, 1605, 1618, 1620, 1634	SFX10	Unknown	23
16	1640, 1654	SFX11	Unknown	10
17	<u>1701</u> , 1715, 1717, 1729, 1732, 1746	SFX12	Unknown	1
18	1712, 1726, 1728, 1740, 1742, <u>1754</u> , 1756, <u>1757</u> , 1770, 1771, 1772, 1786	SF1/type I	Trichovirin II type	18, 16, 17, 34
19	1756, 1770, <u>1784</u> , 1785, <u>1786</u> , <u>1798</u> , <u>1799</u> , <u>1800</u> , 1801, 1813, 1814, 1815, 1828, 1829	SF1/type II	Trichokindin/trichorzin MA type	2, 3, 5, 8, 12, 25, 26, 29, 32
20	1774, 1775, <u>1789</u> , 1790, <u>1803</u> , 1804, 1816, 1817, 1818, 1833	SF1/type unknown	Unknown	33
21	1797, 1811, 1825	SF1/type III	Trichostromaticin type	10, 28, 31‡
22	1804, 1818, 1832, 1833, <u>1834</u> , 1846, 1848, 1849, 1862, 1863	SF1/type II	Trichorzin PA type	6, 15
23	1898, 1899, <u>1913</u> , 1927, 1928	SF1/type VI	Trichostrigocin type	30
24	1912, 1918	SF1/type VII	Trikoningin type	23
25	1946, 1959, <u>1960</u> , <u>1962</u> , 1973, 1974, <u>1975</u> , 1976, 1987, 1989, <u>1990</u> , <u>2002</u> , 2004, <u>2005</u> , 2016, 2019, 2020, 2023	SF1/type VIII	Paracelsin type	7, 11, 14, 20, 24, 27
26	1933, 1947, 1948, 1949, 1950, 1963, 1964, <u>1978</u> , 1980, 1982, 1986, <u>1993</u> , 2000, 2001,	SF1/type IX	Alamethicin type	1, 19, 22

*SFX1–SFX12 describe the 12 unknown mass sets that do not match SF1–9 defined by Chugh & Wallace (2001).

†Strain numbers refer to Table 1.

‡We detected no products in this survey; data were taken from Degenkolb *et al.* (2006a).

sequence data. While there are a small number of positions with defined functional properties, such as Gln/Asn or Pro, the vast majority of positions have variable aliphatic side chains. These include Ala, with successive additions of methylenes to arrive at Aib, Val, Iva, Leu or Ile. Respective peptaibol synthetase modules may have either defined or less defined specificities, thus leading to sets of compounds differing in masses by 14 Da. A phylogenetic analysis of adenylate domains indeed indicates a clustering of all of

these domains (results not shown). Non-conservative substitutions, like Ala–Ser transitions introducing a functional hydroxy group or Leu–Phe replacing an aliphatic side chain by an aromatic residue, then lead to differing mass deviations. Observed mass groups can be considered as diagnostic data to identify either new compounds with yet unknown masses, or mass ranges extending from known groups to more or less hydrophobic analogues.

Of the 26 sets of mass peaks identified, 12 correlate with mass sets of known compounds. Most of the remaining unknowns are found in single strains only, and only a few are widely distributed. Most strains produce two or more peptaibol groups. Twenty strains of 18 species produce 11-residue compounds of the trichovirin/trichorzin type (group 6), and nine species produce 18- and 20-residue peptaibols of the trichokindin/trichorzin type (group 19) as well as the paracelsin/alamethicin type (groups 25 and 26). Hence this survey reveals that the diversity of peptaibol structures in *Trichoderma/Hypocrea* is far from being recognized. Of the 26 mass sets, 14 cannot be correlated with any known peptaibols, and are likely to be as yet undescribed compounds.

Distribution of SF1-type compounds

SF1 contains 24 named compounds with their respective analogues (Chugh & Wallace, 2001), of which 21 originate from strains of the genus *Trichoderma/Hypocrea*. An additional SF1 group of compounds, the trichoareocins, has recently been described by Degenkolb *et al.* (2006b). All of these peptaibols range in size from 18 to 20 amino acid residues.

18-residue peptaibols. We found five groups characterized by 18-residue compounds, i.e. groups 18, 19, 20, 21 and 22 (Table 2).

Group 18 has been defined on the basis of 12 analogues of the trichovirin structural family, which are formed by *T. viride* NRRL 5243 (= *H. lixii*) in submerged cultivation (Fig. 2; Jaworski *et al.*, 1999). We have detected similar

masses in *T. asperellum*, *H. virens* (the former ex-type strain of *T. flavofuscum*, a species name which has been abandoned) and *H. tawa*. Masses of trichovirin analogues may overlap with masses of the trichorzin HA type, but the dominating analogues have lower masses due to their lower Aib content. Our mass data correlate with the results of Brückner *et al.* (1985) for *T. asperellum* (then called *T. viride*) NRRL 5242, in which a family of 13 trichotoxins was described.

Group 19 contains 18-residue peptaibols with major mass peaks in the range m/z 1784–1800, corresponding to trichokindins and trichorzins of the MA type. This group is a major structural component in *Trichoderma* and has been detected in *H. aureoviridis*, *H. gelatinosa*, *H. lixii*, *T. longipile*, *T. oblongisporum*, *H. semiorbis*, *T. spirale* and *T. tomentosum*. A comparison of the 14 known structures provides the consensus peptide sequence **AcUSAU(U,J)Q(I,U)LU(A,G)(U,J)UPLUUQ(L,I,V)ol**, where Ac=acetyl-, U=Aib, J=Iva and all others are standard single-letter amino acid code; the terminal 'ol' indicates reduction to the amino alcohol, and invariant residues are shown in bold. Comparing this structure to similar SF1 compounds (group 18, trichorzin A/trichotoxin type; group 20, trichostromaticin type, with Gln in position 7 instead of 6) we found Ser-2→Gly-2 and Ser-2→Ala-2 transitions, and Phe/Trp aromatic terminal amino alcohols (group 22, trichorzin PA type) replacing Leu/Val, leading to different masses, but with identical invariant residues displaying similar properties (Fig. 2).

Group 20 contains an unknown mass set found only in *T. viride* (= *T. viridescens*) ATCC 28020.

Group 18 - Gly(Ala) in position 2

Trichotoxins	AcUGULUQUUUUAUPLUUQVo1	<i>T. asperellum</i> NRRL 5242
	A E A JE	
Trichorzins HA	AcUGAUUQUVUGLUPLUUQLo1	' <i>H. lixii</i> ' M-903602
	J J	
Trichotoxins A40	AcUGULUQUUUUAUPLUJQVo1	<i>H. lixii</i> CBS 226.95
	A E A A UE	
Trichotoxin	AcUGULUQUUUUAUPLUUQVo1	' <i>H. lixii</i> '
	A E A A JE	
Trichovirins II	AcUGALUQJUVUGJUPLUUQLo1	<i>H. lixii</i> NRRL 5243
	A U	
	J A	

Group 19 - Ser in position 2

Hypomurocins	AcUSALUQUVUGUPLUUQVo1	<i>H. atroviridis</i> IFO 31288
	A J L	
Trichokindins	AcUSAUQUJLUUAUPLUUQIo1	' <i>H. lixii</i> '
	J U J L	
Trichorzins MA	AcUSAUQUJLUGLUPLUUQVo1	' <i>H. lixii</i> ' M-922835
	J J	

Group 21 - Ala in position 2, Gln in position 7

Trichostromaticins	AcUAUAUQUVUGUPLUUQLo1	<i>T. stromaticum</i> CBS 101875
	VV V	

Group 22 - terminal W,F

Trichorzins PA	AcUSAUJQUVUGLUPLUUQWo1	' <i>H. lixii</i> ' M-902608
	J J F	

Fig. 2. Known 18-membered peptaibols of SF1 and their positioning in the mass groups defined in Table 2. Sequences are given in standard single-letter code (Ac=acetyl-, U=Aib, J=Iva, ol represents the C-terminal amino alcohol, V and L in trichostromaticins stand for Vxx and Lxx, representing either Val/Iva or Leu/Ile, respectively). Alternative positional replacements of analogues within this compound group are shown in the lines below. Large groups of analogues are formed due to relaxed substrate binding of the respective adenylate domains. Strains in quotation marks have not been clearly identified and have been described as *T. harzianum*/*H. lixii* (compare with Tables 3 and 4).

Group 21 contains the recently described trichostromatic-type peptaibols found in *T. stromaticum* BBA 70636/70638 (Degenkolb *et al.*, 2006b). We found a similar mass set in the ex-type strain of *T. pubescens* DAOM 162162.

Finally, group 22 corresponds to the trichorzin PA type, where a change of the terminal amino acid to Trp leads to a mass increase of 87 Da. The PA type has been found in the putative strain '*H. lixii*' M-902608, selected by Roquebert and colleagues for its antagonistic properties (Duval *et al.*, 1997; Leclerc *et al.*, 1998a, b). We detected corresponding masses in *H. strictipilosa* (ex-type strain of *T. fasciculatum* CBS 118.72, a species now abandoned) and *H. hunua*.

19-residue peptaibols. Mass groups 23 and 24 correspond to the trichostrogocin and trikoniningin types, respectively. Our data with the ex-type strain of *T. strigosum* are consistent with a detailed analysis of the same strain and strain *T. cf. strigosum* CBS 119777 (Degenkolb *et al.*, 2006a) which produce trichostrogocins and tricholongins, respectively.

20-residue peptaibols. This group of long peptaibols can be separated into two mass groups (25 and 26), depending on whether the position 2 amino acid is Ala (paracelsin type) or Pro (alamethicin type). Differentiation of these two groups by molecular ions is not trivial, as the 16 Da difference can be reduced to only two mass units when conservative Ala/Aib, Aib/Val or Aib/Iva exchanges take place, and terminal deaminations at Gln may further lead to an increase in a single mass unit. We found the paracelsin type as expected in *T. reesei*, and its close phylogenetic neighbours *H. schweinitzii* (= *T. citrinoviride*) and *T. longibrachiatum*, but also in members of other clades/sections, such as *T. pubescens* CBS 119776, *H. pachybasioideis* and *T. strictipile*.

In *H. jecorina*, the 20-residue paracelsin SF1 type has been investigated in strain QM 9414 (=ATCC 26921=IMI 192656) (Brückner & Graf, 1983; Brückner *et al.*, 1984), and additional paracelsins have been detected in a commercial preparation obtained from Fluka BioChemika (Pocsfalvi *et al.*, 1997). Our investigation of *H. jecorina* QM 6a, the progenitor of the cellulase-overproducing mutant QM 9414, has confirmed the presence of paracelsins. In addition, we detected two compound sets of 18 and 19 residues, presumably derived by N-terminal cleavage and deletion of 1 amino acid residue. Since the genome sequence database of *H. jecorina* QM 9414 (<http://gsphere.lanl.gov/trire1/trire1.home.html>) contains only two putative NRPS genes encoding 20- and 14-residue peptaibol synthetases, respectively (H. von Döhren & others, unpublished), these two sets of peptaibols originate as two products from one synthetase and one additional product by peptidolytic modification.

SF1 compounds of 20 (longibrachin) and 19 residues (tricholongins) have been described in a putative *T.*

longibrachiatum isolate (Leclerc *et al.*, 2001; Rebuffat *et al.*, 1991; Auvin-Guette *et al.*, 1992). We detected masses corresponding to longibrachins in the *T. longibrachiatum* ex-type strain CBS 816.68. *T. longibrachiatum* and *T. ghanense*, another member of section Longibrachiatum, have been shown to form eight longibrachin analogues, and various C-terminal degradation products have been investigated in detail by Krause (2006). Masses of the alamethicin type (group 26) have been detected in *H. atroviridis* P1=ATCC 74058, *T. brevicompactum* and *T. hamatum*.

Distribution of SF4 type of peptaibols

SF4 contains 11- and 14-residue compounds. The latter, best known as harzianins, have a characteristic spacing of three Pro residues in positions 5, 9 and 13. Interestingly, the 11-residue group can be derived from harzianins by internal deletions of the amino acids Leu, Aib and Pro. Work on *H. lixii* (Augeven-Bour *et al.*, 1997) and *H. jecorina* (H. von Döhren & others, unpublished) suggests that both types of compounds originate from one peptide synthetase. Known compound structures have been compiled in Fig. 3.

11-residue peptaibols. Groups 2–6 (Table 2) correspond to 11-residue peptaibols; however, only group 6 compounds are known. These structures contain 5 invariable and 6 variable amino acid residues (Fig. 3). In addition to the two *H. lixii* strains mentioned above, the group 6 mass set has been found in 19 of the strains

11-residue peptaibols

Group 6 peptaibols

Trichovorins	AcUNVVUPLLUPVo1 QLL L	' <i>T. viride</i> '
	II	
Trichorzins	AcUNILUPILUPVo1 Q L	' <i>H. lixii</i> '
Harzianin HBI	AcUNLIUPJLUPLo1	' <i>H. lixii</i> '
Harzianin HKVI	AcUNIIUPLLUPLo1	' <i>T. pseudokoningii</i> '
Trichofumin A,B	AcVQLVUPLLUPLo1 L	' <i>Trichoderma</i> sp.'
Hypomurocins	AcUQVVUPLLUPLo1 IL I I	<i>H. atroviridis</i>

Fig. 3. Known 11-membered peptaibols of SF4. Sequences are given in standard single-letter code (Ac=acetyl-, U=Aib, J=Iva, ol represents the C-terminal amino alcohol). Alternative positional replacements of analogues within this compound group are shown in the lines below. Large groups of analogues are formed due to relaxed substrate binding of the respective adenylate domains. Strains in quotation marks have not been clearly identified, and have been described as *T. harzianum*/*H. lixii* (compare with Tables 3 and 4).

investigated (Table 1), and is thus the most widely distributed type of peptaibol. For descriptive purposes, members of SF4 group 6 can be characterized as 11-residue peptaibols of the type AcX(N,Q)XXUPXXUPXol, with fixed positions 5 (Aib), 6 (Pro), 9 (Aib) and 10 (Pro), Asn or Gln in position 2, X being variable amino acids, and a terminal leucinol or valinol. The monoisotopic molecular masses are 1161, 1175, 1189 and 1203 Da.

14-residue peptaibols. Of the 14-residue type, 12 harzianins have been described with the general structure AcU(N,Q)LUP(A,S)(V,I)UP(J,U)LUP(L,V)ol, all isolated from the undefined *H. lixii* strain M-903603 (Rebuffat *et al.*, 1995; Goulard *et al.*, 1995). The similar trichovirin I type with the general structure Ac(U,J)(N,Q)LUP(A,S)VUP(J,U)LUPLol has been described in seven variants from an undefined strain of *T. viride* (Brückner & Koza, 2003). Note the conserved residues in the overall structure AcxxLUPxxUPxLUPxol, with all other residues undergoing conservative replacements with the exception of residue 6, which can be Ala or Ser. All 19 variants provide a set of eight monoisotopic masses of 1385, 1399, 1401, 1413, 1415, 1427, 1429 and 1443 Da. Respective MH^+ , MNa^+ or MK^+ ions were observed in groups 8, 9 and 11, however, with additional unknown masses. As the peaks of highest intensity range from m/z 1403 to 1468 and 1482 to 1496 (Table 2), these three groups have been defined accordingly. Groups 9 and 12 are in the same mass range, but do not correspond to known compounds. The low mass SF4 type has been detected in *T. croceum* and *T. hamatum*, an intermediate type in *H. virens*, and the high mass type in eight strains, including *T. longipile*, *T. tomentosum*, *T. oblongisporum* and *H. semiorbis*, all belonging to section Pachybasium B. We did not detect the SF4 type in our *H. lixii* and *T. viride* strains.

Distribution of SF5 type of peptaibols

The SF5 type comprises several lipopeptaibols. A group of octanoyl-capped 11-residue peptaibols, first described as trichogin/trikoningin from *T. longibrachiatum* and *T. koningii* (Auvin-Guette *et al.*, 1992, 1993), has recently been extended by 14 lipostrigocins isolated from strains of *T. strigosum* and *T. pubescens* (Degenkolb *et al.*, 2006b). In addition, 7- and 10-residue lipopeptaibols (lipopubes-cins) have been described, which are likely to originate from a closely related peptaibol synthetase, or even a single synthetase. Also, trichodecenins, 7-residue decenoyl-capped peptaibols, isolated from an uncharacterized *T. viride* (Fujita *et al.*, 1994), are structurally related.

All 11-residue compounds have the structure Oc(U,Vx)G(Vx,Lx)UGG(Vx,Lx)UG(I,Lx)Lol, where Oc= octanoyl, Vx=Val/Iva and Lx=Leu/Ile, with possible monoisotopic masses of their Na^+ adducts of 1061, 1075, 1089, 1103 Da. This group has indeed been detected as expected in *T. koningii*, *T. strigosum*, *T. pubescens*, and also in *T. hamatum*, which are all members of the section

Trichoderma. We have not detected this group in *T. longibrachiatum* from section Longibrachiatum.

Phylogenetic analysis of peptaibol formation and correlation with species phylogeny

Having assessed the peptaibol pattern formed by the 28 species of *Trichoderma*, we investigated whether the formation of any of the groups would correlate with individual phylogenetic clusters or branches. To this end, we used a fragment of the RNA polymerase subunit B gene, which has previously been shown to result in the most reliable and resolved phylogeny of all marker genes available today (Druzhinina *et al.*, 2005; Druzhinina & Kubicek, 2005). The results (Fig. 4) confirm the established clusters for sections Longibrachiatum, Trichoderma and the Pachybasium B clade, as described by Kullnig-Gradinger *et al.* (2002). Correlation of the individual groups of peptaibols with these clusters appears to be poor.

Ten groups were found only in a single species (group 3, *T. viride*; group 5, *T. hamatum*; group 9, *H. viridescens*; group 12, *H. lixii* CBS 273.78; group 16, *H. minutispora*; group 17, *H. atroviridis*; group 12, *H. hunua*; group 20, *T. viride*; group 23, *T. strigosum*; group 24, *T. koningii*). Group 6 was most abundant (found in nine species) and was distributed throughout all phylogenetic clades. However, some groups are more consistently found in some phylogenetic clades, notably group 1, which was only found in members of section Trichoderma, and group 19, which was formed only by species from section Pachybasium B. Also, members of group 25 were consistently found in species of section Longibrachiatum, albeit they were also formed by *H. pachybasioides* and *T. pubescens*. When the chain length of the peptaibols in the respective subfamilies is taken into consideration, it becomes apparent that members of Pachybasium B only produce 18-residue peptaibols, whereas 19-residue peptaibols are only formed by members of section Trichoderma, with 20-residue compounds being formed by the Pachybasioides (=Polysporum) clade, section Trichoderma and section Longibrachiatum.

One might argue that the fine structure (=sequence) of peptaibols within the same group might still be different enough to attribute individual compounds to selected phylogenetic clades. We tested this with group 25 and 26, which both include 20-residue peptaibols (section Longibrachiatum – longibrachins, trichobrachins, trichoaur-eocins, paracelsins and saturnisporins; section Trichoderma – atroviridins; section Pachybasium B – polysporins and alamethicins). To this end, the peptaibol amino acid sequence was transformed into a binary matrix, which also took the ambiguity of individual positions into account, and then this matrix was transformed into an artificial nucleotide sequence on which we performed neighbour-joining analysis. The results show that this approach grouped the 20-residue peptaibols into five clusters whose members shared the highest similarity (Fig. 5). These clades exhibited no relationship to gene phylogeny, however. We conclude that

Fig. 4. Phylogenetic tree of *Trichoderma/Hypocrea* species based on *rpb2* gene sequence alignment, and correlation with production of peptaibol subfamily members. Some groups appear to be restricted to certain sections or clades. Shaded boxes indicate peptaibol groups, which are consistently found only in one of the clades or sections of *Trichoderma/Hypocrea*. Thus group 2 has been found only in section *Trichoderma*, groups 11 and 19 only in strains from section *Pachybasium* B, while groups 6 and 25 are found in certain strains of all sections.

Fig. 5. Phylogenetic analysis of 20 residue peptaibols. The approach used is described in detail in Methods. Sequences are given in standard single-letter code (Ac=acetyl-, U=Aib, J=Iva, ol represents the C-terminal amino alcohol). Alternative positional replacements of analogues within this compound group are shown in the lines below. Producers of trichokonins, atroviridins and alamethicins (shaded area) belong to section *Trichoderma*, the boxed area combining trichosporins and polysporins are section *Pachybasium* B strains, and all others have been found in section *Longibrachiatum*. Thus no clear correlation is obtained from this sequence analysis approach.

concordant with the evolution of the genus, but still contains a background of ancient species-specific patterns.

Patterns and further classification of peptaibol structures – typing of SF1 peptaibols

Peptaibol sequences have been aligned using CLUSTAL W to define nine distinguishable subfamilies (Chugh & Wallace, 2001). In the meantime, the number of known peptaibol sequences has more than doubled. Relevant new structures have been included in the Discussion. In the genus *Trichoderma/Hypocrea*, so far only SF1, 4, 5 and 9 have been described (Szekeres *et al.*, 2005).

SF1 comprises more than half of all the structures which are characterized by chain lengths between 18 and 20 residues. These ‘long’ peptaibols all share a Gln residue in position 6 or 7 which fundamentally contributes to the channel-forming properties (Chugh *et al.*, 2002). In addition, one or two Gln/Glu residues are consistently located in the C-terminal region, preceding the terminal alcohol.

To further classify members of SF1, various approaches are possible. We have started from plain chemical systematics, which defines types of 18, 19 or 20 residues, and have then further compared these types with respect to the positions of functional side chains and sequence similarities. By using this approach, we can define nine major types, as shown in Fig. 6, within this family. Types I and II are all 18-residue peptaibols with Gln in position 6 and either Gly or Ser/Ala in position 2. It is also intriguing that all of them contain Aib in position 1, and the sequence Pro-Leu-Aib in positions 13–15. Type III peptaibols are 18-residue compounds with Gln in position 7. It is noted that in this case Ala is conserved in position 2, and there is an amino acid insertion at position 6 with simultaneous loss of 1 amino acid between residues 9 and 12. Type IV peptaibols are 18-residue peptaibols, similar to type III, but with an additional Gln in position 11. Among the 19-

residue peptaibol types (V–VII), type V is characterized by a Phe inserted in position 1, thus moving Gln at position 6 to position 7. Type VI mostly displays a Phe in position 3, but this position is sometimes substituted by Aib. Types VIII and IX are all 20-residue compounds with either Pro or Ala/Gly in position 2, respectively, and which also show conservation of Aib-1, Gln-7, Pro-14, Aib-16, Gln-18 and Gln-19. It is intriguing to note that type VII peptaibols may have been derived by deletion of residue 6 from type VIII (trichokonin V). However, peptaibols of this type are also found as a unique set of compounds (e.g. trichorzianins and trikoningin).

If the distribution of these types within different species and clades of *Trichoderma/Hypocrea* (cf. Druzhinina *et al.*, 2006) is compared with their established phylogeny, some concordance is observed: e.g. species from section Longibrachiatum produce both type VIII and type IX compounds, but none of the shorter types I–VI. On the other hand, strains of the section *Trichoderma* B contain the longer types VIII and IX, the short types I–III and type IV, which is not produced by strains from other sections. In addition, *H. pachybasioides* has been reported to produce type VIII and IX peptaibols (Iida *et al.*, 1990, 1993, 1999; New *et al.*, 1996; Fujita *et al.*, 1981), but the actual strain has not yet been identified by sequence analysis. Notably, only one type of SF1 peptaibol is found in each strain, with the exception of type VII which probably arose by a biosynthetic deletion of the position 6 residue from a corresponding type VIII peptaibol.

Conserved features within SF1 type peptaibol synthetases

In view of the fact that peptaibol synthetases belong to the NRPSs, they are expected to contain, in addition to the unique N-terminal acetyltransferase domain and the C-terminal reduction domain, 18–20 modules of a sequence

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
I	Ac	U	G			Q	U						P	L	U		Q			
II	Ac	U	SA			Q	U						P	L	U		Q			
III	Ac	U	A	AU		Q	U						P	L	U		Q			
IV	Ac	U	A	U		Q	U			Q			P	L	U		Q			
V	Ac	F	U	AS	A		Q	G					P		U		Q			
VI	Ac	U	G	F	U		Q	U					P		U		Q	Q		
VII	Ac	U	AG	AU	U		Q	U					P		U		Q	Q		
VIII	Ac	U	AG		A		Q	U					P		U		Q	Q		
IX	Ac	U	P		A		Q	U					P		U		Q	Q		

Fig. 6. Classification of 18- to 20-residue peptaibols. The acetylated N terminus is on the left, the reduced C terminus is on the right. Amino acids are given in standard single-letter code (Ac=acetyl-, U=Aib, J=Iva); two letters at a single position indicate that both variants are observed. Each box represents a position within the peptaibol. The diagram also summarizes the expected organization of the respective NRPSs, each box representing a module. Phe positions (F) have been framed since they represent unique replacements, indicating module replacements. Open boxes can be assigned to any of the amino acids Gln, Ala, Ser, Val, Leu, Ile, Iva or Aib, and in the case of terminal positions Val, Leu, Phe or Trp.

of an amino-acid-activating adenylation domain, a carrier protein domain and a condensation domain. Each of the positional boxes in Fig. 6 in theory therefore predicts the organization of modules in the respective NRPSs. So far, the gene encoding trichovirin synthetase from *H. virens* (section Pachybasium B) has been isolated and the amino acid sequence of the respective NRPS has been deduced. In addition, putative peptaibol synthetase genes have been detected in the genome sequences of *T. reesei* (section Longibrachiatum) and *H. atroviridis* (section Trichoderma) (H. von Döhren & others, unpublished). These three synthetases produce 'long' peptaibols of SF1 type I (trichovirins), SF1 type VIII (paracelsins) and SF1 type IX (atroviridins). In addition, the terminal region of an SF1 type I peptaibol synthetase from *H. lixii* (section Pachybasium B) has been sequenced (Vizcaino *et al.*, 2005). No genes encoding peptaibol synthetases synthesizing products of SF4, 5 and 9 have been cloned yet.

Phylogenetic analysis of conserved adenylation domains of SF1 type synthetases

Since the peptaibols produced by these strains differ in some of their modules, the use of the full-length amino acid sequence for phylogenetic analysis is not possible. We therefore decided to only use positionally conserved domains in the respective synthetases. As shown in Fig. 6, such domains are located in the acyl transferase modules, in the Aib-adding module 2 (with the exception of type V found in boletusin, peptaivirins and chrysospermins produced by strains whose species identity is uncertain), in the positionally conserved central Gln module (position 6 or 7), in the Pro modules (positions 13 or 14), in the Aib module (positions 15 or 16) and in the late Gln modules (positions 18 and 19), although the origin of these tandem Gln modules in types VIII and IX may be due to module duplication and their use in phylogenetic analysis would thus be corrupted. We therefore compared nine Gln-activating domains from trichovirin synthetase, paracelsin synthetase, atroviridin synthetase and a type I NRPS from *H. lixii*, as well as early Aib-activating domains from the first NRPS modules, Pro-activating domains and late Aib-activating domains following next-but-one to the Pro domains. The phylogenetic analysis shows that the domains cluster with high bootstrap support not only with respect to the recognized amino acid, but also with respect to their relative position within the peptaibols (Fig. 7a). Similar phylogenetic patterns were obtained when only the Gln-activating domains or only the Pro- and Aib-activating domains were analysed (data not shown). When the branching pattern within these clusters was compared to the established species phylogeny based on *rpb2* or ITS1 and 2, discordance was noted (Fig. 7b). Within most clusters, the domains from *H. virens* and *H. jecorina* were most similar, and *H. lixii* was most distant. These data indicate that the phylogeny of the adenylation domains for these amino acids does not parallel the evolution of these three species.

Fig. 7. Phylogenetic evaluation of domain evolution based on amino acid sequences. (a) Neighbour-joining analysis (1000 bootstrap replicates) of all Pro, Aib (early and late) and Gln (early and late) domains. Adenylate domains of paracelsin synthetase (PAR), harzianin type synthetase (HAR) from *H. jecorina*, trichovirin synthetase (TVA) and atroviridin synthetase (ATRO), with the numbering indicating the domain position in the respective synthetase, show distinct clustering in Gln-, Aib- and Pro branches. Furthermore, Gln and Aib branches show subclusters of early and late positions. (b) Phylogeny of species based on the *rpb2* nucleotide sequence. The tree structure is typical and supported by ITS1 and 2: *H. virens* and *H. harzianum* are the most closely related. Whether *T. reesei* (= *H. jecorina*) or *T. atroviride* is closer to *H. harzianum*/*H. virens* is not resolved with only these four species, as both *T. reesei* and *T. atroviride* are roughly the same genetic distance from each other and from *H. harzianum*/*H. virens*.

Phylogenetic analysis gives rise to an extended substrate-binding pocket consensus

All Gln-activating domains have identical non-ribosomal substrate-binding sequences, DGGMVGGNYK, except for domain 18 of atroviridin synthetase, where Val is replaced by Ile. A comparison of 34 amino acids in close vicinity to the

Table 3. Peptaibol production in phylogenetically defined *Trichoderma* strains

Species	Strain	Subfamily*	Compound name†	No. of amino acid residues	Monoisotopic masses of subfamily members
<i>T. harzianum</i> = <i>H. lixii</i>	34563=CBS 226.95	SF1	Trichotoxins	18	1690, 1704, 1718 (M)
<i>T. muroiana</i> = <i>T. atroviride</i> / <i>H. atroviridis</i>	IFO 31288	SF1	Hypomurocins B	18	1702, 1718, 1732, 1746 (M)
	IFO 31288	SF4	Hypomurocins A	11	1061, 1075, 1089 (M)
<i>T. reesei</i> = <i>H. jecorina</i>	QM 9414, ATCC 26921, IMI 192656	SF1	Paracelsins	20	1908, 1922, 1936 (MH ⁺)
<i>T. asperellum</i>	NRRL 5242	SF1	Trichotoxins A-50	18	1660, 1674, 1688, 1702, 1716, 1730 (M)
<i>T. polysporum</i>	TMI 60146	SF9	Trichopolys	7	664, 678, 692
<i>T. atroviride</i>		SF1	Atroviridins	20	1962, 1976, 1990 (MH ⁺)
<i>T. atroviride</i>	P1, ATCC 36042	SF1	Trichorzianins	19	1920, 1924, 1938, 1948, 1949, 1963 (M)
<i>T. aureoviride</i> = <i>T. citrinoviride</i>	IMI 91968‡	SF1	Trichoauereocins	20	1922, 1936, 1950, 1964 (MH ⁺)
<i>T. cf. strigosum</i>	BBA 69577= CBS 119777	SF1	Tricholongins	19	1908, 1911, 1925 (MH ⁺)
		SF4	Trikoningsins/ trichogin (12)§	11	1061, 1075, 1089, 1103 (MNa ⁺)
		SFx	Lipostrigocins A (6)	7	
		SF5	Lipostrigocins B (9)	11	
<i>T. strigosum</i>	CBS 348.93	SF4	Trikoningsin type, trichobrevin type	11	1075 (MNa ⁺)
		SF1	Trichostrigocins (2)	19	1891 (MH ⁺)
<i>T. erinaceus</i>	DAOM 230019= CBS 117088	SF1	Trichostrigocins (2)	19	1891 (MH ⁺)
		SF4	Trikoningsin typell, trichobrevin type	11	1075 (MNa ⁺)
<i>T. cf. pubescens</i>	BBA 68989= CBS 119776	SF1	Tricholongins (3)¶	19	1911, 1925 (M)
		SF5	Lipostrigocins B (3)	11	1075, 1089, 1103 (MNa ⁺)
<i>T. stromaticum</i>	BBA 70638=CBS 101875	SF1	Trichostromaticins (5)	18	1761, 1775, 1789, 1803, 1817 (MH ⁺)
		SFx	Trichocompactins (2)**	8	768, 770, 784 (MH ⁺)
<i>T. spirale</i>	CBS 346.93		Trichobrevin BIIIc**		(1127, 1113, 1143) 1157 (MNa ⁺)
<i>T. brevicompactum</i>	CBS 109720	SF1	Alamethicins F30 (3)	20	1962, 1963, 1976, 1977 (MH ⁺)
		SFx	Trichocompactins (6)	8	726, 740, 754, 756, 770, 784 (MH ⁺)
		SFx	Trichocryptins B (8)	12	1210, 1224, 1238, 1252, 1268, 1280 (MNa ⁺)
		SFx	Trichocryptins A (12)	11	1125, 1139, 1153, 1167 (MNa ⁺)
	CBS 119570=IBT 40840	SF1	Alamethicins F30 (3)	20	1962, 1963, 1976, 1977 (MH ⁺)
		SFx	Trichocompactins (6)	8	726, 740, 754, 756, 770, 784 (MH ⁺)
		SFx	Trichocryptins A (7)	11	1153, 1167 (MNa ⁺)
	CBS 119569=IBT 40839	SF1	Alamethicins F30 (3)	20	1962, 1963, 1976, 1977 (MH ⁺)
		SFx	Trichocompactins (5)	8	726, 740, 754, (MH ⁺)
		SFx	Trichocryptins A (7)	11	1153, 1167, 1181 (MNa ⁺)
	CBS 112445	SF1	Alamethicins F30 (5)	20	1950, 1962, 1963, 1964, 1976, 1977 (MH ⁺)
		SFx	Trichocompactins (6)	8	726, 740, 754, (MH ⁺)
		SFx	Trichocryptins A (7)	12	1139, 1153, 1167 (MNa ⁺)
		SFx	Trichocryptins B (6)	12	1224, 1238, 1252, 1266, 1280 (MNa ⁺)
<i>T. cf. brevicompactum</i>	CBS 119576=ATCC 90237	SFx	Trichocompactins (10)	8	756, 770, 784 (MH ⁺) – Gly-Ser exchange
		SFx	Trichobrevins A (11)	11	1099, 1113, 1127, 1141 (MNa ⁺)
		SFx	Trichobrevins B (8)	11	1129, 1143, 1157 (MNa ⁺)
	CBS 119577=IBT 40863	SF1	Alamethicins F30 (>4)	20	1950, 1962, 1963, 1964, 1976, 1977 (MH ⁺)
		SFx	Trichocryptins A	12	1139, 1153, 1167 (MNa ⁺)
		SFx	Trichocryptins B	12	1224, 1238, 1252, 1266, 1280 (MNa ⁺)
	NRRL 3199 (formerly <i>T. viride</i>)	SF1	Alamethicins F30 (4)	20	1950, 1964, 1976, 1977 (MH ⁺)

Table 3. cont.

Species	Strain	Subfamily*	Compound name†	No. of amino acid residues	Monoisotopic masses of subfamily members
		SFx	Trichoferins (7)	9	1193, 1207 (MH ⁺)
		SFx	Trichocompactins (6)	8	740, 754, 756, 770 (MH ⁺)
		SFx	Trichobrevins A (11)	11	1099, 1113, 1127, 1141 (MNa ⁺)
		SFx	Trichobrevins B (8)	11	1129, 1143, 1157 (MNa ⁺)

*SFx are subfamilies not defined by Chugh & Wallace (2001).

†Numbers in parentheses indicate the number of analogues described.

‡C. P. Kubicek (unpublished data).

§Trikoningin KBI also from *T. koningii* 903589, trichogin GA IV also from *T. longibrachiatum* MNHN 3431, and in *H. vinosa* CBS 247.63 and *H. lacteal* CBS 853.70.

||Presumably identical to trikonigoin KBII from *T. koningii* 903589.

¶Tricholongin has also been found in *T. longibrachiatum* MNHN 3431/*T. cf. strigosum* BBA 69577.

**Reported for *T. cf. brevicompactum* ATCC 90237=IBT 40863=NRRL 3199.

binding pocket (Rausch *et al.*, 2005) with the phylogenetic separation of early or central Gln domains, late Gln domains and late additional Gln domains (Table 5) enables refinement of the substrate-binding pocket consensus.

The Aib codes, as defined by Stachelhaus *et al.* (1999) and Challis *et al.* (2000), of early and late domains differ by a single Tyr→Phe replacement; however, in the extended binding site analysis 3 of the 34 amino acids are different. Extended binding site analysis of the Pro adenylation domain reveals the complete identity of all 34 residues. Phylogenetic and comparative sequence analysis thus permits not only a prediction of the substrate profile, but also the classification of the domain position within the synthetase.

Nucleotide sequence evolution of substrate-activating domains

To learn more about the evolution of the peptaibol synthases, we have investigated their interspecies nucleotide diversity, rate of nucleotide substitution and nucleotide sequence phylogeny. The latter provided the same phylogenetic tree structure as the amino acid sequences (data not shown). Applying Tajima's test implemented in the program DnaSP confirmed that they are undergoing neutral evolution ($D=2.65008$ and 1.874744 for the Aib/Pro- and Gln-activating domains, respectively), which coincides with a high K_S/K_N (differences in synonymous to non-synonymous substitutions) ratio (>5) which indicates the operation of purifying selection. Nucleotide diversity (p) was between 0.34 and 0.36, which correlates well with a similar value for nucleotide diversity in the exons of *tef1* and *rpb2* (0.33 and 0.32, respectively). In summary, the activating domains for Aib, Pro and Gln have evolved in a similar neutral way as housekeeping genes.

DISCUSSION

Here, we have applied a survey-type approach which permits the rapid analysis of a large number of samples to

define peptaibol production profiles. Compared to the recently described peptaibiomics approach (Degenkolb *et al.*, 2006a; Krause *et al.*, 2006), our approach only partially identifies the peptaibols, but the information can be obtained from extremely small amounts of biomass and much more rapidly. Below, we discuss the consistency of our data with the available structural data and the implications of molecular genetic data for the study of the evolution of these NRPS systems.

Peptaibol production data survey and correlation of groups defined in our survey with known compounds

We have compiled available data on peptaibols described for *Trichoderma/Hypocrea* (Table 3 and 4), including some mass data permitting comparison to Table 2. Unfortunately, as stressed in the Introduction, the species identity of several strains in the Birkbeck database is highly doubtful, and we present these data and additional data therefore in two separate tables; species identities quoted in Table 4 must be used with caution. Having defined 26 peptaibol groups by their mass profiles, we first need to ask if our data correlate with available data and if additional compounds are known from strains outside our selection.

In a recent study, Krause *et al.* (2006) investigated 11 defined strains of *Trichoderma/Hypocrea* species by solid phase extraction/liquid chromatography/electrospray ionization/mass spectrometry (SPE/LC-ES-MS) (*T. asperellum* CBS 433.97, *T. inhamatum* CBS 345.96=*H. lixii*, *T. aggressivum* f. *europaeum* CBS 100526, *T. stromaticum* CBS 101875, *H. semiorbis* CBS 244.63, *H. vinosa* CBS 247.63, *H. dichromospora* CBS 337.69, *H. gelatinosa* CBS 724.87, *H. nigricans* MUCL 28439=*H. lixii*, *H. muroiana* MUCL 28442 and *H. lactea* CBS 853.70). Molecular masses and partial sequences of 48 peptaibols have been compiled, only seven of which match data compiled in the Birkbeck database, the remaining being either new analogues or new types of compounds. Interestingly these 48 compounds can

Table 4. Peptaibol production in phylogenetically undefined production strains

Species	Strain	Subfamily*	Compound name	No. of amino acid residues	Monoisotopic masses of subfamily members
<i>T. harzianum</i>	M-903602	SF1	Trichorzin HA	18	1703, 1717, 1731, 1745
	M-922835	SF1	Trichorzin MA	18	1733, 1747, 1761
	M-902608	SF1	Trichorzin PA	18	1780, 1794, 1819, 1833, 1847
	PC01	SF1	Trichotoxins	18	1689, 1703, 1717
	Not deposited†	SF1	Trichokindins	18	1747, 1761, 1775
	M-903603	SF4	Harzianins	14	1386, 1400, 1402, 1414, 1416, 1428, 1430, 1444
	M-903603	SF4	Harzianin HB	11	1161, 1175
<i>T. koningii</i>	Not deposited‡	SF4	Trichorozins	11	1161, 1175, 1189
	Not deposited‡	SF1	Trichokonins	20	1865, 1908, 1922, 1936, 1950
	Not given§	SF1	Trikoningin	19	1989
<i>T. longibrachiatum</i>	LCP 853431	SF5	Trikoningins	11	1038, 1052
	MNHN3431	SF1	Longibrachin	20	1936, 1937, 1950, 1951, 1964
	M 3431	SF1	Tricholongins	19	1909, 1923
<i>H. peltata</i>		SF5	Trichogins	11	1066
<i>T. polysporum</i>		SF1	Hypelcins	20	1914, 1928, 1929, 1942, 1943
		SF1	Polysporins	20	1962, 1990, 2004
<i>T. viride</i> = <i>T. cf. harzianum</i>		SF1	Trichosporins	20	1908, 1922, 1936, 1950, 1952, 1964, 1966
		SF1	Suzukacillins	20	1922
<i>T. asperellum</i>	NRRL 5243	SF1	Trichovirins I	14	1383, 1397, 1401, 1411, 1415, 1425, 1429
	NRRL 5243	SF1	Trichovirins II	18	1703, 1717, 1731, 1745
		SF4	Trichorovins	11	1147, 1161, 1175, 1189
		SF1	Trichocellin	20	1922, 1923, 1936, 1937, 1950, 1964
<i>T. asperellum</i>	Not deposited	SF1	Trichosporins	20	1964
<i>T. saturnisporum</i>		SF5	Trichodecenins	7	754
<i>T. pseudokoningii</i>		SF1	Saturnisporins	20	1922, 1936, 1950
		SF4	Pseudokonins	10	1059
	MVHC 662	SF4	Harzianin HK VI	11	1197 (Na ⁺), 1213 (MK ⁺)
<i>Trichoderma</i> sp.		SF1	Peptaivirins	19	
<i>Trichoderma</i> sp.	HKI 0276	SF1	Trichovirins	19	
		SF4	Trichofumins	13	1445, 1459
		SF1	Trichofumins	11	1189, 1203
Unidentified	KGT142	SF1	Peptaivirins	19	1886, 1909
<i>T. virens</i>		SF1	Trichorzins	18	1705, 1719, 1733

*SFx are subfamilies not defined by Chugh & Wallace (2001).

†Isolated from soil, Nara, Japan (Iida *et al.*, 1994).

‡Isolated from wood material inoculated with *Ganoderma lucidum*.

§*T. koningii* isolate from Uruguay (M. F. Roquebert, Museum National d'Histoire Naturelle, Laboratoire de Cryptogamie, Paris, France).

||Provided by Yakult Central Institute for Microbiological Research (Fujita *et al.*, 1994).

be classified into six defined groups as shown in Table 1 and six new groups not yet defined. As only four of the above strains were contained in our pool, the results of Krause *et al.* (2006) can be used to amend our data and extend the number of mass-defined groups to 32 within a total of 39 strains.

Mass peaks may be composed of several peptaibols. In the case of *H. gelatinosa*, this study defined compounds with masses of 1176 and 1204 Da with partial sequences Vxx(=Val/Iva)-Lxx(=Leu/Ile)-Lxx-Aib-Vxx and Gln-Lxx-

Lxx-Aib, respectively. The Na⁺ adduct of the 1204 Da compound would correlate with *m/z* 1227 and the sequence data of group 6. The partial sequence and mass data of the 1176 Da compound are unknown. As the respective Na⁺ adduct would also fit into group 6, these data warn us that unknown compounds may be hidden in the known masses. In addition to these two 11-residue peptaibols, we also detected group 11 and group 19 peptaibols. In the case of *H. semiorbis*, Krause *et al.* (2006) described 20-residue compounds with masses of 1937, 1951 and 1965 Da, corresponding to group 25 compounds

Table 5. Comparison of substrate binding codes of conserved domains in peptaibol synthetases

Domain	10 aa code*	Extended code†
Early Aib domains		
TV1	DLGYLAGVFK	SAIAFDLSIGEGYLLFAGEINGWGPAETCVFSTL
PAR1	DLGYLAGVFK	SAIAFDLSIGEGYLLFAGEINGWGPAETCVFSTL
ATRO1	DLGYLAGVFK	SAIAFDLSIGEGYLLFAGEVNGWGPAETCVFSTL
Late Aib domains		
TV15	DLGFLAGVFK	SAIAFDLSIGEAFLLLAGEINGWGPAETCVFSTL
PAR16	DLGFLAGVFK	SAIAFDLSIGEAFLLLAGEVNGWGPAETCVFSTL
ATRO15	DLGFLAGVFK	SAIAFDLSIGEAFLLLAGEVNGWGPAETCVFSTL
Pro		
TV13	DVLFCLGLICK	SAYSFDVGVLVDVFLCLGGEHGLYGPAEASICAWN
PAR14	DVLFCLGLICK	SAYSFDVGVLVDVFLCLGGEHGLYGPAEASICAWN
ATRO13	DVLFCLGLICK	SAYSFDVGVLVDVFLCLGGEHGLYGPAEASICAWN
Early Gln		
TV6	DGGMVGGNYK	SGHSFDGSLGETMLVLGGEVNGYGPAEACNYATT
PAR7	DGGMVGGNYK	SGHSFDGSLGETMLVLGGEVNGYGPAEACNYATT
ATRO6	DGGMVGGNYK	SGHSFDGSLGETMLVLGGEVNGYGPAEACNYATT
Late Gln		
TV17	DGGMVGGNYK	TGYSFDGSLGEGMLVLGGEFNGYGPAEACNYATT
PAR18	DGGMVGGNYK	TGYSFDGSLGEGMLVLGGEYNGYGPAEACNYATT
PAR19	DGGMVGGNYK	SGYSFDGSLGEAMLVLGGEINGYGPAEACNYATW
ATRO17	DGGMVGGNYK	TAYSFDGSLGEGMLVLGGEFNGYGPAEACNYATI
ATRO18	DGGMIGGNYK	SGYSFDGSLGEAMLILGGEINGYGPAEACNYATW
SALP2_17	DGGMVGGNYK	TGYSFDGSLGEGMLVLGGEFNGYGPAEACNYATT

*The 10 aa code is defined by 10 positionally conserved putative binding pocket residues of the respective adenylate domains.

†The extended code defined by Rausch *et al.* (2005) extracts all amino acids within a distance of 8 Å of the predicted amino acid binding pocket of the respective adenylate domain.

as respective Na⁺ adducts with *m/z* 1960, 1974 and 1988. Surprisingly, we detected groups 6, 11, 13 and 19. Here, apparently, strain identity needs to be confirmed.

Strains of one species may produce different peptaibols.

Krause *et al.* (2006) found that *H. lixii* CBS 273.78 produced three sets of compounds with masses of 1146, 1675 and 1703/1717/1731 Da. The long peptaibols clearly match group 18, the trichovorin II type, while the others have no match, presumably representing new groups. This agrees with *H. lixii* CBS 226.95 which produces 18-residue trichotoxins, also members of this mass group. In our study, *H. lixii* CBS 273.78 produced compounds of groups 6, 14 and 19. These data indicate that strains genetically identified as *H. lixii* may still differ in their peptaibol production profiles. However, *H. lixii* is a complex of several cryptic species (Chaverri *et al.*, 2003) which are known to differ in their biochemical and physiological properties (C. P. Kubicek & I. S. Druzhinina, unpublished data).

Putative bona fide *H. lixii* strains (Table 4), trichorzins of the HA (group 18) and MA types (group 19) have been described from strains M-903602 (soil sample, Uruguay) and M-922835 (soil sample, Malaysia) (Goulard *et al.*, 1995; Hlimi *et al.*, 1995). A set of trichotoxins (group 18) with 11 of the 22 analogues already known from *H. lixii*

NRRL 5243 (Przybylski *et al.*, 1984; Brückner *et al.*, 1984, 1985) have been described from *Trichoderma* sp. strain PC01 (Suwan *et al.*, 2000). In addition, harzianins with 11 and 14 residues (groups 6 and 11) and trichorozins (group 6) have been described from other putative strains of *H. lixii*. These data are consistent with the conclusion that *H. lixii* produces either group 18 or 19 compounds, but neither group is restricted to this species. The transition from group 18 to 19 involves a substrate specificity shift of domain 2 from Gly to Ser. Such a shift might occur by point mutation or domain exchange. The unique group 22, corresponding to harzianins of the PA type with a terminal aromatic residue, described for the undefined strain M-902608, has also been detected in *T. hunua* and *T. strictipilosa*, both members of section Pachybasium B. These data may indicate that the occurrence of 18-residue peptaibols terminating in an aromatic amino alcohol may be restricted to this section.

A similar conclusion can be reached considering the recent study of Degenkolb *et al.* (2006b) comparing various strains of *T. brevicompactum* (Table 3). All four strains produced alamethicins, trichocryptins A and trichobrachsins, while only two strains produced trichocryptins B in addition. Comparing three strains of *T. cf. brevicompactum*, the same study describes trichocompactins and trichobrevins in CBS

119576, alamethicins and trichocryptins in CBS 119577, and alamethicins, trichoferins, trichocompactins and trichobrevins A and B in NRRL 3199.

In a recent peptaibomics study, plant-protective *Trichoderma* strains of different origins (*T. strigosum*, *T. pubescens* and *T. stromaticum*) were compared (Degenkolb *et al.*, 2006a). The authors concluded from the differences between the strains and similarities between different species that peptaibol sequences are of limited relevance for chemotaxonomy. Thus *T. strigosum* CBS 119777, a German compost isolate, was found to produce tricholongins B, 19-residue peptaibols of type VI/SF1, while *T. strigosum* CBS 348.93, a forest soil isolate from North Carolina, produced trichostrogocins, 19-residue type VII/SF1 compounds. There are three significant module changes in these types, exchanging Gly with Ala, Phe with Aib and Aib with Lxx (Fig. 6). Interestingly, a tricholongin has also been detected in *T. cf. pubescens* BBA 66989. The 11-residue lipostrogocins were detected in all three strains. Both species belong to the section *Trichoderma*. Other type VII/SF1 compounds, the 19-residue trichorzianins and trikoningins, are produced by *H. atroviridis* ATCC 36042 and *T. koningii*, respectively, also members of this section.

Another example of significant alteration of a structure is the production of lipostrogocins B and linopubescins in *T. cf. pubescens* BBA 66989 and *T. pubescens* CBS 345.93, respectively (Degenkolb *et al.*, 2006a). Interconversion of the 11- and 10-residue compounds would require alteration (Gly/Vxx) and deletion/insertion of one module, probably involving extensive alterations of the respective NRPS gene.

From the limited data available, it thus seems possible that certain structural types are restricted to sections of the *Trichoderma/Hypocrea* group. It will be interesting to investigate whether the occurrence of similar compounds in different species is due to coevolution or to horizontal gene transfer.

Failure to detect certain compounds. The unique trichostromaticin type of 18-residue peptaibols (group 21, type III of SF1) has been found in *T. stromaticum* BBA 70636/70638 (Degenkolb *et al.*, 2006a). In this case we failed to detect peptaibols, although similar growth conditions were used. In *T. stromaticum* CBS 101875, in addition to group 21 trichostromaticocins C, group 13 peptaibols have been described (Krause *et al.*, 2006). The unknown 1538 Da compound detected might be identical to compounds we detected in *H. atroviridis* ATCC 74058, *H. citrina*, *T. hamatum*, *T. longipile*, *T. pubescens*, *H. semiorbis* and *H. strictipilosa*.

In a putative strain of *T. pubescens* (BBA 66989=DAOM 162162), Degenkolb *et al.* (2006b) found a 19-residue compound of the tricholongin type and group 1 compounds of the lipostrogocin type. We detected only group 1 and 13 mass peaks in this strain, and failed to find the long peptaibol. *T. pubescens* CBS 119776, however, was found to

produce a 20-residue peptaibol of the group 25 type. Thus, as in *H. lixii*, strains of *T. pubescens* produce different groups of 19- and 20-residue peptaibols.

Failures in compound detection can be related to either strain-specific production profiles of further isolates, whose species identity has been verified by sequence analysis, or loss of production by gene loss or expression failure. Both cases have been well documented in cyanobacteria, where numerous isolates from single biotopes showed a large variety of chemotypes. As in the case of *Trichoderma/Hypocrea*, the compounds derived from NRPS genes in cyanobacteria have been used for chemotype definition (Welker *et al.*, 2004a, b, 2006).

Possible identities of undefined strains. Mass group 24 corresponds to the 19-residue trikoningin type, and trikoningin KA has been found in a putative *T. koningii* isolate from Uruguay (Auvin-Guette *et al.*, 1993). We detected masses corresponding to trikoningin in *T. koningii* CBS 460.96. We did not see, however, the 20-residue trichokonins, belonging to mass group 25, which have been described from another putative *T. koningii* isolate (Huang *et al.* 1995a, b, 1996; Landreau, 2001; Landreau *et al.*, 2002; Table 3). *T. koningii* has frequently been confused with *Trichoderma* species from section *Longibrachiatum* (Lieckfeldt *et al.*, 1999).

Trichorzianins (SF1 type VII), which are 19-residue peptaibols with 13 analogues in the biocontrol strain *T. atroviride* P1 (described as *H. lixii* in this study; Pocsfalvi *et al.*, 1998), of which eight have already been described (Bodo *et al.*, 1985; Rebuffat *et al.*, 1989), and the 19-residue peptaivirins (SF1, type V) from an unidentified *Trichoderma* species (Yun *et al.*, 2000) have not been included in this study, but would need to be described in two additional mass groups (Table 4). In this study we failed to detect these compounds in strain P1. Instead, we found 20-residue peptaibols of group 26 corresponding to the alamethicin type, including atroviridins, well known from strains *H. atroviridis* P1 and ATCC 36042 (Oh *et al.*, 2000). Unless the strain has been confused in the earlier study, these differences must be due to different culture conditions.

The structures of atroviridins and trichorzianins deviate in various positions, so that a direct transition by gain or deletion of one module is not possible. Interestingly, strain *H. atroviridis* IFO 31288, formerly described as *H. muroiana*, produces the 18-residue hypomurocins (Becker *et al.*, 1997), members of group 19, which cannot be distinguished from other members of this group, trichokindins and trichorzins type MA, by molecular ions alone. Thus all three characterized strains of *H. atroviridis* produce different types of SF1 peptaibols, 18-residue hypomurocins (type II, group 19), 19-residue trichorzianins (type VII, no group assigned) and 20-residue atroviridins (type IX, group 26). Other group 26 producers are strains of *T. brevicompactum*, which have been

extensively studied by Degenkolb *et al.* (2006b) and shown to produce alamethicins of the F-30 type. Contrary to *H. atroviridis*, section *Trichoderma*, these belong to section *Pachybasium* B. There are only minor differences between atroviridins and alamethicins: the specificity of position 17 is shifted from Aib to Iva/Aib, and in the case of alamethicin, Glu-18 may replace Gln-18.

Twenty-residue compounds of the paracelsin type (group 25, type VIII) are the only long peptaibols found in section *Longibrachiatum*. We have also found masses corresponding to group 25 in section *Pachybasium* strains *H. pachybasioides* DAOM 167068 (formerly *T. croceum*) and *H. strictipilosa* CBS 347.93. Group 25 compounds have also been described for two other putative strains of *H. pachybasioides*, i.e. trichosporins (Iida *et al.*, 1990, 1993, 1999; Sharman *et al.*, 1996) and polysporins (type VIII) (New *et al.*, 1996; Fujita *et al.*, 1981).

These data imply that group 25 peptaibols are the only long types found in section *Longibrachiatum*, while strains in section *Pachybasium* produce a variety of long types, including this group. Group 26 on the other hand has so far only been found in sections *Trichoderma* and occasionally in *Pachybasium* B.

More than one peptaibol produced from a single synthetase? The case of SF4 11- and 14-residue compounds. Both 11-residue and 14-residue peptaibols have been assigned to SF4 as their structural similarity implies some relationship which has not yet been defined. Group 6 compounds with the structure AcX(N,Q)XX-UPXXUPXOH (Fig. 3) and groups 8, 9 and 11 with the general structure AcX(N,Q)LUPXXUPXLUPXOH have identical terminal regions (underlined) and share six of the central residues (XXUPXX). Biosynthetically, a deletion of modules 3–5 (LUP) of a putative 14-module NRPS would lead to group 6 compounds. Two 11-residue harzianins, HB I and HK VI, have been isolated from *H. lixii* M-903603 (Augeven-Bour *et al.*, 1997) and *T. pseudokoningii* MVHC 662 (Rebuffat *et al.*, 1996), respectively. While HB I represents 14-residue SF4 peptaibols of the same strain with an internal deletion of 3 residues, HK is a member of a prominent family of 11-residue peptaibols which have not been structurally characterized. Both structures could be derived from the same deletion. As both 11- and 14-residue compounds have been detected in one strain, M-903603, a single NRPS might be suspected. It is not known, however, if the *T. pseudokoningii* strain also produces 14-residue compounds.

We have detected the formation of both 11- and 14-residue compounds of groups 6 and 9, 10 or 11 in several strains of section *Pachybasium* B (Table 2). However, group 6 compounds are widely distributed in other sections as well. Indeed, a more detailed analysis of *H. jecorina* (section *Longibrachiatum*) showed the presence of a 14-module NRPS, although some strains were found to produce only 11-residue group 6 compounds, while other strains

produced both 11- and 14-residue peptaibols (H. von Döhren & others, unpublished). It remains to be shown if these NRPSs show section- or strain-specific alterations.

Phylogenetic analysis of peptaibol synthetases hints at a complex history. A comparison of sequence data for conserved amino-acid-activating domains of long SF1 peptaibols reveals a similar phylogenetic distance in the peptaibol synthetase genes to that in other housekeeping genes, indicating a comparable rate of evolution. Genetic data from NRPS systems within the *Trichoderma/Hypocrea* group may thus reflect, at least within the conserved regions, the distances obtained from the comparison of other genes.

These findings only partially agree with recent work on cyanobacterial NRPS systems. The characteristic conserved regions of the NRPS genes of *Cyanobacterium* species synthesizing cyclic 7-residue microcystins, and which are produced by different sections of the cyanobacteria, have been compared using sequences from both 16S rRNA and a fragment of the RNA polymerase I gene (Rantala *et al.*, 2004). Phylogenetic analysis indicates concordant evolution of these housekeeping genes with the NRPS genes for their entire evolutionary history, and does not support horizontal transfer between the genera. Our study indicates that a similar form of evolution has also taken place in the peptaibol synthetase genes.

A more complex view of peptaibol chemotypes. The sporadic distribution of bacterial peptide and toxin gene clusters has been ascribed to repeated losses during diversification processes. In the case of peptaibols, we also see a patchwork-like distribution of products, and even strains of the same (confirmed) species may display variable product patterns. Such distributions have been analysed by employing MALDI-TOF MS typing of single filaments or microcolonies of cyanobacteria characterizing single clones within defined biotopes (Fastner *et al.*, 2001; Welker *et al.*, 2002, 2004a, b, 2006; Welker & von Döhren, 2006). In one example, 18 clones isolated from single filaments of *Planktothrix* from a single lake represented 15 chemotypes with respect to at least 11 different non-ribosomal peptides, including 25 analogues. The number of peptides produced has been found to range from one to four, indicating a limited coding space within the bacterial genome. Likewise, our data indicate a variety of chemotypes of individual species, but the number of peptaibols produced seems to be restricted to at most four types.

In bacteria, these observations have been interpreted as an exchange of NRPS gene clusters between strains of the same species, since transposase-related sequences have frequently been found at cluster boundaries. Comparing the genomes of *Bacillus subtilis* 168 and the closely related *Bacillus amyloliquefaciens* FZB 42, three and six NRPS/PKS clusters were found, respectively (Koumoutsis *et al.*, 2004). Two NRPS clusters and operons were positionally conserved with respect to *B. subtilis*, and several transposase sequences are contained in it. As for fungi, transposon

relics have so far only been found in the ergovaline biosynthetic clusters of the endophyte *Epichloe* (Fleetwood *et al.*, 2007). Interestingly in *Trichoderma*, similar genes are found 5' and 3' of the atroviridine synthetase of *H. atroviridis* and of a putative harzianine synthetase in *H. jecorina* which encode SF1 and SF4 peptaibol types, respectively (H. von Döhren, unpublished data). These findings indicate that transfers and exchanges of peptaibol synthetase genes may also contribute to the chemodiversity of peptaibols in the *Trichoderma/Hypocrea* group, and account for the diverse chemotypes of single species. However, the diversity of the peptaibols formed also appears to be related to certain pools of compounds restricted to sections or clades. It is possible that the formation of certain chemotypes could be traced back to gene transfer events between closely related species. We must admit, however, that this has been shown here only for a small number of genes, and exceptions caused by horizontal recombination events or recombination events between synthetases of different subfamilies cannot be fully ruled out at this stage.

Conclusions

Many peptaibiotics have been named after the *Trichoderma* species producing them. Unfortunately, the identity of most of these *Trichoderma* strains must be treated with extreme caution, due to the difficulties of identifying these species by morphological means only (cf. Druzhinina *et al.*, 2006). Krause *et al.* (2006), investigating five defined, different species of *Trichoderma*, recently showed that even closely related species can produce significantly different peptaibol patterns. Two studies comparing in detail two or more different strains from various species even demonstrated strain-specific production of certain peptaibols (Degenkolb *et al.*, 2006a, b). The present study, which made use of a set of 28 phylogenetically characterized species, basically supports these findings. However, a more detailed inspection shows that some generalizations are still possible, considering in more detail the structural types of peptaibols and conserved structures of peptaibol synthetases. Some types of compounds defined by mass groups or sequence-based types seem to be restricted to single sections or clades. The validity of these predictions can be documented by the findings that a strain of *H. peltata* has been reported to produce group 25 peptaibols, which we interpret to be typical for section Longibrachiatum. Although an *rpb2* sequence for *H. peltata* was unavailable for this study, ITS1 and 2 sequence analysis places it in a sister clade close to section Longibrachiatum (Dodd *et al.*, 2002; Druzhinina *et al.*, 2005) which is consistent with this prediction. In addition sequence data of the respective peptaibol synthetase genes may provide information on their phylogenetic background.

ACKNOWLEDGEMENTS

This work was supported by a fellowship from the Deutsche Forschungsgemeinschaft (Do270/10), by the European Community

(EUROFUNG) and by a grant from the Austrian Science Foundation to C. P. K. (P19960-B16).

REFERENCES

- Augeven-Bour, I., Rebuffat, S., Auvin, C., Goulard, C., Prigent, Y. & Bodo, B. (1997). Harzianin HB I, an 11-residue peptaibol from *Trichoderma harzianum*: isolation, sequence, solution synthesis and membrane activity. *J Chem Soc Perkin Trans I* 10, 1587–1594.
- Auvin-Guette, C., Rebuffat, S., Prigent, Y. & Bodo, B. (1992). Trichogin A IV, an 11-residue lipopeptaibol from *Trichoderma longibrachiatum*. *J Am Chem Soc* 114, 2170–2174.
- Auvin-Guette, C., Rebuffat, S., Vuidepot, I., Massias, M. & Bodo, B. (1993). Structural elucidation of trikoningins KA and KB, peptaibols from *Trichoderma koningii*. *J Chem Soc Perkin Trans I* 2, 249–255.
- Becker, D., Kiess, M. & Brückner, H. (1997). Structures of peptaibol antibiotics hypomurocin A and B from the ascomycetous fungus *Hypocrea muroiana* Hino et Katsumoto. *Liebigs Annalen-Recueil*, 767–772.
- Benedetti, E., Bavoso, A., Di Blasio, B., Pavone, V., Pedone, C., Toniolo, C. & Bonora, G. M. (1982). Peptaibol antibiotics: a study on the helical structure of the two sequences of emerimicins III and IV. *Proc Natl Acad Sci U S A* 79, 7951–7954.
- Bissett, J. (1984). A revision of the genus *Trichoderma*. I. Section Longibrachiatum sect. nov. *Can J Bot* 62, 924–931.
- Bissett, J. (1991a). A revision of the genus *Trichoderma*. II. Infrageneric classification. *Can J Bot* 69, 2357–2372.
- Bissett, J. (1991b). A revision of the genus *Trichoderma*. III. Section Pachybasium. *Can J Bot* 69, 2373–2417.
- Bissett, J. (1991c). A revision of the genus *Trichoderma*. IV. Additional notes on section Longibrachiatum. *Can J Bot* 69, 2418–2420.
- Bissett, J. (1992). *Trichoderma atroviride*. *Can J Bot* 70, 639–641.
- Bodo, B., Rebuffat, S., El Hajji, M. & Davoust, D. (1985). Structure of trichorzianine A IIIc, an antifungal peptide from *Trichoderma harzianum*. *J Am Chem Soc* 107, 6017–6019.
- Brückner, H. & Graf, A. (1983). Paracelsin, a peptide antibiotic containing alpha-aminoisobutyric acid, isolated from *Trichoderma reesei* Simmons. Part A. *Experientia* 39, 528–530.
- Brückner, H. & Koza, A. (2003). Solution phase synthesis of the 14-residue peptaibol antibiotic trichovirin I. *Amino Acids* 24, 311–323.
- Brückner, H., Graf, H. & Bokel, M. (1984). Paracelsin; characterization by NMR spectroscopy and circular dichroism, and hemolytic properties of a peptaibol antibiotic from the cellulolytically active mold *Trichoderma reesei* Simmons. Part B. *Experientia* 40, 1189–1197.
- Brückner, H., König, W. A., Aydin, M. & Jung, G. (1985). Trichotoxin A40. Purification by counter-current distribution and sequencing of isolated fragments. *Biochim Biophys Acta* 827, 51–62.
- Challis, G. L., Ravel, J. & Townsend, C. A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem Biol* 7, 211–224.
- Chaverri, P., Castlebury, L. A., Samuels, G. J. & Geiser, D. M. (2003). Multilocus phylogenetic structure within the *Trichoderma harzianum*/ *Hypocrea lixii* complex. *Mol Phylogenet Evol* 27, 302–313.
- Chugh, J. K. & Wallace, B. A. (2001). Peptaibols: models for ion channels. *Biochem Soc Trans* 29, 565–570.
- Chugh, J. K., Brückner, H. & Wallace, B. A. (2002). Model for a helical bundle channel based on the high-resolution crystal structure of trichotoxin A50E. *Biochemistry* 41, 12934–12941.

- Degenkolb, T., Berg, A., Gams, W., Schlegel, B. & Gräfe, U. (2003). The occurrence of peptaibols and structurally related peptaibiotics in fungi and their mass spectrometric identification via diagnostic fragment ions. *J Pept Sci* **9**, 666–678.
- Degenkolb, T., Gräfenhan, T., Berg, A., Nirenberg, H. I., Gams, W. & Brückner, H. (2006a). Peptaibiotics: screening for polypeptide antibiotics (peptaibiotics) from plant-protective *Trichoderma* species. *Chem Biodivers* **3**, 593–610.
- Degenkolb, T., Gräfenhan, A., Nirenberg, H. I., Gams, W. & Brückner, H. (2006b). *Trichoderma brevicompactum* complex: rich source of novel and recurrent plant-protective polypeptide antibiotics (peptaibiotics). *J Agric Food Chem* **54**, 7047–7061.
- Dodd, L. S., Lieckfeldt, E., Chaverri, P., Overton, B. E. & Samuels, G. J. (2002). Taxonomy and phylogenetic relationships of two species of *Hypocrea* with *Trichoderma* anamorphs. *Mycol Prog* **1**, 409–428.
- Druzhinina, I. & Kubicek, C. P. (2005). Species concepts and biodiversity in *Trichoderma* and *Hypocrea*: from aggregate species to species clusters? *J Zhejiang Univ Sci B* **6**, 100–112.
- Druzhinina, I. S., Kopchinskiy, A. G., Komon, M., Bissett, J., Szakacs, G. & Kubicek, C. P. (2005). An oligonucleotide barcode for species identification in *Trichoderma* and *Hypocrea*. *Fungal Genet Biol* **42**, 813–828.
- Druzhinina, I. S., Koptchinskiy, A. & Kubicek, C. P. (2006). The first one hundred *Trichoderma* species characterized by molecular data. *Mycoscience* **47**, 55–64.
- Duval, D., Rebuffat, S., Goulard, C., Prigent, Y., Becchi, M. & Bodo, B. (1997). Isolation and sequence analysis of peptide antibiotics trichorzins PA from *Trichoderma harzianum*. *J Chem Soc Perkin Trans I* **14**, 2147–2154.
- Erhard, M., von Döhren, H. & Jungblut, P. (1997). Rapid typing and elucidation of new secondary metabolites of intact cyanobacteria using MALDI-TOF mass spectrometry. *Nat Biotechnol* **15**, 906–909.
- Fastner, J., Erhard, M. & von Döhren, H. (2001). Oligopeptide diversity in natural populations of *Microcystis* (Cyanobacteria) by typing single colonies using MALDI-TOF mass spectrometry. *Appl Environ Microbiol* **67**, 5069–5076.
- Fleetwood, D. J., Scott, B., Lane, D. A., Tanaka, A. & Johnson, D. A. (2007). A complex ergovaline gene cluster in epichloe endophytes of grasses. *Appl Environ Microbiol* **73**, 2571–2579.
- Fujita, T., Takaishi, Y. & Okamura, A. (1981). New peptide antibiotics, trichopolyns I and II, from *Trichoderma polysporum*. *J Chem Soc Chem Commun* **12**, 585–587.
- Fujita, T., Wada, S., Iida, A., Nishimura, T., Kanai, M. & Toyama, N. (1994). Fungal metabolites. XIII. Isolation and structural elucidation of new peptaibols, trichodecenins-I and -II, from *Trichoderma viride*. *Chem Pharm Bull (Tokyo)* **42**, 489–494.
- Goulard, C., Hlimi, S. & Rebuffat, S. (1995). Trichorzins HA and MA, antibiotic peptides from *Trichoderma harzianum*. I. Fermentation, isolation and biological properties. *J Antibiot (Tokyo)* **48**, 1248–1253.
- Hermosa, M. R., Keck, E. J., Chamorro, I., Rubio, M. B., Sanz, L., Vizcaino, J. A., Grondona, I. & Monte, E. (2004). Genetic diversity shown in *Trichoderma* biocontrol isolates. *Mycol Res* **108**, 897–906.
- Hlimi, S., Rebuffat, S., Goulard, C., Duchamp, S. & Bodo, B. (1995). Trichorzins HA and MA, antibiotic peptides from *Trichoderma harzianum*. II. Sequence determination. *J Antibiot* **48**, 1254–1261.
- Huang, Q., Tezuka, Y., Hatanaka, Y., Kikuchi, T., Nishi, A. & Tubaki, K. (1995a). Studies on metabolites of mycoparasitic fungi. IV. Minor peptaibols of *Trichoderma koningii*. *Chem Pharm Bull (Tokyo)* **43**, 1663–1668.
- Huang, Q., Tezuka, Y., Kikuchi, T., Nishi, A., Tubaki, K. & Tanaka, K. (1995b). Studies on metabolites of mycoparasitic fungi. II. Metabolites of *Trichoderma koningii*. *Chem Pharm Bull (Tokyo)* **43**, 223–229.
- Huang, Q., Tezuka, Y., Hatanaka, Y., Kikuchi, T., Nishi, A. & Tubaki, K. (1996). Studies on the metabolites of mycoparasitic fungi. V. Ion-spray ionization mass spectrometric analysis of Trichokonin-II, a peptaibol mixture obtained from the culture broth of *Trichoderma koningii*. *Chem Pharm Bull (Tokyo)* **44**, 590–593.
- Iida, A., Okuda, M., Uesato, S., Takaishi, Y., Shingu, T., Morita, M. & Fujita, T. (1990). Fungal metabolites. Part 3. Structural elucidation and antibiotic peptides, trichosporin-B-IIIb, -IIIc, -IVb, -IVc, -IVd, -VIa and -VIb from *Trichoderma polysporum*. Application of fast atomic bombardment mass spectrometry to peptides containing a unique Aib-Pro peptide bond. *J Chem Soc Perkin Trans I* **12**, 3249–3255.
- Iida, J., Iida, A., Takahashi, Y., Takaishi, Y., Nagaoka, Y. & Fujita, T. (1993). Fungal metabolites. Part 5. Rapid structure elucidation of antibiotic peptides, minor components of trichosporin Bs from *Trichoderma polysporum*. Application of linked scan and continuous-flow fast atom bombardment mass spectrometry. *J Chem Soc Perkin Trans I* **3**, 357–365.
- Iida, A., Sanekata, M., Fujita, T., Tanaka, H., Enoki, A., Fuse, G., Kanai, M., Rudewicz, P. J. & Tachikawa, E. (1994). Fungal metabolites. XVI. Structures of new peptaibols, trichokindins I–VII, from the fungus *Trichoderma harzianum*. *Chem Pharm Bull (Tokyo)* **42**, 1070–1075.
- Iida, A., Mihara, T., Fujita, T. & Takaishi, Y. (1999). Peptidic immunosuppressants from the fungus *Trichoderma polysporum*. *Bioorg Med Chem Lett* **9**, 3393–3396.
- Jaworski, A., Kirschbaum, J. & Brückner, H. (1999). Structures of trichovirins II, peptaibol antibiotics from the mold *Trichoderma viride* NRRL 5243. *J Pept Sci* **5**, 341–351.
- Koumoutsis, A., Chen, X.-H., Henne, A., Liesegang, H., Hitzeroth, G., Franke, G., Vater, J. & Boriss, P. (2004). Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. *J Bacteriol* **186**, 1084–1096.
- Kraus, G. F., Druzhinina, I., Gams, W., Bissett, J., Zafari, D., Szakacs, G., Koptchinski, A., Prillinger, H., Zare, R. & Kubicek, C. P. (2004). *Trichoderma brevicompactum* sp. nov. *Mycologia* **96**, 1059–1073.
- Krause, C. (2006). *Isolierung, Sequenzierung und Wirkungsprüfung von bioaktiven Peptid-Antibiotika aus Schimmelpilzen*. PhD thesis, Justus-Liebig-Universität Gießen, Germany.
- Krause, C., Kirschbaum, J. & Brückner, H. (2006). Peptaibiotics: an advanced, rapid and selective analysis of peptaibiotics/peptaibols by SPE/LC-ES-MS. *Amino Acids* **30**, 435–443.
- Kuhls, K., Lieckfeldt, E., Samuels, G. J., Kovacs, W., Meyer, W., Petrini, O., Gams, W., Börner, T. & Kubicek, C. P. (1996). Molecular evidence that the asexual industrial fungus *Trichoderma reesei* is a clonal derivative of the ascomycete *Hypocrea jecorina*. *Proc Natl Acad Sci U S A* **93**, 7755–7760.
- Kullnig, C., Krupica, T., Woo, S. L., Mach, R. L., Rey, M., Benitez, T., Lorito, M. & Kubicek, C. P. (2001). Confusion abounds over identities of *Trichoderma* biocontrol isolates. *Mycol Res* **105**, 769–772.
- Kullnig-Gradinger, C. M., Szakacs, G. & Kubicek, C. P. (2002). Phylogeny and evolution of the fungal genus *Trichoderma*: a multigene approach. *Mycol Res* **106**, 757–767.
- Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.
- Landreau, A. (2001). *Métabolites d'une souche de Trichoderma koningii Oudemans isolée du milieu maris: étude chimique, biologique et risques pour les coquillages en culture*. PhD thesis, University Nantes, France.

- Landreau, A., Pouchus, Y. F., Sallenave-Namont, C., Biard, J. F., Boumard, M. C., Robiou du Pont, T., Mondegue, F., Goulard, C. & Verbist, J. F. (2002). Combined use of LC/MS and a biological test for rapid identification of marine mycotoxins produced by *Trichoderma koningii*. *J Microbiol Methods* **48**, 181–194.
- Leclerc, G., Rebuffat, S., Goulard, C. & Bodo, B. (1998a). Directed biosynthesis of peptaibol antibiotics in two *Trichoderma* strains. I. Fermentation and isolation. *J Antibiot (Tokyo)* **51**, 170–177.
- Leclerc, G., Rebuffat, S. & Bodo, B. (1998b). Directed biosynthesis of peptaibol antibiotics in two *Trichoderma* strains. II. Structure elucidation. *J Antibiot (Tokyo)* **51**, 178–183.
- Leclerc, G., Goulard, C., Prigent, Y., Bodo, B., Wroblewski, H. & Rebuffat, S. (2001). Sequences and antimycoplasmic properties of longibrachins LGB II and LGB III, two novel 20-residue peptaibols from *Trichoderma longibrachiatum*. *J Nat Prod* **64**, 164–170.
- Lieckfeldt, E., Samuels, G. J., Nirenberg, H. I. & Petrini, O. (1999). A morphological and molecular perspective of *Trichoderma viride*: is it one or two species? *Appl Environ Microbiol* **65**, 2418–2428.
- Lu, B. S., Druzhinina, I., Fallah, P., Chaverri, P., Gradinger, C. M., Kubicek, C. P. & Samuels, G. J. (2004). *Hypocrea/Trichoderma* species with pachybasium-like conidiophores: teleomorphs for *T. minutisporum* and *T. polysporum* and their newly discovered relatives. *Mycologia* **96**, 310–342.
- Nei, M. & Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. New York: Oxford University Press.
- Neuhoﬀ, T., Dieckmann, R., Drushinina, I. S., Kubicek, C. P., Nakari-Setälä, T., Penttilä, M. & von Döhren, H. (2007a). Direct identification of hydrophobins and their processing in *Trichoderma* using Intact-Cell MALDI-TOF mass spectrometry. *FEBS J* **274**, 841–852.
- Neuhoﬀ, T., Berg, A., Besl, H., Schweske, T., Dieckmann, R. & von Döhren, H. (2007b). Peptaibol production by *Sepedonium* strains parasitizing Boletales. *Chem Biodivers* **4**, 1103–1115.
- New, A. P., Eckers, C., Haskins, N. J., Neville, W. A., Elson, S., Hueso-Rodríguez, J. A. & Rivera-Sagredo, A. (1996). Structures of polysporins A-D, four new peptaibols isolated from *Trichoderma polysporum*. *Tetrahedron Lett* **37**, 3039–3042.
- Nicholas, H. B., Jr & McClain, W. H. (1987). An algorithm for discriminating sequences and its application to yeast transfer RNA. *Comput Appl Biosci* **3**, 177–181.
- Oh, S. U., Lee, S. J., Kim, J. H. & Yoo, I. D. (2000). Structural elucidation of new antibiotic peptides, atroviridins A, B and C from *Trichoderma atroviride*. *Tetrahedron Lett* **41**, 61–64.
- Pabel, C. T., Vater, J., Wilde, C., Franke, P., Hofemeister, J., Adler, B., Bringmann, G., Hacker, J. & Hentschel, U. (2003). Antimicrobial activities and matrix-assisted laser desorption/ionization mass spectrometry of *Bacillus* isolates from the marine sponge *Aplysina aerophoba*. *Mar Biotechnol (NY)* **5**, 424–434.
- Pocsfalvi, G., Ritieni, A., Ferranti, P., Randazzo, G., Vekey, K. & Malorni, A. (1997). Microheterogeneity characterization of a paracelsin mixture from *Trichoderma reesei* using high-energy collision-induced dissociation tandem mass spectrometry. *Rapid Commun Mass Spectrom* **11**, 922–930.
- Pocsfalvi, G., Scala, F., Lorito, M., Ritieni, A., Randazzo, G., Ferranti, P., Vekey, K. & Malorni, A. (1998). Microheterogeneity characterization of a trichorzianine-A mixture from *Trichoderma harzianum*. *J Mass Spectrom* **33**, 154–163.
- Przybylski, M., Dietrich, I., Manz, I. & Brückner, H. (1984). Elucidation of structure microheterogeneity of the polypeptide antibiotics paracelsin and trichotoxin A-50 by fast atom bombardment mass spectrometry in combination with selective in situ hydrolysis. *Biomed Mass Spectrom* **11**, 569–582.
- Rantala, A., Fewer, D., Hisbergues, M., Rouhiainen, L., Vaitomaa, J., Börner, T. & Sivonen, K. (2004). Phylogenetic evidence for the early evolution of microcystin synthesis. *Proc Natl Acad Sci U S A* **101**, 568–573.
- Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W. & Huson, D. H. (2005). Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res* **33**, 5799–5808.
- Rebuffat, S., El Hajji, M., Hennig, P., Davoust, D. & Bodo, B. (1989). Isolation, sequence, and conformation of seven trichorzianines B from *Trichoderma harzianum*. *Int J Pept Protein Res* **34**, 200–210.
- Rebuffat, S., Prigent, Y., Auvin-Guette, C. & Bodo, B. (1991). Tricholongins BI and BII 19-residue peptaibols from *Trichoderma longibrachiatum*. *Eur J Biochem* **201**, 661–674.
- Rebuffat, S., Goulard, C. & Bodo, B. (1995). Antibiotic peptides from *Trichoderma harzianum*: harzianins HC, proline-rich 14-residue peptaibols. *J Chem Soc Perkin Trans I* **14**, 1849–1855.
- Rebuffat, S., Hlimi, S., Prigent, Y., Goulard, C. & Bodo, B. (1996). Isolation and structural elucidation of the 11-residue peptaibol antibiotic, harzianin HK VI. *J Chem Soc Perkin Trans I* **16**, 2021–2027.
- Rifai, M. A. (1969). A revision of the genus *Trichoderma*. *Mycol Pap* **116**, 1–56.
- Rozas, J., Sanchez-DelBarrio, J. C., Messegue, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497.
- Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Sharman, G. J., Try, A. C., Williams, D. H., Ainsworth, A. M., Beneyto, R., Gibson, T. M., McNicholas, C., Renno, D. V., Robinson, N. & other authors (1996). Structural elucidation of XR586, a peptaibol-like antibiotic from *Acremonium persicinum*. *Biochem J* **320**, 723–728.
- Stachelhaus, T., Mootz, H. D. & Marahiel, M. A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* **6**, 493–505.
- Suwan, S., Isobe, M., Kanokmedhakul, S., Lourit, N., Kanokmedhakul, K., Soyong, K. & Koga, K. (2000). Elucidation of high micro-heterogeneity of an acidic-neutral trichotoxin mixture by electrospray ionisation quadrupole time-of-flight mass spectrometry. *J Mass Spectrom* **35**, 1438–1451.
- Szekeres, A., Leitgeb, B., Kredics, L., Antal, Z., Hatvani, L., Manczinger, L. & Vágvölgyi, C. (2005). Peptaibols and related peptaibiotics of *Trichoderma*. *Acta Microbiol Immunol Hung* **52**, 137–168.
- Tajima, F. & Nei, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol* **1**, 269–285.
- Vizcaino, J. A., Sanz, L., Cardoza, R. E., Monte, E. & Gutiérrez, S. (2005). Detection of putative peptide synthetase genes in *Trichoderma* species: application of this method to cloning of a gene from *T. harzianum* CECT 2413. *FEMS Microbiol Lett* **244**, 139–148.
- Vizcaino, J. A., Cardoza, R. E., Dubost, L., Bodo, B., Gutiérrez, S. & Monte, E. (2006). Detection of peptaibols and cloning of a putative peptaibol synthetase gene from *Trichoderma harzianum* CECT 2413. *Folia Microbiol (Praha)* **51**, 114–120.
- Welker, M. & von Döhren, H. (2006). Cyanobacterial peptides – nature's own combinatorial biosynthesis. *FEMS Microbiol Lett* **30**, 530–563.
- Welker, M., Fastner, J., Erhard, M. & von Döhren, H. (2002). Applications of MALDI-TOF MS analyses in cyanotoxin research. *Environ Toxicol* **17**, 367–374.
- Welker, M., Brunke, M., Preussel, K., Lippert, I. & von Döhren, H. (2004a). Diversity and distribution of *Microcystis* (Cyanobacteria)

oligopeptide chemotypes from natural communities studied by single-colony mass spectrometry. *Microbiology* **150**, 1785–1796.

Welker, M., Christiansen, G. & Döhren, H. (2004b). Diversity of coexisting *Planktothrix* chemotypes – production of microcystins and other oligopeptides. *Arch Microbiol* **182**, 288–298.

Welker, M., Maršálek, B., Šejnohová, L. & von Döhren, H. (2006). Detection and identification of oligopeptides in *Microcystis* (cyanobacteria) colonies: toward an understanding of metabolic diversity. *Peptides* **27**, 2090–2103.

Whitmore, L. & Wallace, B. A. (2004). The peptaibol database: a database for sequences and structures of naturally occurring peptaibols. *Nucleic Acids Res* **32**, D593–D594.

Wiest, A., Grzegorski, D., Xu, B. W., Goulard, C., Rebuffat, S., Ebbola, D. J., Bodo, B. & Kenerley, C. (2002). Identification of peptaibols

from *Trichoderma virens* and cloning of a peptaibol synthetase. *J Biol Chem* **277**, 20862–20868.

Williams, B. H., Hathout, Y. & Fenselau, C. (2002). Structural characterization of lipopeptide biomarkers isolated from *Bacillus globigii*. *J Mass Spectrom* **37**, 259–264.

Yun, B.-S., Yoo, I.-D., Kim, Y. H., Kim, Y.-S., Lee, S. J., Kim, K. S. & Yeo, W. H. (2000). Peptaivirins A and B, two new antiviral peptaibols against TMV infection. *Tetrahedron Lett* **41**, 1429–1431.

Zhang, J., Rosenberg, H. F. & Nei, M. (1998). Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc Natl Acad Sci U S A* **95**, 3708–3713.

Edited by: N. L. Glass