Feast/famine regulatory proteins of a multiple drug resistant, *Pseudomonas aeruginosa*

By Masashi SUZUKI,^{*)} Hironori ARAMAKI,^{**)} and Hideaki KOIKE^{*)}

(Communicated by Masanori OTSUKA, M. J. A., Oct. 14, 2003)

Abstract: The eubacterium *Pseudomonas aeruginosa* is an opportunistic pathogen, and at the same, possesses a high ability to resist antibiotics and disinfectants. Using the genomic sequence of its standard strain PAO1, determined by another group, genes coding feast/famine regulatory proteins (FFRPs) have been identified. In general, FFRPs regulate transcription of genes, thereby controlling metabolism, growth and environmental adaptation of bacteria, and thus they are obvious targets when designing drugs in order to irradiate *P. aeruginosa*. In the light of findings on FFRPs of other bacteria, amino acid residues of FFRPs of *P. aeruginosa* that are likely to be involved in DNA recognition or interaction with natural ligands have been identified. One of the FFRPs, Pa5977610, has been identified as orthologous to an *E. coli* FFRP, the leucine-responsive regulatory protein (Lrp). Another FFRP, Pa2220251, resembles another *E. coli* FFRP, YbaO, to a lesser extent. No FFRP of *P. aeruginosa* is orthologous to the third *E. coli* FFRP, AsnC. These differences among *P. aeruginosa* FFRPs can be used for selecting appropriate target proteins, thereby minimizing potential interaction of drugs to be developed with the human symbiont *E. coli*.

Key words: Cystic fibrosis; drug design; FFRP; inside-hospital infection; pathogen; quorum sensing.

Introduction. The gram-negative eubacterium *Pseudomonas aeruginosa* has a high ability to cause diseases to susceptible patients¹⁾⁻⁶⁾ (see <http://www.pseudomonas.com/>). Patients suffering from cancer, burns or cystic fibrosis, or those who are deficient in their immune systems are often victimized by a variety of toxic substances produced by this organism, and the results can be lethal. This opportunistic pathogen easily acquires resistance to antibiotics and disinfectants by modifying internal genes or by adopting external genes, and causes outbreaks of serious epidemics even inside hospitals. With the hope of terminating such health hazards, the genomic sequence of the standard PAO1 strain of *P. aeruginosa* was determined.⁶⁾

Many pathogens are parasitic, having lost their adaptability to environmental changes, and able to survive only inside their hosts. However, *P. aeruginosa* is

fairly ubiquitous, adapting to various ecological niches from water and soil to plant and animal tissues.¹⁾⁻⁶⁾ It is believed that this remarkable adaptability originates in its efficient gene-regulatory mechanisms, although details remain unknown. Using its efficient transcription regulation, *P. aeruginosa* even responds to signals from other bacteria of the same species, similarly to intercellular communications inside eukaryotes (i.e. the quorum sensing).⁵⁾

In general, FFRPs regulate transcription of genes thereby controlling metabolism, growth and environmental adaptation of organisms. Thus, although no direct evidence is established, it is possible that FFRPs directly regulate pathogenic functions of *P. aeruginosa*. In fact, the number of FFRPs coded in the genome of *P. aeruginosa* is the largest among eubacteria so far examined, and in other words, many of these FFRPs are unseen in other bacteria.⁷⁾ This fact may well indicate the importance of these proteins. In any case, however, FFRPs are potential targets for designing drugs in order to irradiate this pathogen, since disruption of ordinary metabolic functions through FFRPs will cause enough damage to *P. aeruginosa*.

Up to now, no FFRP has been found coded in eukaryotic genomes, of nuclei, mitochondria or plastids.

*) National Institute of Advanced Industrial Science and Technology (AIST), AIST Tsukuba Center 6-10, 1-1-1, Higashi, Tsukuba, Ibaraki 305-8566, Japan.

**) Department of Molecular Biology, Daiichi College of Pharmaceutical Sciences, 22-1, Tamagawa-cho, Minami-ku, Fukuoka 815-8511, Japan.

†) Correspondence to M. Suzuki at AIST.

Fig 1. Amino acid sequences of FFRPs coded in the genomes of *P. aeruginosa*, *E. coli*, *P. OT3*, and *T. volcanium*. Four of these entries, pot0300646, pot1216151, TVG0307586, pot0175330, are demi-FFRPs, each of which corresponds to the N-terminal half of a full length FFRP. The secondary structural elements (i.e. α -helices 1-3, 11, and 12, β -strands 1 and 11-14) were assigned using crystal structures. The hydrophobic phases inside the α -helices are marked with asterisks (*), with types of residues, Ile, Leu, Val, and Met, found there highlighted in bold. At each position outside the α -helices, the type of residues most frequently found is highlighted in bold. The amino acid sequences of *E. coli* Lrp, YbaO, and AsnC are underlined.

followed by the stop codon positions of the corresponding genes in the genomic sequences (see the NCBI database, <http://www.ncbi.nlm.nih.gov/entrez/>). Archaeal FFRPs are identified using our ARCHAIC code

Fig. 2. Phylogenetic relation of FFRPs (a) and their amino acid residues possibly important for interaction with DNA or ligands (b). (a) An unrooted phylogenetic tree made using the alignment in Fig. 1. Nodes characterized with bootstrap values, 500 or higher, are colored in red. Nodes characterized with smaller bootstrap values but 300 or higher are indicated by closed circles in black. The four demi-FFRPs assembled into a group are highlighted in green. Two other subgroups identified with relatively high bootstrap values are highlighted in red and blue, respectively. The three *E. coli* FFRPs are underlined. In Fig. 1 these FFRPs are aligned by circling around this subfigure counter-clockwise, starting from the point as is indicated by the arrow. (b) Some positions extracted from the alignment in Fig. 1. Left: Amino acid residues positioned inside the DNA recognition helices (i.e. α -helices 3) out of their hydrophobic phases. These are the candidates for residues contacting DNA bases. Right: Residues found at positions, from which *E. coli* Lrp interact with leucine. Green is used for highlighting FFRPs of *P. aeruginosa*, while red for *E. coli* FFRPs. Particular types of residues shared by the closest groups of FFRPs at the same positions are indicated using the same colors. We have identified Pa2472442 as an ortholog of BkdR (the branched-chain keto acid dehydrogenase regulator)²⁴ of *Pseudomonas putida*.

(<http://www.aist.go.jp/RIODB/ARCHAIC/>).

Identification of FFRPs from *P. aeruginosa*.

In the *E. coli* genome⁸⁾ three FFRPs are coded.⁷⁾ Two of them, Lrp (the leucine-responsive regulatory protein, Ec0932312),^{9),10)} and AsnC (the asparagine synthesis *C* gene product, Ec3924173),¹¹⁾ are the best characterized FFRPs. These activate or repress transcription of a number of genes, in many cases depending on extra-cellular leucine or asparagine, respectively. The function of the third *E. coli* FFRP, YbaO (Ec0468065), is not known, but its amino acid sequence resembles that of another FFRP,¹²⁾ Grp (the glutamate uptake regulatory protein), which regulates the uptake of glutamate and glutamine in another bacterium.¹³⁾

By performing a homology search using amino acid sequences of these *E. coli* FFRPs and FFRPs of two archaea, *Pyrococcus* sp. OT3 and *Thermoplasma volcanium* as references (listed in refs. 7) and 14)), 8 FFRPs have been identified as coded in the genome⁶⁾ of *P. aeruginosa* strain PAO1: Pa2220251, Pa2291589, Pa2472442, Pa2914358, Pa4445486, Pa5047278, Pa5372266, Pa5977610. This number eight is the largest among eubacteria so far examined.⁷⁾

During this sequence analysis we have noticed that some other open reading frames, including pot0258936, pot0112597, TVG1466300, TVG1497808, TVG0198521, coded for proteins which were closely related with FFRPs. Their N-terminal halves are essentially the same as the DNA-binding domains of FFRPs, but the C-terminal halves are longer, possibly composing different types of assembly domains. Next to these, Pa0604912 has been found to be close to FFRPs.

Phylogenetic relation of FFRPs. The amino acid sequences of FFRPs have been aligned (Fig. 1), using the CLUSTALW program¹⁴⁾ and the PHYLIP package (<http://evolution.genetics.washington.edu/phylip.html>), followed by, as before,¹⁵⁾ manual improvements with focuses on characteristics expected to form secondary structural elements in the same way as in the three crystal structures of FFRPs.^{7),16),17)} In Fig. 1 hydrophobic residues regularly positioned inside the predicted α -helices are highlighted in bold. For the positions outside the α -helices the most frequent types are highlighted in the same way, showing the overall consistency of this alignment.

By applying tools in the PHYLIP package to the sequence alignment (Fig. 1), an unrooted phylogenetic tree has been made (Fig. 2a). A bootstrap value was calculated for each node in order to evaluate the reliability of the tree. This value defines the number of times that

Fig. 3. A ribbon diagram of a demi-FFRP from *P. OT3* (pot1216151), with the 7 positions genetically identified as important for leucine-binding colored in blue.

the same diversification at the node is observed out of 1,000 trials, while effectively changing the weight of each amino acid position in the sequence alignment. Bootstrap values of 500 or higher characterize the nodes as reliable (highlighted in red in Fig. 2a). Nodes characterized with smaller bootstrap values, 350 or higher, are indicated by closed circles in black (Fig. 2a).

Two of the three *E. coli* FFRPs are related with FFRPs of *P. aeruginosa* by relatively high bootstrap values. *E. coli* Lrp (Ec0932312) is closely related to Pa5977610; the bootstrap value obtained for the node separating the two is 992. *E. coli* YbaO (Ec0468065) is related to Pa2220251, although the corresponding bootstrap value, 422, is not so high. Pa2914358 and Pa5372266, in the blue subgroup in Fig. 2a, deviate furthest from the two *E. coli* FFRPs. No FFRP of *P. aeruginosa* closely relates with *E. coli* AsnC (Ec3924173).

Positions of FFRPs possibly important for interaction with DNA bases. The DNA-binding specificity of an FFRP depends on its type (e.g., Lrp or AsnC).¹⁸⁾ It is highly likely that α -helix 3 is the major component for contacting DNA bases, thereby determining the binding specificity (the DNA recognition helix).^{15),19)} Typically, amino acid residues inside a DNA recognition helix can be classified into three types²⁰⁾⁻²²⁾: those binding the sugar-phosphate backbones of the DNA, thereby fixing the binding geometry of the α -helix

Amino acid positions	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	
	1	1	2	4	4	6	7	8	8	9	9	1	5	7	0	2	4	4	5	5	5	7	9	9	4	5	8	9	6	0
	6	9	0	4	5	6	3	3	5	3	5	3	5	6	2	6	8	9	0	2	3	7	6	7	8	7	4	1	9	0
Ⓐ Facing the pocket	○	○	○	○	○	○	○	○	○	○	○					○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Ⓑ Surrounding <i>E.coli</i> ligand													○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○		
Ⓒ Leucine-binding																										○	○	○	○	○

Fig. 4. Positions possibly important for interaction with types of ligands.

(type I), those important for stabilizing the protein domain, therefore facing away from the DNA into the hydrophobic core of the protein, thereby limiting orientation of the helix relative to the DNA (type II), and finally, as the sum effects of types I and II, those facing into the DNA major groove thereby directly contacting DNA basepairs (type III).

With the known 3D structures of FFRPs,^{7,16,17} all crystallized without DNA, currently we are only able to identify positions 73, 77, 80 as being of type II. The remaining positions, 70-72, 74-76, 78-79, 81-82 are all candidates for type III (Fig. 2b left). Resemblances and differences of residues occupying these positions are consistent with the phylogenetic tree. Namely, *E. coli* Lrp and Pa5977610 share the same types at 8 out of the 10 positions, suggesting the same DNA-binding specificity. YbaO (Ec0468065) and Pa2220251 share the same types at 4 positions. Pa2914358 and Pa5372266 are most different from the two *E. coli* proteins. No FFRP of *P. aeruginosa* shares the same types at more than 2 positions with AsnC (Ec3924173).

Positions of FFRPs possibly important for interaction with ligands. Generally, FFRPs interact with small molecules possessing the size of amino acids. A pocket in a "+"-like shape is created in the center of an FFRP assembly (Fig. 3), and it appears that some residues surrounding this pocket interact with such ligands.

Interaction of *E. coli* Lrp and leucine was studied genetically, and by mutating at each of 7 positions near the pocket, 147, 154, 169, 181, 199, 200, and 201 (defined here as type C positions and shown in Fig. 4), leucine-dependence of the transcription regulation by Lrp was affected.²³ At 6 positions out of these 7, *E. coli* Lrp and Pa5977610 share the same types of residues, and the 7th position 200 is occupied by very similar residues, Ile or Val (Fig. 2b right), suggesting the same leucine-interaction of the two proteins.

A complication added here is that, although the type C positions are important for interaction with leucine, other positions might be involved in interactions with ligands other than leucine. While trying to crystallize an archaeal FFRP, pot1216151, we expressed this protein using an *E. coli* system.¹⁷ Consequently, we have found that this protein was crystallized, in fact, in complex with an unidentified type of ligands, most likely present inside *E. coli* cells (Koike *et al.*, in preparation). In an improved electron density map, two moieties are identified, which are not part of the protein but are larger than solvent molecules. The size of each moiety is similar to that of valine, but no amino acid perfectly fits. By best fitting a valine molecule into each moiety, 18 positions, here defined as of type B, have been identified, whose C, N, S or O atoms are positioned at distances 6 Å or shorter from the same types of atoms of the modeled valine molecule. Some positions are shared by types B and C (Fig. 4), but there are positions specific to each type, suggesting partial overlapping of the two sites.

We therefore have identified the population of residues (here defined as type A and shown in Fig. 4) creating the whole pocket between dimers of pot1216151. These residues cannot be contacted by a probe of radius 5.0 Å, but can be contacted by another probe of radius 1.4 Å to some degree. The criteria used here are accessibility being 5% or more than that when the same types of residues are fully stretched between a pair of glycine residues. In the crystal structure of pot1216151, 4 sets of dimer-interfaces are crystallographically independent, and type A residues are those satisfying the above criteria at least in 2 of the 4 sets.

Positions of FFRPs of *P. aeruginosa*, which interact with ligands other than leucine may well be found inside type A. These would be the obvious sites for targeting drugs.

Conclusion: FFRPs and their amino acid residues possibly targeted. In the genome of *P.*

aeruginosa strain PAO1, 8 FFRPs are coded. One of them, Pa5977610 is identified as an ortholog of *E. coli* Lrp, while Pa2220251 resembles YbaO to some extent. No FFRP of *P. aeruginosa* is orthologous to *E. coli* AsnC. A drug able to irradiate *P. aeruginosa* should better have minimal interaction with the human symbiont *E. coli*. Thus, 6 remaining FFRPs of *P. aeruginosa*, Pa2914358, Pa5372266, Pa2472442, Pa2291589, Pa5047278, Pa4445486, or possibly Pa2220251 can be targets for such drug design. Natural ligands interact with the pocket which is created between two FFRP dimers by a pair of ~30 amino acid residues, and these have been identified as potential sites for interaction with drugs to be developed.

Acknowledgements. This work was supported by the Core Research for Evolutional Science and Technology (CREST) program of the Japan Science and Technology Corporation (JST), and the Mitsubishi Foundation. We thank Dr. Lester Clowney for his critical reading of the manuscript.

References

- 1) Livermore, D. M. (2002) Clin. Infect. Dis. **34**, 634-640.
- 2) Davis, J. C. (2002) Paediatr. Respir. Rev. **3**, 128-134.
- 3) Lambert, P. A. (2002) J. R. Soc. Med. **41** (Suppl.), 22-26.
- 4) Chunha, B. A. (2002) Semin. Respir. Infect. **17**, 250-258.
- 5) Dunny, G., and Leonard, B. A. B. (1997) Ann. Rev. Microbiol. **51**, 527-564.
- 6) Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrener, P. *et al.* (2000) Nature **406**, 959-964.
- 7) Koike, H., Sakuma, M., Mikami, A., Amano, N., and Suzuki, M. (2003) Proc. Japan Acad. **79B**, 63-69.
- 8) Blattner, F. R., Plunkett III, G., Bloch, C. A., Perna, N. T., Burland, V. *et al.* (1997) Science **277**, 1453-1474.
- 9) Calvo, J. M., and Matthews, R. G. (1994) Microbiol. Rev. **58**, 466-490.
- 10) Newman, E. B., and Lin, R. (1995) Annu. Rev. Microbiol. **49**, 747-775.
- 11) Kölling, R., and Lotter, H. (1985) J. Bacteriol. **164**, 310-315.
- 12) Zeng, Q., and Summers, A. O. (1997) Mol. Microbiol. **24**, 229-232.
- 13) Peekhaus, N., Tolner, B. P., and Krämer, R. (1995) J. Bacteriol. **177**, 3934-3940.
- 14) Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucl. Acids Res. **22**, 4673-4680.
- 15) Suzuki, M., Amano, N., and Koike, H. (2003) Proc. Japan Acad. **79B**, 92-98.
- 16) Leonard, P. M., Smits, S. H. J., Sedelnikova, S. E., Brinkman, A. B., de Vos, W. M., van der Oost, J., Rice, D. W., and Rafferty, J. B. (2001) EMBO J. **20**, 990-997.
- 17) Kudo, N., Allen, M. D., Koike, H., Katsuya, Y., and Suzuki, M. (2001) Acta Crystallogr. **D57**, 469-471.
- 18) Suzuki, M. (2003) Proc. Japan Acad. **79B**, 213-222.
- 19) Suzuki, M. and Koike, H. (2003) Proc. Japan Acad. **79B**, 114-119.
- 20) Suzuki, M. (1993) EMBO J. **12**, 3221-3226.
- 21) Suzuki, M., Yagi, N., and Gerstein, M. (1995) Protein Engineer. **8**, 329-338.
- 22) Suzuki, M., and Gerstein, M. (1995) Proteins **23**, 525-535.
- 23) Platko, J. V., and Calvo, J. M. (1993) J. Bacteriol. **175**, 1110-1117.
- 24) Madhusudhan, K. T., Lorenz, D., and Sokatch, J. R. (1993) J. Bacteriol. **175**, 5140-5147.