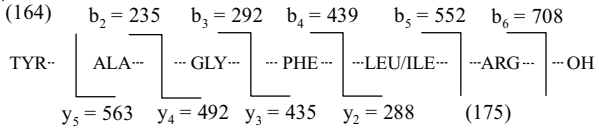
*Chemistry of Natural Compounds, Vol. 47, No. 5, November, 2011 [Russian original No. 5, September-October, 2011]*

# PEPTIDE FROM MARINE NEMATODES, A STRUCTURAL ANALOG OF DALARGIN

**L. A. Elyakova, B. V. Vas′kovskii, and S. I. Vantseva*** UDC 577.112.82.017

Our previous work [1] included a table where the inhibitor-active (IA) fractions from this same polychete of tropical marine *Eunicidae* (collected in July in the Seychelles Islands, Indian Ocean) contained peptides for which the structures were established previously [2] and a peptide with *m*/*z* 727 (e.g., in fractions 3/10 and 4/7 [1]). We performed an additional HPLC stage on a Phenomenex column (Luna 3 micron C-18, 50 × 3 mm) using 10–65% $CH_3CN$ in buffer A (0.1% TFA) in order to isolate the peptide with *m*/*z* 727 (henceforth m727). Peptide m727 eluted with $\tau \sim 10$ min, was clearly separated from other previously described peptides, and was easily detected because of the presence of an aromatic amino acid (AA) that was tyrosine according to the UV spectrum. We used a variety of usual methods to establish its structure. Mass fragmentation of the peptide was successful is assigning the structure as Tyr-Ala-Gly-Phe-Leu/Ile-Arg:




series b (MW of residue or their sum – 17) ($–O_1H_1$);
series y (MW of residue – $H_1$) total MW of peptide amino acids from the C-terminus.

Sequencing refined the structure as Tyr-Ala-Gly-Phe-**Leu**-Arg, i.e., the peptide had an AA sequence corresponding to commercial dalargin (D-Ala-), which exhibits opiate activity. A determination of IA for a sample of commercial dalargin (Sigma) relative to HIV integrase gave $IC_{50}$ $2.5 \cdot 10^{-4}$ M. IA relative to HIV integrase for the natural sample of m727 had a similar value ($1 \times 10^{-4}$ M).

A search in the Entrez system provided by NCBI among protein sequences for the nematode *C. elegans*, the gene structure of which is fully established, showed a double repetition of five of these six AA, namely, (433)-Tyr-Ala-Gly-Phe-Leu-(437) and (205)-Ala-Gly-Phe-Leu-Arg-(209) [3] (Table 1). A search among structures of microorganisms found the peptide YAGFLR in seven sources, among which were vectors of hazardous diseases such as plague (placed No. 1 on the list of quarantine diseases), leprosy, and Far-East (FE) scarlet-like fever [4]. Table 1 presents the data. The peptide YAGFLR is included in sequences of 238 AA and is fully replicated in protein structures of both plague and FE scarlet-like fever (Fig. 1). This same YAGFLR fragment occurred in a different place for leprosy.

| | | | | | |
|---|---|---|---|---|---|
| 1 | MPPAMAEVSD | KLQLSHKVYA | HDYQAFWLWS | GVNPQPALQQ | ANQVYLHQGE |
| 51 | VVIRQRAAWF | QKMGLPSSRL | TLPAMWVTVR | ITTLDVPDDI | LAILIDLPRR |
| 101 | WAAAGNQVIG | LQIDFDAGTY | RLDD**YAGFLR** | RVRTKLDPNF | ALGVTGLLDW |
| 151 | AKTGSIQQLN | ALPIDELVIQ | TYQGRSTVNQ | YSRYLPALLQ | LRLPFKIGLV |
| 201 | QHGEWDPQWE | QYLAASPFYR | GEVVFLLNHL | RSEPANGK | |

Fig. 1. Partial amino-acid sequence for portions of protein structures of plague and FE scarlet-like fever. Numbers on the left are numbers of residues from the amine terminuses of the compared portions. Identical AA residues are shown in bold.

M. M. Shemyakin and Yu. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117997, GSP, Moscow, B-437, fax: (495) 330 56 92, e-mail: svetvan@yandex.ru. Translated from Khimiya Prirodnykh Soedinenii, No. 5, September–October, 2011, pp. 754–755. Original article submitted March 10, 2011.

 0009-3130/11/4705-0864 

TABLE 1. Portions of Primary Structures of Proteins Homologous to Our Peptide Tyr-Ala-Gly-Phe-Leu-Arg

| Organism | Protein fragment* | Amino-acid sequence |
|---|---|---|
| Nematode *Caenorhabditis elegans* | P34315 NCX6 [3] 433-437 | YAGFL |
| Nematode *C. elegans* | P24892 NU4M [3] 205-209 | AGFLR |
| *Yersinia pseudotuberculosis Protobacteria* (FE scarlet-like fever) | 3023438FGH [4] 125-130 | YAGFLR |
| *Yersinia pestis* (plague) | 3211395FGE [4] 142-147 | YAGFLR |
| *Mycobacterium leprae* (leprosy) | 2323363ADL [4] 265-270 | YAGFLR |

*Accession number in these databases is given.

TABLE 2. Elution Times (τ, min) of These Compounds*

| IX<br>Without D-amino acids | XI<br>D-Tyr, D-Ala | VI<br>D-Ala, commerc. dalargin | XXVII<br>D-Ala, D-Arg | XVII<br>D-Ala, D-Phe | XXIII<br>D-Ala, D-Leu |
|---|---|---|---|---|---|
| | | Literature data [5] | | | |
| 40 | 40.7 | 41.2 | 42.8 | 46.4 | 48 |
| | | Our data (HPLC) | | | |
| 21.37 | – | 22.03 | 22.45 | 23.2 | – |

*Natural peptide with m727 and τ = 21.57 min.

We knew of an additional difficulty in working with peptides of natural origin, especially with peptides of marine invertebrates. This is the high probability of encountering their optical isomers with D-amino acids. Therefore, we consulted the discoverers of dalargin in the Russian Cardiology Center, Academy of Medical Sciences. They had compared dalargin with analogs synthesized by them [5] and graciously supplied three samples: (IX) Tyr-Ala-Gly-Phe-Leu-Arg; (XVII) Tyr-**DAla**-Gly-**DPhe**-Leu-Arg (**D-Ala**, **D-Phe**); (XXVII) Tyr-**DAla**-Gly-Phe-Leu-**DArg** (**D-Ala**, **D-Arg**) (retaining the previous numbering [5]).

All compounds found by them in the published HPLC trace eluted in the order : IX, XI, VI, XXVII, XVII, and XXIII with the retention times τ given in Table 2. Under our conditions (Nova Pack C-18 column; 4.6 × 150; 4 μm) the elution order of the compounds was preserved. Sample IX had the shortest τ. Then, samples VI, XXVII, and XVII followed. The natural sample m727 eluted with τ = 21.57 min, i.e., closest to sample IX. This was probably YAGFLR. Table 2 presents all known literature data [5] and our results.

**ACKNOWLEDGMENT**

The work was supported by the ISTC Project 3197p (01.04.06-31.03.08). We thank Doc. Med. Sci. Zh. D. Bespalova (Russian Cardiology Center, AMS) for supplying samples of control compounds and staff members of the V. N. Orekhovich Biomedical Chemistry Research Institute, RAMS, Cand. Biol. Sci. V. G. Zgoda and Cand. Biol. Sci. A. T. Kopylov for performing the mass fragmentation.

**REFERENCES**

1. L. A. Elyakova, B. V. Vas′kovskii, N. I. Khoroshilova, S. I. Vantseva, and Yu. Yu. Agapkina, *Bioorg. Khim.*, **37**, No. 2, 233 (2011).
2. L. A. Elyakova, B. V. Vas′kovskii, E. V. Bocharov, V. L. Tunitskaya, S. N. Kochetkov, and G. B. Elyakov, RF Pat. No. 2,314,818 (2008); *Byull.* No. 2 (2008).
3. BLAST (USA), www.ncbi.nlm.nih.gov/blast/
4. Protein Research Foundation (Japan), http://www.prf.or.jp/seqdb-e.html
5. Zh. D. Bespalova, N. F. Sepetov, O. L. Isakova, M. I. Titov, B. L. Pekelis, A. S. Molokoedov, and V. N. Bushuev, *Bioorg. Khim.*, **13**, No. 2, 160 (1987).