Brief Articles

Similarity Based Virtual Screening: A Tool for Targeted Library Design

Joni K. O. Alvesalo,[†] Antti Siiskonen,[‡] Mikko J. Vainio,[§] Päivi S. M. Tammela,[†] and Pia M. Vuorela*[,§,†]

Drug Discovery and Technology Development Center (DDTC), Faculty of Pharmacy, P.O. Box 56, FIN-00014 University of Helsinki, Finland, Lead Pharmaceuticals Ltd., Helsinki, Finland, and Department of Biochemistry and Pharmacy, Åbo Akademi University, Tykistökatu 6A, FIN-20520 Turku, Finland

Received December 1, 2005

High throughput screening drug discovery utilizes large and expensive compound libraries. As an alternative, a smaller targeted library can be constructed with the aid of the 3D structure of the target molecule. We used the X-ray crystal structure of a protein homologous to the selected target in creation of a small focused library and evaluated inhibition potential of this library against *Chlamydia pneumoniae*, a common pathogen recently linked to atherosclerosis and risk of myocardial infarction.

Introduction

Chlamydia pneumoniae, an obligatory intracellular parasite worldwide, causes pneumonia, bronchitis, sinusitis, and pharyngitis. In addition, more serious diseases such as atherosclerosis[1] have recently been associated with chronic *C. pneumoniae* infection, which has proven to be extremely difficult to diagnose and treat with current antibiotics. High doses and prolonged treatment with currently used antibiotics are often needed in order to achieve a clinical cure, and they also involve the risk of persistence of *C. pneumoniae* in the tissues after treatment.[2] Lack of sensitive antimicrobial agents,[3] difficulty in developing a vaccine, and new evidence of lethal aspects of this bacterium[4] suggest the necessity for finding new potential antichlamydial substances.

To our knowledge, no extensive new antimicrobial agent screens against *C. pneumoniae* have been performed, probably due to the labor-intensive and time-consuming nature of traditional screening methods. The new 96-well plate *C. pneumoniae* assay,[5] based on time-resolved fluorometry, is definitely a step in the higher throughput direction, but without automation it is still quite laborious for the evaluation of large libraries. One effective way to diminish the size of a primary screening library is to create a targeted library by virtually screening molecules that show affinity to the 3D structure model of a selected target. In the case of *C.* pneumoniae, no publicly available X-ray crystal structure exists for any protein. Therefore, we decided to use the X-ray structure of a protein (RNA methyltransferase) with a highly similar amino acid sequence to that of the *C.* pneumoniae target protein (dimethyladenosine transferase) in creation of the targeted library. We did not want to create a comparative model for the *C. pneumoniae* target protein, since we wanted to test whether high level of similarity in the amino acid sequence, and especially in the binding pocket, is enough to guide the creation of a small library of active compounds. We set out to search a *C. pneumoniae* target protein that is crucial for bacterial survival and for which a closely related crystallized nonhuman protein exists. The use of a human protein, as a virtual screening target, could lead to discovery of compounds that are cytotoxic to human cells.

Results and Discussion

Search of public protein sequence databases produced a promising target: *Chlamydia pneumoniae* dimethyladenosine transferase (coded by the *ksgA* gene), which has a highly similar crystallized relative, *Bacillus subtilis* RNA methyltransferase (coded by the *ermC* gene). Both of these prokaryotic methyltransferases also bind their ligands in a similar way.[6] In the International Union of Biochemistry and Molecular Biology (NC-IUBMB) classification, the two proteins belong to the same subclass of rRNA (adenine-N^6-)-methyltransferases (EC 2.1.1.48), whose indispensable functions are related to ribosomal structure and ribosomal methylation. The FlexX program[7] was used to dock ligands to the X-ray crystal structure of ErmC′ RNA methyltransferase (1QAO).[8] Two databases of commercially available compounds, namely Specs (Delft, Netherlands: http://www.specs.net) and Maybridge (Cornwall, England: http://www.maybridge.com/html/m_msc_unit.htm), were screened in silico. The databases contained a total of 3×10^5 molecules. All molecules containing reactive groups, or other groups considered undesirable for drug molecules, were eliminated before virtual screening. After the screening process, the 2000 best binding molecules, ranked by the FlexX program, were analyzed visually and 33 molecules were selected and purchased. These molecules showed the best estimated binding properties and drug-like characteristics[9] and maximum diversity in structure. The 33 selected compounds formed 15 structurally different families with one to nine members in each family. This targeted library was evaluated in the cell-based *C. pneumoniae* assay.[5] The eight most active molecules, coming from seven structural families, were studied in more detail including inhibition efficiency testing in different concentrations and preliminary cytotoxicity determination (Table 1).

The most interesting structural family of ligands contained five molecules (Figure 1), of which two were active (MB4 and MB5) and three were inactive (MB2, MB3, and MB6). Although the differences in activity were large (Table 1), the structural differences were rather small. The only feature separating active

* To whom correspondence should be addressed. Tel: + 358 2 215 4267, Fax: +358 2 215 3280, E-mail: pia.vuorela@abo.fi.
[†] DDTC University of Helsinki.
[‡] Lead Pharmaceuticals Ltd.
[§] Åbo Akademi University.

10.1021/jm051209w CCC: $33.50
Published on Web 03/09/2006

Table 1. Results of *Chlamydia pneumoniae*, Cytotoxicity, and Viability Assays

compound	inhibition % 70 μM	inhibition % 50 μM	inhibition % 5 μM	inhibition % 1 μM	cytotoxicity %	viability %
S1	83.9	79.8	76.0	0.0	1.4	95
S9	93.6	81.8	15.3	0.0	2.3	86.2
MB 1	71.2	68.6	58.8	2.1	0.4	98.2
MB 4	92.6	90.5	41.2	39.5	1.2	98
MB 5	69.1	69.6	19.7	14.6	1.2	95.5
MB 12	72.5	56.1	0.2	0.0	0.0	100.0
MB 13	99.3	85.5	6.8	0.0	1.1	99.3
MB 16	100.0	72.8	0.0	0.0	3.7	83.4

[a] Percentages of inhibition, cytotoxicity, and viability for all compounds showing over 50% inhibition at 50 μM concentration in a *C. pneumoniae* assay. Inhibition values are averages from two experiments, both with six replicates, performed on different days. Cytotoxicity and viability values are averages from four replicates in one experiment with 50 μM substance concentration.

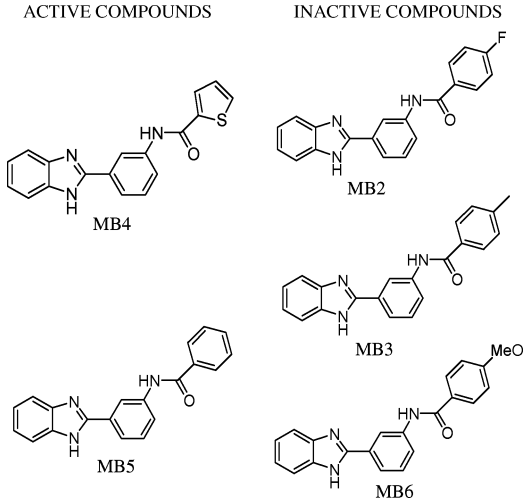

Figure 1. Structural family of ligands with five similar members.

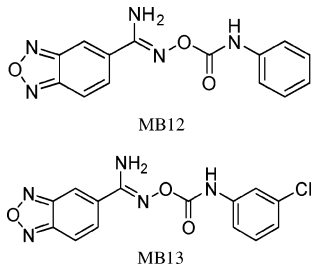

Figure 2. Structural family with two active compounds.

molecules from the inactives was a para-substituent in the phenyl ring of inactive compounds, or replacement of the phenyl with a slightly smaller thiophene ring in the active compound MB4, which might indicate lack of space for the para-substituent in the target molecule binding pocket.

Another interesting structural family contains two active molecules MB12 and MB13 (Figure 2). The sole difference in structure is the chlorine substituent in the third carbon of the phenyl ring, enhancing activity but not cytotoxicity (Table 1). This encourages testing of other substituents in the meta-position of the phenyl ring in order to find even more active compounds.

Virtual screening is widely used to find molecules more likely to bind to the selected target than randomly selected molecules. Andrew et al.[10] estimated that 1−20% of the virtual hits from a pharmacophore-based virtual screening are active in the actual in vitro assay[11−14] depending on the pharmacophore model and the used filtering system. This expected hit rate is considerably higher than the hit rate for a large random library screening; examples of hit rates from the literature are 0.02%,[15] 0.26%,[16] 0.04%,[17] 0.20%,[18] and 0.23%.[19]

The hit rate for our targeted library was 24.2% (8 of 33), which is considerably higher than what would be expected for a random library HTS. Considering the small size of the targeted library, it is unlikely to find as many active yet structurally different molecules as we did, unless the ligands act on the same target protein in the cell-based screening assay.

In this study, virtual screening against the 3D structure of a structural homologue of the target was shown to be a fast and economical way to improve the odds of finding potential drug candidates from compound databases.

Experimental Section

Human Lung (HL) Cell Line and *Chlamydia* Inoculum. HL cells served as host cells in the experiments. The HL cells were cultivated in RPMI 1640 medium (BioWhittaker Europe, Verviers, Belgium) supplemented with 10% of fetal bovine serum of South American origin (BioWhittaker Europe), 2 mM L-glutamine, and 20 μg/mL streptomycin. The cells were kept at 37 °C and 95% humidity in an atmosphere containing 5% of CO_2. *Chlamydia pneumoniae* elementary bodies (EBs), strain Kajaani 7, were stored in sucrose−phosphate−glutamic acid buffer (0.2 M sucrose, 3.8 mM KH_2PO_4, 6.7 mM Na_2HPO_4, 5 mM L-glutamic acid, pH 7.4) at −70 °C.

Infection of HL Cells. HL cells were plated on 96-well plates, 5×10^4 cells/well (Wallac Isoplate 1450−516, PerkinElmer Life and Analytical Sciences/Wallac Oy, Turku, Finland). After 24 h incubation, confluent cell monolayers were inoculated with *C. pneumoniae* EBs (4×10^5 inclusion-forming units per mL) using centrifugation (550 × g, 60 min, Heraeus Multifuge 3 S, Kendro Laboratory Products GmbH, Hanau, Germany). The desired concentration of the agent studied was prepared in antibiotic-free cell culture media supplemented with 0.5 μg/mL cycloheximide (Sigma, St.Louis, MO), an 80-s ribosome inhibitor, and added to the cell cultures after centrifugation. The cultures were incubated for 72 h in 5% CO_2 at 37 °C. After incubation, the cells were fixed with methanol for the *C. pneumoniae* assay.

***Chlamydia pneumoniae* Assay.** The assay is described in detail elsewhere.[5] In short, the infected HL cells were fixed in 96-well plates and labeled with a genus-specific murine monoclonal antibody (Argene SA, Varilhes, France), containing an europium label (PerkinElmer/Wallac Oy, Turku, Finland) at 100 ng/mL concentration. After 30-min incubation at 37 °C, the plates were washed six times with a Biohit BW50 plate-washer (Biohit Plc, Finland) using 300 μL of Wallac DELFIA Wash Solution (PerkinElmer Life and Analytical Sciences/Wallac Oy) each time. An aliquot of 100 μL of DELFIA Enhancement Solution (PerkinElmer) was added to each well, and the plates were shaken for 5 min at low speed on a DELFIA Plateshake (PerkinElmer). Signals associated with the *Chlamydia* antibody were measured from the wells with a Wallac Victor^2 multilabel counter (PerkinElmer).

Cytotoxicity and Viability Assays. The cytotoxicity assay, Promega CytoTox 96 Non-Radioactive Cytotoxicity Assay (Promega, Madison, WI), was used to measure the release of lactate dehydrogenase (LDH) from the HL cells to the culture media. LDH is released from the cell upon loss of plasma membrane integrity or necrosis. Cytotoxicity percentage was obtained by comparing the amount of LDH from cell culture medium after 2 h of exposure

Table 2. Active Site Homology[a]

B. subtilis ErmC' RNA methyltransferase	10 Gln	11 Asn *	12 Phe	13 Ile *	36 Glu *	38 Gly *	39 Ser	40 Gly	59 Glu *	60 Ile	84 Asp *	85 Ile *	101 Asn *	103 Pro
C. pneumoniae dimethyladenosine transferase	26 Gln	27 Asn	28 Phe	29 Leu	52 Glu	54 Gly	55 Pro	56 Gly	75 Glu	76 Lys	95 Asp	96 Ala	112 Asn	114 Pro

[a] *Residues participating in hydrogen bonding. Shading dentoes correspondence between ErmC′ RNA methyltransferase and *C. pneumoniae* dimethyladenosine transferase active-site amino acids.

to the tested substance, to amount of LDH from unexposed control cell culture medium.

The viability assay uses intracellular LDH to measure the number of viable HL cells after 3 days of exposure to the substances tested. After culture media removal, cells were washed twice and lysed, so that only intracellular LDH from viably cells was measured. Percentages were obtained by comparing the unexposed-culture LDH amount to the exposed-culture LDH. This assay was also performed with the Promegas CytoTox 96 Non-Radioactive Cytotoxicity Assay according to manufacturer instructions.

Docking. The peptide chain from the X-ray crystal structure 1QAO was used as the target protein. The structure was preprocessed in the program Sybyl, version 6.5 (SYBYL 6.5, Tripos Inc., 1699 South Hanley Rd., St. Louis, MO 63144). Atom types of carboxylic acid oxygens were set to O.co2 and those of basic nitrogens to N.4. Hydrogens were added to the model, and their orientation was optimized using the Tripos force field energy minimization while all non-hydrogen atoms were not allowed to move. The ligand position in 1QAO was used to define the active site cavity.

Default FlexX (Sybyl 6.5) parameters were used. 2000 top ranking molecules were further inspected visually to exclude any structures with improbable docking orientation, and 33 molecules with good docking orientations were chosen for biological testing.

Homology. Sequence homology: Statistics of a BLAST alignment between *Chlamydia pneumoniae* dimethyladenosine transferase (coded by the *ksgA* gene)[20] and *Bacillus subtilis* RNA methyltransferase (coded by the *ermC* gene)[21] amino acid sequences are as follows: score = 70.1 bits (170), expect = 5×10^{-11}, identities = 67/258 (25%), positives = 119/258 (45%), and gaps = 31/258 (12%). *B. subtilis* amino acid numbers are those of the crystal structure[8] numbers, and *C. pneumoniae* amino acids are numbered using UniProt-database (http://www.ebi.uniprot.org/) sequence numbers, accession number: Q9Z6K0.

Active site homology: BLAST alignment of the ErmC′ RNA methyltransferase and *C. pneumoniae* dimethyladenosine transferase active-site amino acid sequences (Table 2). Residues participating in hydrogen bonding according to Figure 3 in Schluckebier et al.[8] are marked with an asterisk. Positions of identical residues are shaded.

Targeted Library. International Union of Pure and Applied Chemistry (IUPAC) names of all target library compounds:

S = Specs (Delft, Netherlands) and MB = Maybridge (Cornwall, England).

S1 = 2-{4-amino-6-[(4-chlorophenyl)amino]-1,3,5-triazin-2-yl}-4-chlorophenol

S2 = *N*-1,3-benzodioxol-5-yl-2-[(5-methyl-1*H*-benzimidazol-2-yl)thio]acetamide

S3 = 2-(1*H*-benzimidazol-2-ylthio)-*N*-(3-hydroxyphenyl)acetamide

S4 = *N*-(3-fluorophenyl)-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)-acetamide

S5 = 2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)-*N*-phenylacetamide

S6 = *N*-1,3-benzodioxol-5-yl-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)acetamide

S7 = 2-[(1*H*-benzimidazol-2-ylmethyl)thio]-3*H*-imidazo[4,5-*b*]-pyridine

S8 = *N*-[3-(acetylamino)phenyl]-4-methoxybenzamide

S9 = 3-[(2,5-dichlorophenoxy)methyl]-*N*-2,3-dihydro-1,4-benzodioxin-6-ylbenzamide

S10 = *N*-(2,4-difluorophenyl)-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)acetamide

S11 = *N*-(4-fluorophenyl)-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)-acetamide

S12 = *N*-1,3-benzodioxol-5-yl-2-[(5-methoxy-1*H*-benzimidazol-2-yl)thio]acetamide

MB1 = *N*-(1,3-benzodioxol-5-ylmethyl)-6-phenylthieno[3,2-*d*]-pyrimidin-4-amine

MB2 = *N*-[3-(1*H*-benzimidazol-2-yl)phenyl]-4-fluorobenzamide

MB3 = *N*-[3-(1*H*-benzimidazol-2-yl)phenyl]-4-methylbenzamide

MB4 = *N*-[3-(1*H*-benzimidazol-2-yl)phenyl]thiophene-2-carboxamide

MB5 = *N*-[3-(1*H*-benzimidazol-2-yl)phenyl]benzamide

MB6 = *N*-[3-(1*H*-benzimidazol-2-yl)phenyl]-4-methoxybenzamide

MB7 = *N*-2,3-dihydro-1,4-benzodioxin-6-ylbenzamide

MB8 = *N*-2,3-dihydro-1,4-benzodioxin-6-yl-4-methoxybenzamide

MB9 = *N*-(4-morpholine-4-aminobenzyl)-3,4-dihydro-2*H*-1,5-benzodioxepine-7-carboxamide

MB10 = 9-{2-[(1,3-benzodioxol-5-ylmethyl)amino]ethyl}-9*H*-purin-6-amine

MB11 = *N*-{4-[(pyrimidin-2-ylamino)sulfonyl]phenyl}acetamide

MB12 = *N*′-[(anilinocarbonyl)oxy]-2,1,3-benzoxadiazole-5-carboximidamide

MB13 = *N*′-({[(3-chlorophenyl)amino]carbonyl}oxy)-2,1,3-benzoxadiazole-5-carboximidamide

MB14 = 2-[(2,1,3-benzoxadiazol-5-yloxy)methyl]-*N*-(4-chlorophenyl)-1,3-thiazole-4-carboxamide

MB15 = 2-[(2,1,3-benzoxadiazol-5-yloxy)methyl]-*N*-(4-nitrophenyl)-1,3-thiazole-4-carboxamide

MB16 = N^2,N^4-diphenyl-1,3,5-triazine-2,4,6-triamine

MB17 = *N*-benzyl-9*H*-purin-6-amine

MB18 = *N*-(2-furylmethyl)-9*H*-purin-6-amine

MB19 = *N*-(3-nitrophenyl)-1,3,5-triazine-2,4-diamine

MB20 = 4-chloro-*N*-(2-{[(6-nitro-4*H*-1,3-benzodioxin-8-yl)-methyl]thio}phenyl)benzamide

MB21 = *N*-(2-{[(6-nitro-4*H*-1,3-benzodioxin-8-yl)methyl]thio}-phenyl)thiophene-2-carboxamide

Acknowledgment. We are grateful to Prof. Maija Leinonen for donation of the HL cell line and *Chlamydia* inoculum, and to the National Technology Agency of Finland for financial support (grant 40083/03).

References

(1) Saikku, P.; Leinonen, M.; Mattila, K.; Ekman, M. R.; Nieminen, M. S.; Mäkelä, P. H.; Huttunen, J. K.; Valtonen, V. Serological evidence of an association of a novel Chlamydia, TWAR, with chronic coronary heart disease and acute myocardial infarction. *Lancet* **1988**, *8618*, 983−986.

(2) Gieffers, J.; Rupp, J.; Gebert, A.; Solbach, W.; Klinger, M. First-choice antibiotics at subinhibitory concentrations induce persistence of Chlamydia pneumoniae. *Antimicrob. Agents Chemother.* **2004**, *4*, 1402−1405.
(3) Dunbar, L. M. Current issues in the management of bacterial respiratory tract disease: the challenge of antibacterial resistance. *Am. J. Med. Sci.* **2003**, *6*, 360−368.
(4) Leinonen, M.; Saikku, P. Evidence for infectious agents in cardiovascular disease and atherosclerosis. *Lancet Infect. Dis.* **2002**, *1*, 11−17.
(5) Tammela, P.; Alvesalo, J.; Riihimäki, L.; Airenne, S.; Leinonen, M.; Hurskainen, P.; Enkvist, K.; Vuorela, P. Development and validation of a time-resolved fluorometric immunoassay for screening of antichlamydial activity using a genus-specific europium-conjugated antibody. *Anal. Biochem.* **2004**, *1*, 39−48.
(6) Seppälä, H.; Skurnik, M.; Soini, H.; Roberts, M. C.; Huovinen, P. A novel erythromycin resistance methylase gene (ermTR) in Streptococcus pyogenes. *Antimicrob. Agents Chemother.* **1998**, *2*, 257−262.
(7) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *3*, 470−489.
(8) Schluckebier, G.; Zhong, P.; Stewart, K. D.; Kavanaugh, T. J.; Abad-Zapatero, C. The 2.2 A structure of the rRNA methyltransferase ErmC′ and its complexes with cofactor and cofactor analogs: implications for the reaction mechanism. *J. Mol. Biol.* **1999**, *2*, 277−291.
(9) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3−25.
(10) Good, A. C.; Krystek, S. R.; Mason, J. S. High-throughput and virtual screening: core lead discovery technologies move towards integration. *Drug Discovery Today* **2000**, *12 Suppl 1*, 61−69.
(11) Wang, S.; Milne, G. W.; Yan, X.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. Discovery of novel, non-peptide HIV-1 protease inhibitors by pharmacophore searching. *J. Med. Chem.* **1996**, *10*, 2047−2054.
(12) Hong, H.; Neamati, N.; Wang, S.; Nicklaus, M. C.; Mazumder, A.; Zhao, H.; Burke, T. R., Jr.; Pommier, Y.; Milne, G. W. Discovery of HIV-1 integrase inhibitors by pharmacophore searching. *J. Med. Chem.* **1997**, *6*, 930−936.
(13) Kaminski, J. J.; Rane, D. F.; Snow, M. E.; Weber, L.; Rothofsky, M. L.; Anderson, S. D.; Lin, S. L. Identification of novel farnesyl protein transferase inhibitors using three-dimensional database searching methods. *J. Med. Chem.* **1997**, *25*, 4103−4112.
(14) Marriott, D. P.; Dougall, I. G.; Meghani, P.; Liu, Y. J.; Flower, D. R. Lead generation using pharmacophore mapping and three-dimensional database searching: application to muscarinic M(3) receptor antagonists. *J. Med. Chem.* **1999**, *17*, 3210−3216.
(15) Zolli-Juran, M.; Cechetto, J. D.; Hartlen, R.; Daigle, D. M.; Brown, E. D. High throughput screening identifies novel inhibitors of Escherichia coli dihydrofolate reductase that are competitive with dihydrofolate. *Bioorg. Med. Chem. Lett.* **2003**, *15*, 2493−2496.
(16) Lewis, L. M.; Engle, L. J.; Pierceall, W. E.; Hughes, D. E.; Shaw, K. J. Affinity capillary electrophoresis for the screening of novel antimicrobial targets. *J. Biomol. Screen.* **2004**, *4*, 303−308.
(17) Chapman, R. L.; Stanley, T. B.; Hazen, R.; Garvey, E. P. Small molecule modulators of HIV Rev/Rev response element interaction identified by random screening. *Antiviral Res.* **2002**, *3*, 149−162.
(18) Carey, K. L.; Westwood, N. J.; Mitchison, T. J.; Ward, G. E. A small-molecule approach to studying invasive mechanisms of Toxoplasma gondii. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *19*, 7433−7438.
(19) Pilger, B. D.; Cui, C.; Coen, D. M. Identification of a small molecule that inhibits herpes simplex virus DNA Polymerase subunit interactions and viral replication. *Chem. Biol.* **2004**, *5*, 647−654.
(20) Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Hyman, R. W.; Olinger, L.; Grimwood, J.; Davis, R. W.; Stephens, R. S. Comparative genomes of Chlamydia pneumoniae and C. trachomatis. *Nat. Genet.* **1999**, *4*, 385−389.
(21) Monod, M.; Denoya, C.; Dubnau, D. Sequence and properties of pIM13, a macrolide-lincosamide-streptogramin B resistance plasmid from Bacillus subtilis. *J. Bacteriol.* **1986**, *1*, 138−147.

JM051209W