DOI 10.1002/pmic.200600280

RESEARCH ARTICLE

Structural protein analysis of the polyvalent staphylococcal bacteriophage 812

Luděk Eyer[1], Roman Pantůček[1], Zbyněk Zdráhal[2], Hana Konečná[2], Petr Kašpárek[1], Vladislava Růžičková[1], Lenka Hernychová[3], Jan Preisler[4] and Jiří Doškař[1]

[1] Department of Genetics and Molecular Biology, Faculty of Science, Masaryk University, Kotlářská 2, Brno, Czech Republic
[2] Department of Functional Genomics and Proteomics, Faculty of Science, Masaryk University, Kamenice 5, Brno, Czech Republic
[3] Proteome Center for the Study of Intracellular Parasitism of Bacteria, Faculty of Military Health Sciences, University of Defence, Třebešská 1575, Hradec Králové, Czech Republic
[4] Department of Analytical Chemistry, Faculty of Science, Masaryk University, Kotlářská 2, Brno, Czech Republic

Phage 812 is a polyvalent phage with a very broad host range in the genus *Staphylococcus*, which makes it a suitable candidate for phage therapy of staphylococcal infections. This proteomic study, combining the results of both 1-DE and 2-DE followed by PMF, led to the identification of 24 virion proteins. Twenty new proteins, not yet identified by proteome analysis of closely related staphylococcal phages K and G1 were identified using this approach. Fifteen proteins were assigned unambiguously to the head–tail genome module; the remaining nine proteins are encoded by genes of the left or right arms of the phage genome. As expected, the most abundant proteins in the electrophoretic patterns are the major capsid protein, the major tail sheath protein and proteins identical to ORF 50 and ORF 95 of phage K, although their function is only putative. Identification of these 20 new proteins contributes substantially to a detailed characterization of phage virions, knowledge of which is necessary for rational phage therapy.

Received: April 14, 2006
Revised: August 1, 2006
Accepted: September 29, 2006

Keywords:
Bacteriophage / Phage therapy / *Staphylococcus*

1 Introduction

Polyvalent staphylococcal bacteriophages of the family *Myoviridae* have recently attracted considerable interest with respect to their practical application for phage therapy. This approach offers a promising alternative to antibiotic treatment of staphylococcal infections and could be based on various phage features [1–4]. Phages K, G1 and Twort, whose complete genome sequences are known, are the best-explored phages of the species Twort. They all belong to the family *Myoviridae* [5, 6] along with the phage 812, the phenotype of which is known in detail and which has a very broad host range within the genus *Staphylococcus*. The genome DNA sequencing of phage 812 is in progress [7]. At the nucleotide level, phages K and G1 are 90% identical, whereas the relationship between Twort and G1 or K is not as close (51–52%). From comparison of the sequenced portions of the phage 812 genome, and from the restriction map, it is evident that phage 812 is closely related to lytic phages K and G1 (Fig. 1). Their genome sizes exceed 125 kbp and they show a similar modular organization. Another characteristic in common is the presence of group I introns in different genes [6, 8–10]. The genomes of polyvalent staphylococcal phages include about 200 putative ORFs. More than half of these ORFs code

Correspondence: Professor Jiří Doškař, Department of Genetics and Molecular Biology, Faculty of Science, Masaryk University, Kotlářská 2, 61137 Brno, Czech Republic
E-mail: doskar@sci.muni.cz
Fax: +420-549-492-570

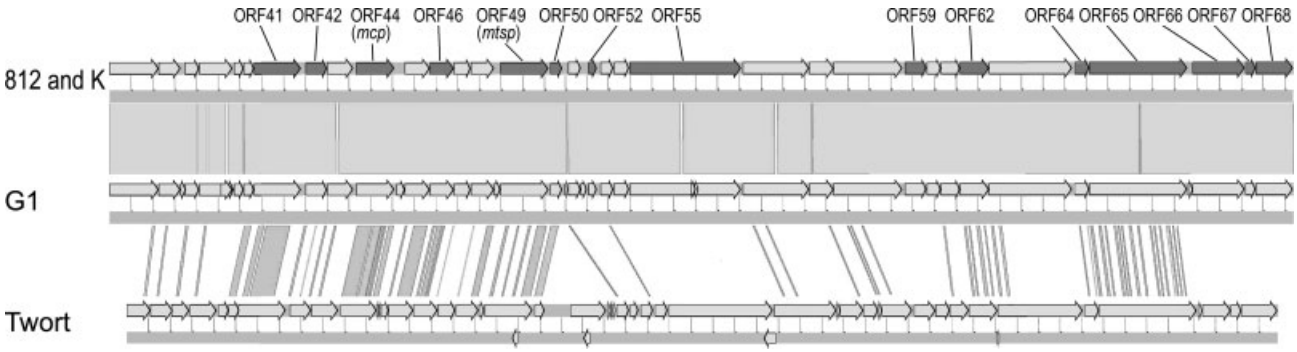

Figure 1. Nucleotide sequence comparison of the packaging and the head–tail genome modules of related phages 812, K (GenBank nucleotide position 34 848–78 346), G1 (13–43 524), and Twort (1 284–43 570) displayed using the Artemis Comparison Tool. The dark arrows represent ORFs encoding proteins identified in the phage 812 and are marked according to those of phage K.

for proteins which are not incorporated into the phage particle. All of these phages are characterized by a similar morphology (an isometric head and relatively rigid tail, with a contractile sheath and base plate) and are assigned to the genus of '*SPO1-like viruses*'.

Both a strong lytic activity and a potential for lysis of the highest number of staphylococcal strains are important for the phage therapy. The number of sensitive strains could be increased by isolating phage host range mutants. In our earlier paper [7], we characterized in detail mutants of the polyvalent phage 812 in a host range; these phages are able to lyse up to 95% of all *Staphylococcus aureus* strains and 43% strains of other *Staphylococcus* species. Therefore, phage 812 and its mutants are considered good candidates for phage therapy of staphylococcal infections as anti-staphylococcal vaccines [11] and preparations with high lytic activity [12–15]. The phage coat consists of a relatively large number of proteins. As phage proteins can also interact directly with mammalian cells and with both the humoral and the innate immune system [16], and because some gastrointestinal or allergic symptoms occurred in a few cases during phage therapy trials [17], detailed knowledge of phage coat protein immunogenicity is required for the development of a new generation of therapeutic phage preparations with precisely defined parameters. Generally, weak sequence similarities of tailed phage genes (about 15–20%) [18] complicate the use of genomic methods to identify functions and properties of phage gene products. Thus, we decided to use a proteomics approach in our work.

For the reasons above, this work covers both characterization of the phage proteome in detail and the identification of phage virion proteins. Proteomic methods are under continuous development – most of the tools used for protein identification have shown to be useful in specific cases [19, 20]; however, none of the techniques are applicable to all the cases. Using different tools for analysis of the same sample usually leads to more comprehensive results [21]. For analysis of a simple phage proteome, we have applied two closely related separation techniques, 1-DE and 2-DE, followed by MALDI MS PMF. To identify the maximum number of proteins, an optimization of 1-DE and 2-DE separation parameters was necessary, because these techniques have not been extensively used for staphylococcal phage proteome analysis yet.

2 Materials and methods

2.1 Apparatus and chemicals

Ultracentrifuge UP 65 was from MLW (Leipzig, Germany). The Protean II xi Cell, Protean II xi Multi-Cell, QuantityOne software, PDQuest 7.3.0 software and the GS-800 Calibrated Densitometer were from BioRad (Hercules, CA, USA). Multiphor II was from GE Healthcare Biosciences (Uppsala, Sweden). Personal Densitometer was from Molecular Dynamics (Sunnyvale, CA, USA). Bruker Reflex IV (MALDI-TOF system), XMASS 5.1.5 and MS Biotools 2.0 software were manufactured by Bruker Daltonik (Bremen, Germany). SpeedVac Vacuum Concentrator SPD111V was from ThermoSavant (Holbrook, NY, USA).

SDS, Tris-base, DNase I, SigmaMarker Wide Range (6.5–205 kDa), SigmaMarker High Range (36–205 kDa), TEMED, APS, ProteoSilver Plus Silver Stain Kit, Ampholyte (pH 9–11), CHAPS and urea were obtained from Sigma (St. Louis, MO, USA). The Colloidal Blue Staining Kit was from Invitrogen (San Diego, CA, USA). Bio-Safe Coomassie G-250, 1-D SDS-PAGE Standards Low Range (14.4–97.4 kDa) and 2-D SDS-PAGE Standards (p*I* range, pH 4.5–8.5; MW range, 17.5–76 kDa) were from BioRad. DTT was from Serva Electrophoresis (Heidelberg, Germany). Bromophenol blue and CsCl were from Merck (Darmstadt, Germany). Immobiline DryStrips, DeStreak, Pharmalyte (pH 3–10) and Pharmalyte (pH 8–10.5) were obtained from GE Healthcare Biosciences. Rotiphorese gel 30 was obtained from Carl Roth (Karlsruhe, Germany). Nutrient Broth CM1 was from Oxoid (Basingstoke, UK). Trypsin was from Promega (Madison, MA, USA). Other chemicals were at least of proanalysis quality and were obtained from local manufacturers.

2.2 Strain and growth conditions

Bacteriophage 812 and its host *S. aureus* SA812 (= CCM 4028) were obtained from Dr. J. Pillich (Institute of Biophysics, Brno, Czech Republic). The phage 812 was propagated on strain *S. aureus* SA812, cultivated aerobically at 37°C in Nutrient Broth CM1, as described by Pantůček *et al.* [7]. The prepared phage lysate was used for sample preparation.

2.3 Sample preparation for 1-DE and 2-DE

Phage lysate (total volume 500 mL, ~10^{10} PFU/mL) was centrifuged at low speed to remove cellular debris. The phages were harvested from the lysate by ultracentrifugation at 55 000 × *g* for 2.5 h. The pellet was resuspended in phage buffer containing 10 mM NaCl, 10 mM $CaCl_2$ and 50 mM Tris-HCl (pH 8) and stored overnight at 4°C. Subsequently, the bacteriophage 812 was purified by CsCl density gradient centrifugation (1.45 g/mL, 1.50 g/mL and 1.70 g/mL; CsCl was dissolved in phage buffer). CsCl-purified phages were dialyzed against water for 24 h; any phage DNA released, causing high sample viscosity, was removed by addition of DNase I (100 μg/mL). The purified phages were concentrated in a vacuum concentrator.

2.4 1-DE

An aliquot of the concentrated phage proteins was boiled for 3 min in sample buffer (0.175 M Tris-HCl (pH 6.8), 15% w/v glycerol, 5% w/v SDS, 4.65% w/v DTT, a trace of bromophenol blue). Vertical 1-DE was performed on 15%T and 7.5%T acrylamide gels. Electrophoresis was performed on a Protean II xi Cell at constant power (15 mA in the stacking gel and 30 mA during separation) in running buffer (0.025 M Tris-HCl, 0.192 M glycine, 0.1% w/v SDS). The gels were 16 × 20 × 0.1 cm^3. The protein molecular weight markers were SigmaMarkers, wide and high range. Proteins were stained with the Bio-Safe Coomassie G-250 or with ProteoSilverPlus kit, scanned using a laser densitometer and the resulting traces were processed with QuantityOne and PDQuest software.

2.5 2-DE

Bacteriophage proteins were precipitated overnight in 20% TCA in acetone (–18°C) containing 0.2% DTT [22], and then solubilized in IEF buffer containing 9 M urea, 4% w/v CHAPS, 70 mM DTT, 40 mM Tris-base and 5% v/v carrier ampholytes (pH 9–11). The protein concentration in IEF buffer was determined by a modified BCA assay [23]. Immobiline DryStrips, 18 cm, with a nonlinear pH 3–10 gradient were used. In-gel rehydration with 75 μg of protein (for silver stained gels) or with 350 μg of protein (for gels stained by Colloidal Blue) was carried out overnight. Rehydration solution contained 2 M thiourea, 6 M urea, 4% w/v CHAPS, 40 mM Tris-base, 12 μL/mL DeStreak, 0.003% w/v bromophenol blue, 1% v/v pharmalyte (pH 3–10) and 0.5% v/v pharmalyte (pH 8–10.5). The following IEF running conditions were used: 300 V for 30 min, 600 V for 30 min, 1000 V for 30 min, 2000 V for 30 min, 2500 V for 30 min, 3000 V for 30 min, 3500 V for 3 h and 5000 V for 12 h. In the second dimension, a gradient gel (9–16%T) was used. Electrophoresis was run on the Protean II xi Multi-Cell, at a constant current of 5 mA/gel for 1.5 h, and then at 40 mA/gel for 5 h. For a 2-DE reference map of the bacteriophage, analytical gels were silver stained by a protocol incompatible with MS analysis [24]. The micropreparative gels were stained either with Colloidal Blue Staining Kit or silver staining protocol compatible with MS analysis, according to Shevchenko *et al.* [25]. The p*I*s and molecular weights of proteins were approximated using 2-D SDS-PAGE standards. The analytical and micropreparative gels were scanned using a laser densitometer and processed by PDQuest software.

2.6 Tryptic digestion and MALDI-TOF MS analysis

Proteins to be analysed were excised from 1-DE and 2-DE gels. After destaining, the proteins were incubated with trypsin at 37°C for 16 h. Digested peptides were extracted from gels using 50% ACN in water containing 5% formic acid. This digestion protocol is based on the procedure described by Shevchenko *et al.* [25]. MALDI-MS analyses were performed on a Bruker Reflex IV operated in the reflector mode with the detection of positively charged ions. The matrix was a CHCA solution prepared according to Havliš *et al.* [26], used in combination with an AnchorChip target to enhance sensitivity. Samples (1 μL) were mixed with matrix solution on the target at 2:1 ratio. The known autoproteolytic products of trypsin were used as internal standards for mass calibration. If these products were absent, an external calibration procedure was employed. The mixture of five known peptides covered the mass range of 660–2465 Da. The XMASS 5.1.5 and MS Biotools 2.0 software were used for data processing. MALDI TOF/TOF MS analysis was carried out using Applera 4700 Proteomics Analyser. Spectra were acquired in positive mode with helium as the collision gas. Precursor selection was limited to 15 ions with the highest intensities. MS/MS data were evaluated using GPS Explorer Workstation.

2.7 Data processing

MASCOT 2.0 (MatrixScience, London, UK) search engine was used for processing the MS data. PMF was done against the NCBI protein database (release 20051121) as well as against translated genome sequence data of bacteriophage K and genome sequence data of bacteriophage 812. One enzyme missed cleavage and a mass tolerance of 60 ppm were set for all searches. Oxidation of methionine and carbamidomethylation of cysteine as optional and fixed modifications, respectively, were allowed.

The program COGnitor (www.ncbi.nlm.nih.gov/COG) was used to determine the functions of identified proteins. BLASTp 2.2.13 and BLAST 2 sequences (www.ncbi.nlm.nih.gov/blast) were applied for amino acid sequence comparison of identified proteins. Nucleotide sequences of packaging and head–tail modules of phages K (GenBank accession No. AY176327), G1 (AY954969), 812 (EF136581) and Twort (AY954970) were compared by Artemis Release 8 (www.sanger.ac.uk/Software/Artemis). Putative conserved domains were identified by InterProScan 4.2 (www.ebi.ac.uk/InterProScan).

3 Results and discussion

3.1 Proteins detected by 1-DE

Due to the relative simplicity of phage proteomes, 1-DE was selected for initial proteomic analysis. Approx. 40 gel bands of proteins in the 10–150 kDa range were localized (Fig. 2). Because the currently determined nucleotide sequences of the head–tail genome module of phage 812 show 100% identity with the sequences of phage K and a high degree of similarity (99%) with phage G1 (Fig. 1), sequences of these phages were also used for PMF. PMF identified 11 proteins (Table 1). Based on their similarity with proteins of related phages, ten proteins can be assigned to the group of structural proteins. No function could be assigned to one detected protein using this approach.

In addition to proteins migrating at positions on the gel corresponding to their molecular weights, several bands were seen to have identified proteins of considerably higher molecular weights than the gel slice. With a single exception, all of these bands contained fragments of identified proteins as indicated by fingerprints of peptides that span only a portion of a protein. Interestingly, the major tail sheath protein (mtsp; 64.7 kDa) was also identified in bands at ~30 kDa in addition to the band at ~64 kDa. Using modified conditions (7.5% gel), three sub-bands with MW in the 30–35 kDa region were resolved (data not shown). PMF with MALDI TOF MS together with more specific MALDI TOF/TOF MS analysis revealed that two of the sub-bands contained the N-terminal part of the protein and the remaining sub-band contained the C-terminal part of the protein. This cleavage of the major tail sheath protein approximately at the centre may be related to a phenomenon noticed in the tail sheath protein of *Lactobacillus* phage LP65 [27]. Structural modifications of the protein may be linked to the contraction of the sheath tail. Post-translational cleavage of structural proteins of some bacteriophages was also detected [28].

3.2 Proteins detected by 2-DE

The high separation power of 2-DE played a major role in an advanced phage proteome analysis. Gels stained with Colloidal Blue showed about 120 separated spots of which the 42 most intensive spots were further analysed by MS. Gels stained with MS-incompatible silver staining showed about 700 spots (Fig. 3). However, the number of spots obtained by MS-compatible silver staining was significantly lower; only 190 of the most intense spots were further processed by MS analysis (Fig. 4). PMF identified 22 proteins, 9 of them represented proteins found by 1-DE approach earlier (Table. 1). Major proteins were present in numerous chained spots as charged mass variants. Database search showed sequence similarities of 17 proteins with the corresponding proteins of phages K and G1. Remaining five proteins showed similarities only with the phage G1 proteins (ORF 055, ORF 138, ORF 172, ORF 189 and ORF 207).

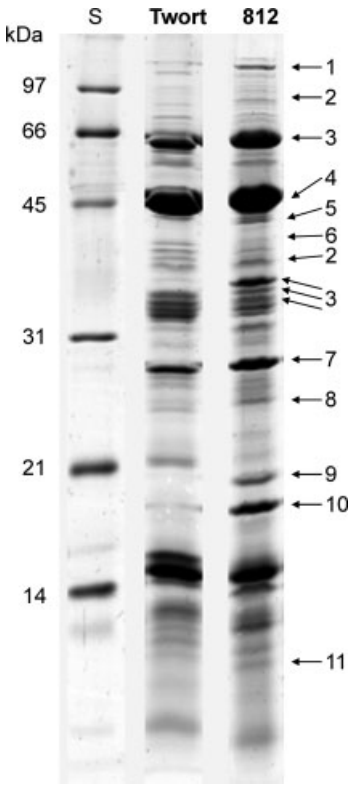

Figure 2. Coomassie G-250 stained 1-DE pattern of proteins from phages 812 and Twort; S = SDS-PAGE Standards Low Range (BioRad). Arrows and numbers indicate bands containing the identified proteins of the phage 812. Band numbers are given in Table 1.

Bioinformatic analysis revealed that nucleotide sequences of the phage G1 coding ORFs mentioned above were also present in the genome sequence of the phage K [6]; however, they were not annotated as ORFs [5]. In four cases, suggested sequences are localized always between two known ORFs of the phage K (ORF 207 of the phage G1 between ORF 33 and ORF 34 of the phage K, ORF 189 of the phage G1 between ORF 96 and ORF 97 of the phage K, ORF 172 of the phage G1 between ORF 18 and ORF 19 of the phage K and ORF 138 of the phage G1 between ORF 7 and ORF 8 of the phage K).

Table 1. Characteristics of the phage 812 identified proteins

Identified protein[a)]	Function [5, 6, 27]	Putative conserved domain	MW (kDa)[b)]	$MW_{exp.}$ (kDa)[c)]		p*I*[d)]	$pI_{exp.}$[e)]	MASCOT score[f)]	*n*/%[g)]	Band/spot number[h)]	
				1-DE	2-DE					1-DE	2-DE
K: ORF 23[i)]	Unknown	–	23.0*	–	38.2	4.0*	<4.5	57	4/26	NI	122
K: ORF 41	Portal protein	Phage portal	64.0*	–	53.9; 55.6	6.1*	5.90; 5.85	217	17/32	NI	42, 44
					45.8		>8.5				78
K: ORF 42	Major capsid protein	Caudovirus prohead protease	28.6*	–	13.8	4.9*	4.76	96	7/31	NI	182
812: mcp, K: ORF 44	Major capsid protein	–	51.2	45.5	44.8–51.6	5.0	4.79–5.82	236	19/61	4	46, 49, 51, 52, 55, 57–60, 62, 63, 65–69, 71, 74, 75, 80, 81, 85, 88
					37.2–42.7		4.82–4.92				104, 111, 112, 115, 116, 119, 131
					29.7; 27.7		5.03; 5.15				149, 158
812: ORF 4, K: ORF 46	Unknown	–	33.7*	–	37.2	5.6*	5.64	100	6/35	NI	132
812: mtsp, K: ORF 49	Major tail sheath protein	–	64.5	63.3 36.1	62.5–65.4	4.7	4.95–5.62	360	28/56	3	18, 20–29
				35.5 33.8	41.6–44.8		4.86–4.99				48, 86, 92, 98, 107
					38.2;34.9–35.7		4.73;4.93–5.66				130, 133, 135–137
812: ORF 8, K: ORF 50	Unknown	–	15.9	18.7	21.4;21.9	5.3	5.38;5.15	126	8/59	10	191, 192
K: ORF 52[i)]	Unknown	–	12.2*	11.2	–	5.9*	–	53	3/25	11	NI
K: ORF 55	Tape measure protein	Lysozyme-like domain, ATPase	143.7*	120.2	–	9.1*	–	334	10/14	1	NI
K: ORF 59	Unknown	–	29.3*	–	33.0	7.8*	>8.5	202	14/49	NI	140
K: ORF 62	Tail protein	–	39.2*	39.5	42.9–44.5	4.6*	5.09–5.34	110	7/24	6	90, 101–103, 105
K: ORF 64	Unknown	–	19.2*	20.4	<17.5	5.3*	~5.50	222	13/85	9	visible only on CBB stained gels
K: ORF 65	Virulence-associated protein	Sialidase	129.0*	89.2 38.2	>76.0 75.9 38.5	4.9*	4.94–4.96 4.71 5.51	316	35/45	2	1–3 6 121
K: ORF 66	Unknown	–	72.5*	–	65.0–67.7 46.7	6.5*	5.76- 6.72 6.96	381	29/56	NI	13, 14, 16, 17 76
K: ORF 67	Unknown	–	14.6*	–	18.2	4.7*	4.72	174	11/91	NI	174
K: ORF 68	Unknown	–	50.4*	43.1	43.5–44.5	6.0*	5.74–5.89	384	25/75	5	89, 91, 93, 94
K: ORF 95	Unknown	–	23.2*	28.7	31.5 16.3	4.7*	4.88 6.41	237	14/80	7	141 178
K: ORF 96	Major tail protein	Bacterial Ig-like domain (group 2), invasin/ intimin	17.8*	25.9	35.8 – 40.3	4.3*	< 4.5	123	6/50	8	70, 109, 134
					29.0		<4.5; >8.5				147, 163
K: ORF 111[i)]	Nicotinamid phosphoribosyl transferase	Nicotinate phosphoribosyl transferase	56.1*	–	38.5	5.3*	5.27	73	7/9	NI	123
G1: ORF 055	Unknown	–	25.8**	–	30.6	4.9**	4.94	112	8/35	NI	146
G1: ORF 138	Structural	–	11.8**	–	14.7	6.3**	5.87	104	5/67	NI	181
G1: ORF 172	Unknown	–	10.1**	–	13.9	5.7**	5.68	168	9/83	NI	183

Table 1. Continued

Identified protein[a)]	Function [5, 6, 27]	Putative conserved domain	MW (kDa)[b)]	$MW_{exp.}$ (kDa)[c)]		p*I*[d)]	p$I_{exp.}$[e)]	MASCOT score[f)]	*n*/%[g)]	Band/spot number[h)]	
				1-DE	2-DE					1-DE	2-DE
G1: ORF 189	Major tail protein	–	7.8**	–	16.0	4.4**	<4.5	104	5/92	NI	175
G1: ORF 207	Unknown	–	8.0**	–	11.9	5.7**	5.52	134	6/79	NI	190

a) As the phage K and G1 genomes are highly similar, ORFs 23-111 of the phage K also correspond to proteins in the phage G1 (not listed in Table 1). Matching proteins of the phage G1 are shown in Table 2.
b) MW calculated from the protein database (expected MW).
c) Observed MW.
d) p*I* calculated from the protein database (expected p*I*).
e) Observed p*I*.
f) The highest number is shown in the case of protein identification in multiple spots or in both 1-DE and 2-DE.
g) Peptides matched/sequence coverage (the highest number shown in the case of protein identification in multiple spots).
h) According to Figs. 2 and 4.
i) Protein was positively identified against phage K database but did not passed against NCBI database. We consider it as tentatively identified and further confirmation is necessary.
* MW and p*I* derived from ORF sequence of the phage K.
** MW and p*I* derived from ORF sequence of the phage G1.
NI = not identified by the given type of electrophoresis

Moreover, ORF 138 and ORF 189 of the phage G1 show sequence similarity with ORF 96 of the phage K [6]. The sequence of 81 nucleotides of the ORF 055 of the phage G1 shows 100% identity with the corresponding sequence of ORF 108 of phage K (Table 2).

There were two proteins (showing sequence similarities with the ORF 52 and ORF 55 of the phage K) identified only on 1-DE gel, presumably because they were not among the

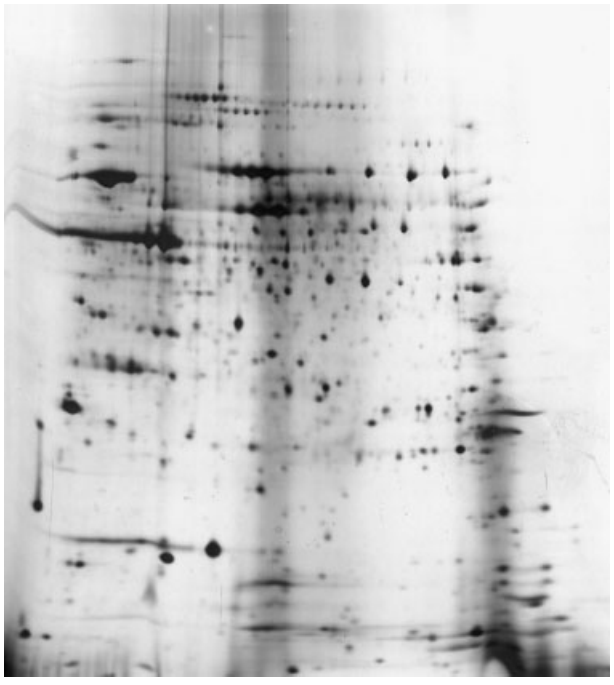

Figure 3. 2-DE spot pattern of proteins from phage 812; pH gradient 3–10; MS-incompatible silver staining.

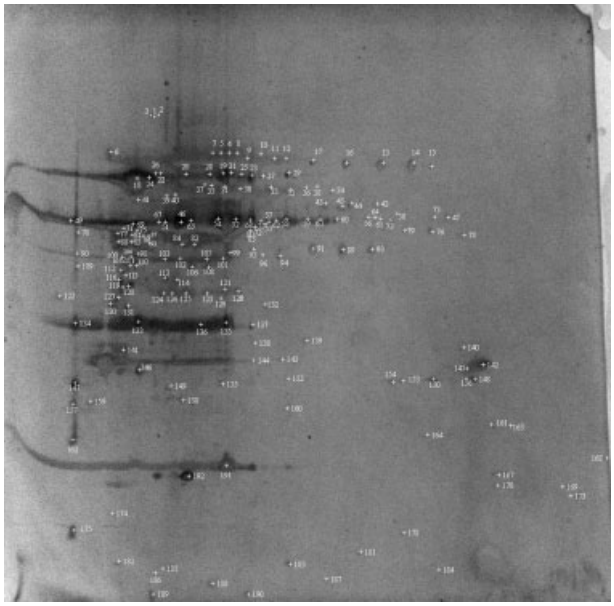

Figure 4. 2-DE spot pattern of proteins from phage 812; pH gradient 3–10; MS-compatible silver staining. Analyzed spots are marked. Spot numbers are given in Table 1.

190 most intense spots selected for the analysis. In the case of proteins separated by both 1-DE and 2-DE, protein fragments might be present (*e.g.* the major capsid protein and protein identified as ORF 65 of the phage K discussed above). In addition, some of the proteins separated by 2-DE occur as a series of spots of the same mobility, but varying in p*I* values (*e.g.* the major capsid protein, the major tail sheath protein and ORF 8). This could be caused by protein post-translation modifications or by point mutations in the nucleotide

Table 2. Assignment of the phage 812 identified proteins to ORFs of related phages K, G1, Twort, LP65, A511 and P100

K [5]	G1 [6]	Twort [6]	LP65 [27]	A511 [37]	P100 [38]
ORF 23	ORF 063	ORF 070	–	–	–
ORF 41	ORF 014	ORF 012	ORF 112	–	gp 14
ORF 42	ORF 048	ORF 040	ORF 111	ORF 1	gp 15
ORF 44	ORF 016	ORF 016	ORF 109	Cps	gp 17
ORF 46	ORF 034	ORF 034	ORF 107	ORF 4	gp 20
ORF 49	ORF 011	ORF 011	ORF 103	Tsh	gp 24
ORF 50	ORF 105	ORF 106	ORF 102	ORF 8	gp 25
ORF 52	ORF 141	ORF 125	ORF 100	ORF 9	gp 26
ORF 55	ORF 001	ORF 002	ORF 98	–	gp 28
ORF 59	ORF 043	ORF 041	ORF 129, 130	–	gp 31
ORF 62	ORF 027	ORF 026	ORF 95	–	gp 34
ORF 64	ORF 079	ORF 073	ORF 91	–	gp 36
ORF 65	ORF 002	ORF 003	ORF 90	–	gp 37
ORF 66	ORF 008	ORF 024	–	–	–
ORF 67	ORF 117	ORF 107	–	–	–
ORF 68	ORF 017	ORF 017, 024	–	–	–
ORF 95	ORF 059	ORF 056	–	–	gp 77
ORF 96	ORF 080	–	–	–	–
	ORF 138	ORF 001, 178			
	ORF 189	ORF 178			
ORF 111	ORF 018	–	–	–	gp 64
ORF 108	ORF 055	–	–	–	–
–	ORF 172	ORF 150	–	–	–
–	ORF 207	ORF 222	–	–	–

sequence of the relevant ORFs [29]. 2-DE also confirmed anomal migration of the major tail sheath protein in the gel. As expected, the major tail sheath protein fragments were easily separated using 2-DE due to their differences in p*I*.

3.3 Functions of the identified proteins

Consequently, 24 proteins of the phage 812 were identified by both the separation techniques (Table 1). Among the identified proteins of the phage 812, there are four proteins found earlier by electrophoretic analysis of the proteome and also by N-terminal sequencing of the related phage K [5], including ORF 44, ORF 49, ORF 50 and ORF 95. Twenty proteins were identified for the first time. Based on the high level of homology of head–tail modules and the other genomic sequences and on the similarity of restriction endonuclease patterns for phages 812 and K, it can be concluded that the proteins identified in our study are most likely to be a part of the virion structure in polyvalent bacteriophages of the species Twort.

Altogether, 15 proteins of the phage 812 are encoded by genes belonging to the head–tail module. The majority of the phage capsid consists of the major capsid protein (mcp; 51.2 kDa). It can be concluded from its localization in the most intense band that it is present with the highest copy number in the phage particle. In addition to the major capsid protein, one more protein was found; it was incorporated into the capsid structure (28.6 kDa) containing the Caudovirus prohead peptidase domain. This group of peptidases belongs to the MEROPS peptidase family U35 (clan U-). This family contains the prohead protease from HK97 and related phages or prophages. It is generally encoded adjacent to the gene for the capsid protein that is processed, and in some cases may be fused to it [30]. The portal protein (64.0 kDa) contains a conserved domain typical for phage portal proteins, forms a dodecamer and is located at a five-fold vertex of the viral capsid. The portal complex forms a channel through which the viral DNA is packaged into the capsid and exits during infection [31]. The second most intense band is the major tail sheath protein (64.5 kDa), which makes the tail contraction possible during the process of infection. The tape measure protein (143.7 kDa) determines the tail length, being used as a template for the tail polymerisation and remaining inside the tail tube. These proteins are the longest proteins encoded by the phages. The polypeptide chain lengths of the tape measure proteins of the phages 812 or K (1351 amino acids) and the phage Twort (1269 amino acids) correspond to lengths of their tails (205 and 150 nm in the phages 812 and Twort, respectively). Some of these proteins share a conserved domain, which appears to have the features of lytic transglycosylases [32]. The protein identical to ORF 65 of the phage K (129.0 kDa) was assigned as the virulence-associated protein. This protein sequence contains the sialidase domain. Sialidases (neuraminidases) hydrolyse the nonreducing, terminal sialic acid linkage in various natural substrates such as glycoproteins, glycolipids, gangliosides and polysacharides [33]. In viruses, sialidases enable the transport of the virus through mucin, the release of the virus from the infected host cell and prevent self-aggregation of virus particles [34]. The protein identical to ORF 62 of the phage K (39.2 kDa) was assigned as the tail protein. The functions of the other identified proteins encoded by genes of the head–tail module are not known.

Proteins encoded by the genes identified as ORF 95, ORF 96 and ORF 111 of the phage K and ORF 055 and ORF 189 of the phage G1 are located on the right arm of the phage DNA. Although these genes are not a part of the head–tail module, ORF 96 and ORF 189 code for other structural proteins annotated as the major tail proteins (17.8 and 7.8 kDa). The bacterial immunoglobulin-like domain (group 2) is a component of ORF 96. Proteins that contain this domain are found in a variety of bacterial and phage surface proteins such as intimins. Intimin is a bacterial cell-adhesion molecule that mediates the intimate bacterial host-cell interaction [35]. ORF 111 of the phage K was assigned to nicotinamide phosphoribosyl transferase (56.1 kDa). Nicotinamide phosphoribosyl transferase is the rate-limiting enzyme that catalyzes the first reaction in the NAD salvage synthesis [36]. The presence of the enzyme as a structural component of the

phage particle is questionable. Genes encoding proteins identical with ORF 23 of phage K and ORF 138, ORF 172 and ORF 207 of phage G1 occur on the left arm of the phage genome. Their functions are unknown.

3.4 Similarity of the phage 812 with related phages

Identified proteins of the phages 812 and K show 99–100% similarity with the corresponding proteins of the phage G1 and approximately 60% similarity with proteins of the phage Twort. Genome size and structure, and many similarities on the protein sequence level occur also among other phages of *Myoviridae* family. Their hosts are members of the low-GC branch of Gram-positive bacteria, *e.g. Listeria monocytogenes* (phages A511 [37] and P100 [38]) and *Lactobacillus plantarum* (phage LP65) (Table 2). By application of 1-DE, six major proteins of the phage LP65 were discovered. 2-D gels revealed the same number of proteins, later identified by MS [27].

4 Concluding remarks

As proved by gel electrophoresis and PMF, the virion of polyvalent staphylococcal phage 812 is formed of at least 24 proteins. Although the phage proteome is much simpler compared to that of cellular organisms, using 2-DE (with both CBB and silver staining) in addition to 1-DE significantly increased the number of identified proteins. Using this simple methodic approach, most of the structural proteins were observed in phage 812. Fifteen of the identified proteins are encoded by genes of the head–tail genome module. Twenty proteins of phage 812, which were not identified yet in related lytic phages K and G1, set the basis for a detailed analysis of the phage virion structure and are a prerequisite for their efficient and safe application in phage therapy.

We gratefully acknowledge the financial support of the Czech Science Foundation (Grant No. 203/03/0515), the Ministry of Education, Youth and Sports of the Czech Republic (MSM 0021622415), and grant LSHM-CT-2006-019064 from the European Union. We also would like to thank Gabriela Lochmanová for image analysis and data processing by PDQuest, to prof. Barry L. Karger, Tomáš Rejtar and Hsuan-shen Chen for their help with TOF/TOF MS experiments and to Roger Kautz for his helpful suggestions.

5 References

[1] Sulakvelidze, A., Alavidze, Y., Morris, J. G., *Antimicrob. Agents. Chemother.* 2001, *45*, 649–659.

[2] Liu, J., Dehbi, M., Moeck, G., Arhin, F. *et al.*, *Nat. Biotechnol.* 2004, *22*, 185–191.

[3] Embleton, M. L., Nair, S. P., Heywood, W., Menon, D. C. *et al.*, *Antimicrob. Agents Chemother.* 2005, *49*, 3690–3696.

[4] Wills, Q. F., Kerrigan, C., Soothill, J. S., *Antimicrob. Agents Chemother.* 2005, *49*, 1220–1221.

[5] O'Flaherty, S., Coffey, A., Edwards, R., Meaney, W. *et al.*, *J. Bacteriol.* 2004, *186*, 2862–2871.

[6] Kwan, T., Liu, J., DuBow, M., Gros, P., Pelletier, J., *Proc. Natl. Acad. Sci. U. S. A.* 2005, *102*, 5174–5179.

[7] Pantůček, R., Rosypalová, A., Doškař, J., Kailerová, J. *et al.*, *Virology* 1998, *246*, 241–252.

[8] Landthaler, M., Shub, D. A., *Proc. Natl. Acad. Sci. USA* 1999, *96*, 7005–7010.

[9] Landthaler, M., Begley, U., Lau, N. C., Shub, D. A., *Nucleic Acids Res.* 2002, *30*, 1935–1943.

[10] Golden, B. L., Kim, H., Chase, E., *Nat. Struct. Mol. Biol.* 2005, *12*, 82–89.

[11] Pillich, J., Tovarek, J., Fait, M., *Z. Orthop. Ihre. Grenzgeb.* 1978, *116*, 40–46.

[12] O'Flaherty, S., Ross, R. P., Meaney, W., Fitzgerald, G. F. *et al.*, *Appl. Environ. Microbiol.* 2005, *71*, 1836–1842.

[13] O'Flaherty, S., Coffey, A., Meaney, W., Fitzgerald, G. F., Ross, R. P., *J. Bacteriol.* 2005, *187*, 7161–7164.

[14] O'Flaherty, S., Coffey, A., Meaney, W. J., Fitzgerald, G. F., Ross, R. P., *Lett. Appl. Microbiol.* 2005, *41*, 274–279.

[15] Jikia, D., Chkhaidze, N., Imedashvili, E., Mgaloblishvili, I. *et al.*, *Clin. Exp. Dermatol.* 2005, *30*, 23–26.

[16] Dabrowska, K., Switala-Jelen, K., Opolski, A., Weber-Dabrovska, B., Gorski, A., *J. Appl. Microbiol.* 2005, *98*, 7–13.

[17] Slopek, S., Durlakowa, I., Weber-Dabrowska, B., Kucharewicz-Krukowka, A. *et al.*, *Arch. Immunol. Ther. Exp.* 1983, *31*, 293–327.

[18] Ackermann, H.-W., Elzanowski, A., Fobo, G., Stewart, G., *Arch. Virol.* 1995, *140*, 1871–1884.

[19] Wang, H., Hanash, S., *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* 2003, *787*, 11–18.

[20] Foret, F., Preisler, J., *Proteomics* 2002, *2*, 360–372.

[21] Wang, J. Q., Xue, Y. F., Feng, X. L., Li, X. L. *et al.*, *Proteomics* 2004, *4*, 136–150.

[22] Görg, A., Obermaier, C., Boguth, G., Csordas, A. *et al.*, *Electrophoresis* 1997, *18*, 328–337.

[23] Brown, R. E., Jarvis, K. L., Hyland, K. J., *Anal. Biochem.* 1989, *180*, 136–139.

[24] Hochstrasser, D. F., Merril, C. R., *Appl. Theor. Electrophor.* 1988, *1*, 35–40.

[25] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., *Anal. Chem.* 1996, *68*, 850–858.

[26] Havliš, J., Thomas, H., Šebela, M., Shevchenko, A., *Anal. Chem.*, 2003, *75*, 1300–1306.

[27] Chibani-Chennoufi, S., Dillmann, M.-L., Marvin-Guy, L., Rami-Shojaei, S., Brüssow, H., *J. Bacteriol.* 2004, *186*, 7069–7083.

[28] Laemmli, U. K., *Nature* 1970, *227*, 680–685.

[29] Hubálek, M., Hernychová, L., Brychta, M., Lenčo, J. *et al.*, *Proteomics* 2004, *4*, 3048–3060.

[30] Duda, R. L., Martincic, K., Hendrix, R. W., *J. Mol. Biol.* 1995, *247*, 636–647.

[31] Moore, S. D., Prevelige, P. E. Jr., *Curr. Biol.* 2002, *12*, R96–98.

[32] Lehnherr, H., Hansen, A.-M., Ilyina, T., *Mol. Microbiol.* 1998, *30*, 454–457.

[33] Monti, E., Preti, A., Venerando, B., Borsani, G., *Neurochem. Res.* 2002, *27*, 649–663.

[34] Johansson, B. E., Brett, I. C., *J. Biochem.* 2003, *134*, 345–352.

[35] Kelly, G., Prasannan, S., Daniell, S., Fleming, K. *et al.*, *Nat. Struct. Biol.* 1999, *6*, 313–318.

[36] Imsande, J., *J. Biol. Chem.* 1961, *236*, 1494–1497.

[37] Loessner, M. J., Scherer, S., *J. Bacteriol.* 1995, *177*, 6601–6609.

[38] Carlton, R. M., Noordman, W. H., Biswas, B., de Meester, E. D., Loessner, M. J., *Regul. Toxicol. Pharmacol.* 2005, *43*, 301–312.

