FEMS Microbiology Letters 91 (1992) 257–264

Published by Elsevier

FEMSLE 04811

Modular design of the *Enterococcus hirae* muramidase-2 and *Streptococcus faecalis* autolysin

Bernard Joris [1], Serge Englebert [1], Chien-Peng Chu [2], Reiko Kariyama [2], Lolita Daneo-Moore [2], Gerald D. Shockman [2] and Jean-Marie Ghuysen [1]

[1] *Centre d'Ingénierie des Protéines, Université de Liège, Institut de Chimie, B6, Liège, Belgium, and* [2] *Department of Microbiology and Immunology, Temple University School of Medicine, Philadelphia, U.S.A.*

Received 24 December 1991
Accepted 11 January 1992

Key words: Muramidase; *Enterococcus hirae*; *Streptococcus faecalis*; Autolysin

1. SUMMARY

The mature forms of the extracellular muramidase-2 of *Enterococcus hirae* and *Streptococcus faecalis* autolysin have very similar primary structures. Each consists of an active-site-containing N-terminal domain fused to a multiple-repeat C-terminal domain. Polypeptide segments occurring at equivalent places in these two bacterial wall lytic enzymes have homologues in two phage lysozymes and in three functionally unrelated proteins, illustrating the principle that protein molecules frequently are constructed from modules that are linked in a single polypeptide chain.

2. INTRODUCTION

The *Streptococcus faecalis* gene that codes for an autolysin of unknown specificity [1] and the *Enterococcus hirae* ATCC 9790 gene that codes for the extracellular muramidase-2 (Genbank accession number: M77639) [2,26] have been cloned and sequenced, giving access to primary structures of high similarity. Research aimed at identifying proteins which would be structurally related to the *S. faecalis* and *E. hirae* enzymes led to the conclusion that a significant extent of similarity exists, at least at the level of defined polypeptide segments, with the *Bacillus subtilis* ϕPZA lysozyme [3], the homologous *Bacillus* gene 15 ϕ29 lysozyme [4], the *Salmonella typhimurium* FlgJ flagellar protein [5], the *Listeria monocytogenes* pathogenicity-associated P60 protein [6] and *Staphylococcus aureus* protein A [7].

3. METHODS

Comparison of the primary structures of the *S. faecalis* and *E. hirae* enzymes with the protein sequence database was made using the FASTA and TFASTA algorithm of the software package of the Genetics Computer Group (GCG) (University of Wisconsin Biotechnology Centre, Madi-

Correspondence to: J.-M. Ghuysen, Centre d'Ingénierie des Protéines, Université de Liège, Institut de Chimie, B6, B-4000 Sart Tilman (Liège 1), Belgium.

son, WI) [8]. Sequences similar to a query sequence were identified using a word size of one. The proteins were compared by using the DIAGON plot procedure [9] and hydrophobic cluster analysis [10]. Homologies between aligned amino acids correspond to a threshold of 0.8 according to Gribskov and Burgess [11]. Hydrophobic cluster analysis rests upon a representation of the protein sequences on an α-helical two-dimensional pattern in which hydrophobic residues tend to form clusters that usually correspond to secondary structural elements. Clusters of similar shapes, sizes and relative positions express similarity in the polypeptide folding of the proteins.

4. RESULTS

4.1. Molecular organization of the E. hirae muramidase and S. faecalis autolysin

Intersequence comparisons show that: i) the *E. hirae* muramidase (Fig. 1a) comprises an approx. 250-amino acid N-terminal domain and an approx. 400-amino acid C-terminal domain that consists of six 45-amino acid repeats connected by five non-homologous (S and T rich) intervening sequences; ii) the *S. faecalis* autolysin (Fig. 1b) comprises an approx. 350 amino acid N-terminal domain and an approx. 250-amino acid C-terminal domain that consists of four 68 amino acid contiguous repeats; iii) the two proteins (Fig. 1c) have similar primary structures except for the presence of an approx. 114-amino acid stretch (E52-A166) in the N-terminal domain of the *S. faecalis* autolysin, that is not present in the *E. hirae* muramidase. Providing that this E52-A166 segment is excluded from the analysis, the N-terminal domains of the *E. hirae* muramidase and *S. faecalis* autolysin show almost complete identity in the shape and distribution of the hydrophobic and hydrophilic clusters along the amino acid sequences (Fig. 2a,b). The same conclusion applies to the carboxy-terminal domains of the proteins (Fig. 3a,b) except that, when compared to the *S. faecalis* autolysin, the *E. hirae* muramidase has two additional repeats. The alignments derived from the above analysis, yield 141 identities for approx. 250 aligned amino acids (55%) for the amino-terminal domains of the *E. hirae* muramidase and *S. faecalis* autolysin (Fig. 4); 26 identities for 45 aligned amino acids (54%) for the six repeats of the *E. hirae* muramidase (Fig. 5a); 57 identities for 68 aligned amino acids (84%) for the four repeats of the *S. faecalis* autolysin (Fig. 5b); and 11 identities for 45 aligned amino acids (25%) for all the repeats of the *E. hirae* muramidase and the *S. faecalis* autolysin (as derived from the data shown in Figs. 5a,b).

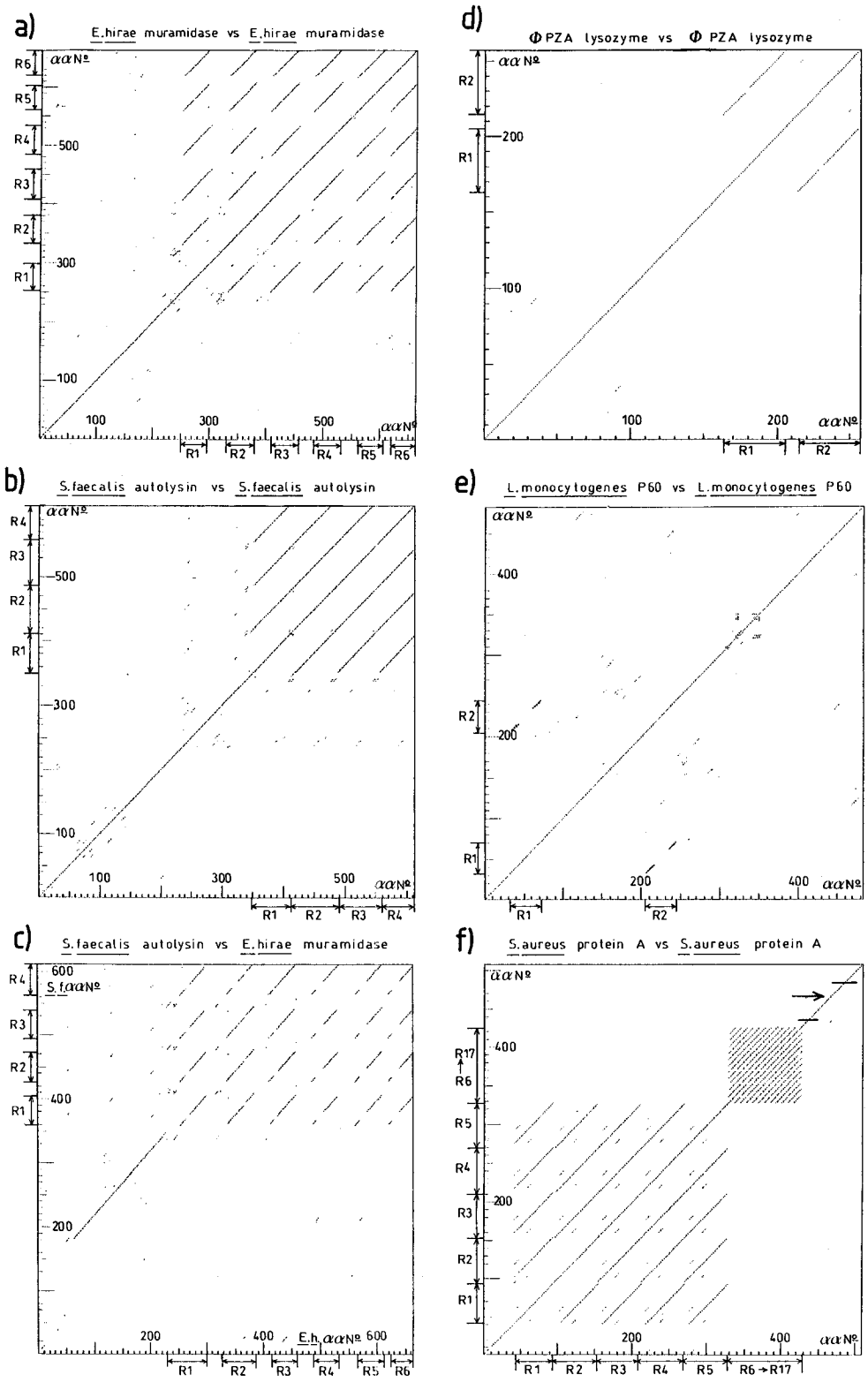
4.2. Homologue of the N-terminal domain

The only protein which showed similarity, at least in part, with the N-terminal domain of the *E. hirae* and *S. faecalis* enzymes, is the 316-amino acid FlgJ flagellar protein of *S. typhimurium*. This protein is of unknown function; it has no repeats and it may be cytoplasmic [5]. As shown in Fig. 2c, the shape and distribution of the hydrophobic and hydrophilic clusters along the carboxy-terminal half of FlgJ (from S152 to F316) closely resemble those exhibited by the T62-G230 segment of the *E. hirae* muramidase (Fig. 2b) and the P180-P350 segment of the *S. faecalis* autolysin (Fig. 2a). Alignment of the three peptide segments (Fig. 4) yields 59 identities/homologies for 160 aligned amino acids (34%).

4.3. Homologues of the C-terminal domain

Four proteins have polypeptide segments that show significant similarity with the repeats of the

Fig. 1. Comparison matrices. Intersequence comparisons of the *E. hirae* muramidase (**a**), the *S. faecalis* autolysin (**b**), the ϕPZA lysozyme (**d**), the *L. monocytogenes* P60 protein (**e**) and the *S. aureus* protein A (**f**) and between the *E. hirae* muramidase and the *S. faecalis* autolysin (**c**). In all cases, the span length is fixed to 11 amino acids and the selected threshold score is such that the probability that similarity between pairs of segments occurs by chance is lower than 1 in 1000. R1—R6: repeats. The arrow in (**f**) indicates the peptide segment which has similarity with the repeats of the *E. hirae* and *S. faecalis* enzymes.


a)
E.hirae muramidase vs E.hirae muramidase
αα№
R6
R5
R4
R3
R2
R1
500
300
100
100
300
500
αα№
R1
R2
R3
R4
R5
R6
d)
Φ PZA lysozyme vs Φ PZA lysozyme
αα№
R2
R1
200
100
100
200
αα№
R1
R2
b)
S.faecalis autolysin vs S.faecalis autolysin
αα№
R4
R3
R2
R1
500
300
100
100
300
500
αα№
R1
R2
R3
R4
e)
L.monocytogenes P60 vs L.monocytogenes P60
αα№
400
R2
200
R1
200
400
αα№
R1
R2
c)
S.faecalis autolysin vs E.hirae muramidase
600
S.faαα№
R4
R3
R2
R1
400
200
200
400
E.h.αα№ 600
R1
R2
R3
R4
R5
R6
f)
S.aureus protein A vs S.aureus protein A
αα№
R17
R6
R5
R4
R3
R2
R1
400
200
200
400
αα№
R1
R2
R3
R4
R5
R6→R17

E. hirae muramidase and *S. faecalis* autolysin: i) the 258-amino acid ϕPZA lysozyme, and its homologue, ϕ29 lysozyme, have two almost contiguous repeats in the carboxy-terminal region of the protein, from K164 to 1204 and from K215 to S258, respectively (Fig. 1d). The two repeats share 26 identities for 41 aligned amino acids (63%) (Fig. 5c); ii) the *L. monocytogenes* 432-amino acid P60 protein has two repeats in the amino-terminal half of the protein from V30 to N72 and from A204 to K246, respectively (Fig. 1e). These two repeats share 16 identities for 38 aligned amino acids (42%) (Fig. 5d) iii); the *S. aureus* 510 amino acid protein A consists of 5 contiguous large repeats spanning from D40 to D330, 11 contiguous small repeats spanning from N331 to N420 and, downstream from this cluster, a non-repeated sequence that extends from K420 to H509 (Fig. 1f, arrow). As shown in Fig. 5, the six repeats of the *E. hirae* muramidase, the four repeats of the *S. faecalis* autolysin, the two repeats of the PZA lysozyme, the two repeats of the *L. monocytogenes* P60 protein and the non-repeated V432-D474 stretch of the *S. aureus* protein A align well, yielding a consensus of 16 identities/homologies for 42 aligned amino acids (38%).

5. DISCUSSION

The *E. hirae* muramidase and *S. faecalis* autolysin have very similar primary structures except that the *S. faecalis* enzyme has an additional 114-amino acid insert in the N-terminal domain. This insert is of unknown function but could

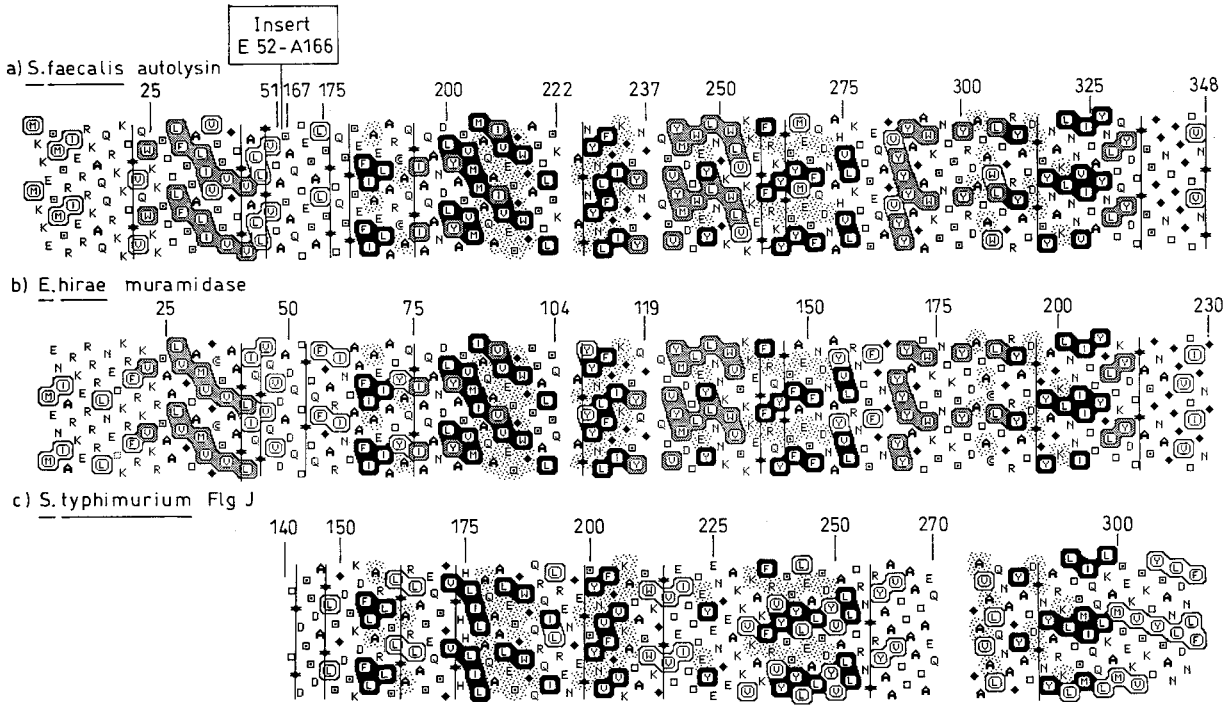


Fig. 2. Hydrophobic and hydrophilic cluster analysis of the amino-terminal domains of the *S. faecalis* autolysin (from M1 to P348) (**a**) and the *E. hirae* muramidase (from M1 to G230) (**b**), and of the carboxy-terminal region of the *S. typhimurium* FlgJ protein (from T140 to F316) (**c**). The hydrophobic residues are encircled and the hydrophobic clusters are also delineated. The hydrophobic residues and clusters conserved in the three sequences are in bold. Those common only to the *S. faecalis* and *E. hirae* proteins are shaded. The hydrophilic residues and clusters conserved in the three sequences are marked by scattered points. The one-letter code is used except for Pro: ∗; Gly: ◆; Cys: ∈; Ser: □ and Thr: □.

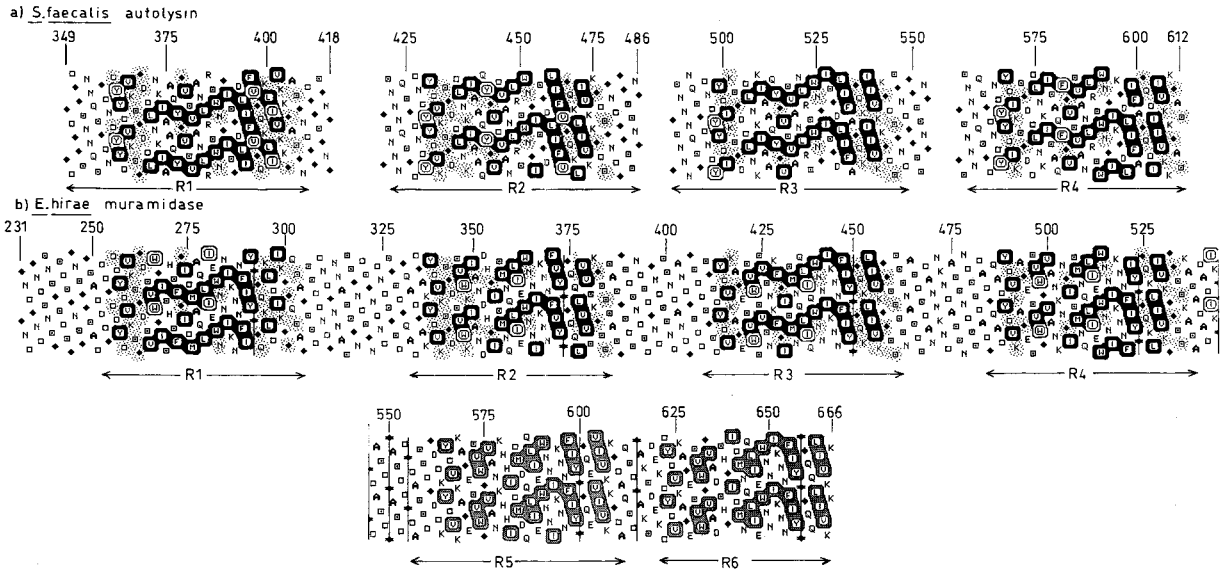


Fig. 3. Hydrophobic- and hydrophilic-cluster analysis of the carboxy-terminal repeats of the *S. faecalis* autolysin (**a**) and the *E. hirae* muramidase (**b**). The hydrophobic residues are encircled and the hydrophobic clusters are also delineated. The hydrophobic residues and clusters conserved in the two sequences are in bold. Those occurring in the repeats 5 and 6 of the *E. hirae* muramidase are shaded. The hydrophilic clusters conserved in the two sequences are marked by scattered points. Pro: ∗; Gly: ♦; Cys: ⋐; Ser: ⊡; Thr: □. R1—R6: repeats.

```
E.h.        ..MENIARKE RRRLNETKRF RKV.KRSAAL VGTAMVGCSV AAPLI.QPVQ   42
S.f.        MKKESMSRIE RRKAQQRKKT PVQWKKSTTL FSSALIVSSV GTPVALLPVT   50

E.h.        VDADQTPTQF GARIN..... ..........  ..........  ..........  61
S.f.        A----------------insert----------------

E.h.        TAAFIAEIAT YAQPIAQAND LYASVMIAQA VVESGWGSSA L...SQAPYY  108
S.f.        PSEFIAELAR CAQPIAQAND LYASVMMAQA IVESGWGAST L...SKAPNY  226
Flg.J       SKDFLARLSL PARLASEQSG VPHHLILAQA ALESGWGQRQ ILRENGEPSY  201
Consensus   ---FIA-I-- -A-------- L---L--AQA -LESGWG--- I------P-Y
                L  L               V   V       V          L

E.h.        NLFGIK..GS YQGQTVYMDT LEYLNNKWVS KKEPFRQYPS FAESFNDNAY  156
S.f.        NLFGIK..GS YNGQSVYMDT WEYLNGKWLV KKEPFRKYPS YMESFQDNAH  274
Flg.J       NVFGVKATAS WKGPVTEITT TEYENGEAKK VKAKFRVYSS YLEALSDYVA  251
Consensus   NLFGIK---S W-G------T -EY-N----- -K--FR-Y-S F-E-F-D---
             V  V       Y                              Y   L

E.h.        VLRNTSFGNG .YYYAGTWKS NTKSYTDATA CLTGRYATDP GYAGKLNNII  205
S.f.        VLKTTSFQAG VYYYAGAWKS NTSSYRDATA WLTGRYATDP SYNAKLNNVI  324
Flg.J       LLTRNP.... ..RYAAVTTA ATAE.QGAVA LQNAGYATDP NYARKLTSMI  294
Consensus   LL-------- ---YA----- -T-----A-A -----YATDP -Y--KL---I
            V

E.h.        TTYGLTKYDT PASGNAGGGV TIGNGGNTGN TSNSGSTSGN SGGSA       250
S.f.        TAYNLTQYDT PSSGGNTGGG TV.NPGTGGS NNQSG.....  .....      358
Flg.J       QQ..LKAMSE KVSKTYSANL DNLF                              316
Consensus   ----L----- --S------- ---------- ---------- -----
```

Fig. 4. Amino acid sequence alignment of the amino terminal domain of the *E. hirae* muramidase (E.h), the amino-terminal domain of the *S. faecalis* autolysin (S.f) and the carboxy-terminal region of the *S. typhimurium* FlgJ protein. The insert E52-A166 of the *S. faecalis* autolysin is excluded from the alignment. Identities occurring along the sequences of the *E. hirae* and *S. faecalis* proteins are marked by black circles (●). The consensus shows the identities and homologies occurring along the sequences of all three proteins.

possibly be a pre-pro-peptide. Given the high extent of similarity in the structure, one may safely conclude that the *E. hirae* and *S. faecalis* peptidoglycan hydrolases have similar specificities and catalytic mechanisms. In all likelihood, their active centre is located in the N-terminal domain and results from the folding of that approx. 200-amino acid polypeptide stretch which shows high similarity in the two enzymes.

Given that the *E. hirae* enzyme (and by extension the *S. faecalis* enzyme) is a muramidase, one would expect that aspartic/glutamic acids are present in the active site. In fact, the N-terminal domain contains several acidic residues which are separated by 13–33 amino acids consistent with similar locations in other lysozymes. Thus E128 of the *E. hirae* enzyme is located 25 amino acids upstream from D153, similar to the 27-amino acid separation of these residues in the putative active site of the *Chalaropsis* (D6-E33) [12] and bacteriophage CPI-1 (D10-E37) muramidases [13–15]. Also, the pairs D50-E68, D125-E139, and E139-D153 in the *E. hirae* enzyme are spaced similarly to the pair E35-D52 in hen egg-white lysozyme [16]. However, the active site containing N-terminal domains of the *E. hirae* and *S. faecalis* enzymes lack overall sequence relatedness with other glycosidases, in particular muramidases (including the *Bacillus* ϕPZA and ϕ29 lysozymes) of known primary structure. Lysozymes exist which have a similar three-dimensional structure despite the absence of amino acid sequence similarity [17].

The *E. hirae* muramidase possesses several amino acid groupings (the tetrad S175NTK, the triad SGN at positions S218, S243, S317 and S320, and the triad KSG at positions K260 and K341) which, at first sight, can be regarded as equivalents of the characteristic motifs found in the serine β-lactamases and penicillin-binding proteins [18]. However, the *S. faecalis* autolysin lacks the SXXK active-site motif. The shape and distribution of the hydrophobic- and hydrophilic-clusters along the amino acid sequences of both *E. hirae* and *S. faecalis* enzymes differ from those found in the serine β-lactamases and penicillin-binding proteins [19,20]. Evidence has been presented that the *E. hirae* muramidase binds penicillin with low affinity [21]. On the basis of the data presented here, it may well be that this

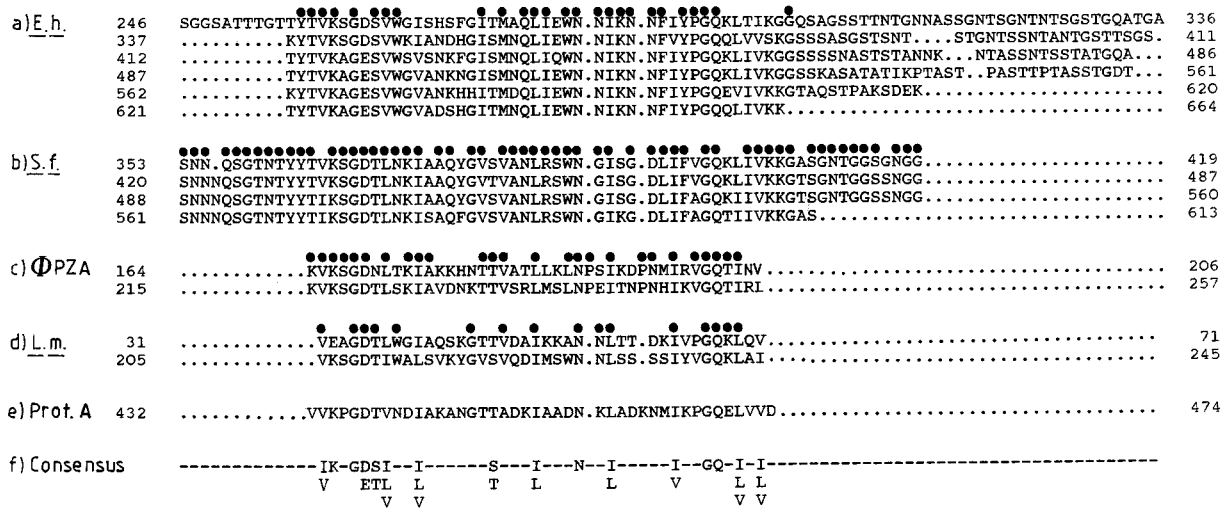


Fig. 5. Amino acid sequence alignment of the carboxy terminal repeats of the *E. hirae* muramidase (**a**); the carboxy-terminal repeats of the *S. faecalis* autolysin (**b**); the carboxy-terminal repeats of the ϕPZA lysozyme (**c**); the internal repeats of the *L. monocytogenes* P60 protein (**d**); and the V432-D474 segment of the carboxy-terminal region of the *S. aureus* protein A (**e**). Identities occurring in groups (**a**), (**b**), (**c**) and (**d**), respectively, are marked by black circles (●). The consensus derived from the 15 aligned peptide segments is shown in (**f**).

protein binds penicillin by a mechanism different from penicilloyl transfer involving an 'active' serine residue.

Based on the principle that protein molecules are frequently constructed from modules and that each of these modules folds into one domain that performs a particular function, one may hypothesize that the carboxy-terminal domain of the *S. typhimurium* flagellar FlgJ protein, which is similar to the active site of the *E. hirae* and *S. faecalis* enzymes, might function as a glycosidase of one kind or another, perhaps a muramidase.

Downstream from the catalytic amino-terminal domain, the *E. hirae* and *S. faecalis* enzymes have carboxy-terminal extended repeat structures. These repeats may be responsible for substrate, i.e. peptidoglycan, binding. They are similar to those of the *Bacillus* ϕPZA and ϕ29 lysozymes, with one segment of protein A and with the two repeats that occur at internal positions in the *L. monocytogenes* P60 protein. Spontaneous mutants of *L. monocytogenes* impaired in the synthesis of P60 lose the ability to invade phagocytic cells and the bacteria grow in long cell chains. The chains disaggregate after treatment with P60 and invasiveness of the treated mutant cells is restored [6]. Peptidoglycan hydrolase activities have been associated with cell separation in other bacterial systems (summarized in ref. 22). The repeats described here may form a particular family of ligand-binding sequences. They do not exhibit similarity to the repeats of: i) the carbohydrate-binding proteins reviewed by Wren [23]; ii) the surface proteins from Gram-positive cocci reviewed by Fischetti et al. [24]; and iii) the pneumococcal peptidoglycan amidase and CPL-1, CPL-7 and CPL-9 muramidases [25]. The repeats of the pneumococcal enzymes appear to participate in the recognition of moieties in the pneumococcal wall [15].

ACKNOWLEDGEMENTS

The work in Liège was supported, in part, by the Belgian programme on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming (PAI No. 19), an Action concertée with the Belgian Government (convention 86/91-90), the Fonds de la Recherche Scientifique Médicale (contract No. 3.4537.88), and a Convention tripartite between the Région wallonne, SmithKline Beecham, U.K., and the University of Liège. B.J. is Chercheur qualifié of the Fonds National de la Recherche Scientifique (FNRS, Brussels). The work at Temple was supported by Public Health Service Grant AI05044 from the National Institute of Allergy and Infectious Diseases.

REFERENCES

[1] Beliveau C., Potvin C., Trudel J., Asselin A. and Bellemare G. (1991) J. Bacteriol. 173, 5619–5623.
[2] Kawamura, T. and Shockman, G.D. (1983). FEMS Microbiol. Lett. 19, 65–69.
[3] Paces V., Uleek C. and Urbanek P. (1986) Gene 44, 107–114.
[4] Garvey K.J., Sardi M.S. and Ito J. (1986) Nucleic Acids Res. 14, 10001–10008.
[5] Jones C.J., Homma M. and Macnab R.M. (1989) J. Bacteriol. 171, 3890–3900.
[6] Köhler S., Leimeister-Wächter M., Chakrabarty T., Lottspeich F. and Goebel W. (1990) Infect. Immun. 58, 1943–1950.
[7] Uhlen M., Guss B., Nilsson B., Gatenbeck S., Philipson L. and Lindberg M. (1984) J. Biol. Chem. 259, 1695–1702.
[8] Devereux J., Haeberli P. and Smithies O. (1984) Nucleic Acids Res. 12, 387–395.
[9] Staden R. (1982) Nucleic Acids Res. 10, 2951–2961.
[10] Gaboriaud C., Bissery V., Bencherett T. and Mornon J.P. (1987) FEBS Lett. 224, 149–155.
[11] Gribskov M. and Burgess R.R. (1984) Nucleic Acids Res. 14, 6745–6763.
[12] Fouche P.B. and Hash J.H. (1978) J. Biol. Chem. 253, 6787–6793.
[13] Carcia E., Garcia J.L., Garcia P., Arraras A., Sanchez-Puelles and Lopez R. (1988) Proc. Natl. Acad. Sci. USA 85, 914–918.
[14] Diaz E., Lopez R. and Garcia L.L. (1990) Proc. Natl. Acad. Sci. USA 87, 8125–8129.
[15] Sanz J.M. and Garcia J.L. (1990) Eur. J. Biochem. 187, 409–416.
[16] Yamada H., Imoto T. and Noshita S. (1982) Biochemistry 21, 2189–2192.
[17] McKenzie H.A. and White F.H. (1991) Adv. Protein Chem. 41, 174–258.
[18] Ghuysen J.M. (1991) Annu. Rev. Microbiol. 45, 37–67.
[19] Palomeque-Messia, P., Englebert S., Leyh-Bouille M., Nguyen-Distèche M., Duez C., Houba S., Dideberg O.,

Van Beeumen J. and Ghuysen J.M. (1991) Biochem. J. 279, 223–230.
[20] Granier B., Duez C., Lepage S., Englebert S., Dusart J., Dideberg O., Van Beeumen J., Frère J.M. and Ghuysen J.M. (1992) Biochem. J. (in press).
[21] Dollinger D.L., Daneo-Moore, L. and Shockman G.D. (1989) J. Bacteriol. 171, 4355–4361.
[22] Rogers, H.J., Perkins, H.R. and Ward, J.B. (1980) Microbial Cell Walls and Membranes, pp. 448–449. Chapman and Hall, London, New York.
[23] Wren B.W. (1991) Mol. Microbiol. 5, 797–803.
[24] Fischetti V.A., Pancholi V. and Schneewind O. (1991) In: Genetics and Molecular Biology of Streptococci and Enterococci. (Dunny G.M., Cleary P.P. and McKey L.L., (Eds.), pp. 290–294. Society for Microbiology, Washington, DC.
[25] Garcia P., Garcia J.L., Garcia E., Sanchez-Puelles, J.M. and Lopez R. (1990) Gene 86, 81–88.
[26] Cho, C.P., Kariyama, R., Daneo-Moore, L. and Shockman, G.D. (1992) J. Bacteriol. (in press).