ELSEVIER

Development and screening of a polyketide virtual library for drug leads against a motilide pharmacophore

Michael A. Siani,* A. Geoffrey Skillman,† Christopher W. Carreras,* Gary Ashley,* Irwin D. Kuntz,† and Daniel V. Santi*

**Kosan Biosciences, Hayward, California, USA*
†Department of Pharmaceutical Chemistry, University of California, San Francisco, California, USA

A virtual library of macrocyclic polyketide molecules was generated and screened to identify novel, conformationally constrained potential motilin receptor agonists ("motilides"). A motilide pharmacophore model was generated from the potent 6,9-enol ether erythromycin and known derivatives from the literature. The pharmacophore for each molecular conformation was a point in a distance-volume space based on presentation of the putative binding moieties. Two methods, one fragment based method and the other reaction based, were explored for constructing the polyketide virtual library. First, a virtual library was assembled from monomeric fragments using the CHORTLES language. Second, the virtual library was assembled by the in silico *application of all possible polyketide synthase enzyme reactions to generate the product library. Each library was converted to low-energy 3D conformations by distance geometry and standard minimization methods. The distance-volume metric was calculated for low-energy conformations of the members of the virtual polyketide library and screened against the enol ether pharmacophore. The goal was to identify novel macrocycles that satisfy the pharmacophore. We identified three conformationally constrained, novel polyketide series that have low-energy conformations satisfying the distance-volume constraints of the motilide pharmacophore.*

Color Plates for this article are on pages 539–540.

Corresponding author: Michael A. Siani, Kosan Biosciences, 3832 Bay Center Place, Hayward, CA, USA. Tel.: 510-732-8400 X 223; fax: 510-732-8401.

E-mail address: siani@kosan.com (M.A. Siani).

Keywords: polyketide, virtual library, motilide, virtual screening, macrolactone

INTRODUCTION

Screening of Virtual Libraries

Traditional medicinal chemistry has focused on optimization of a known drug lead. Medicinal chemists were quite successful without any consideration of computational representation. However, in the past two decades, computational prediction of some molecular properties has become facile. Increasingly, the computational predictions, particularly binding potency predictions, are being incorporated by medicinal chemists in selecting which compounds to synthesize. During the past decade, medicinal chemistry has been revolutionized by combinatorial chemistry and high-throughput screening. Initially, these efforts focused on nucleic acid, peptide, and peptide-mimetic libraries,[1,2] but soon were followed by more "drug-like" organic, heterocyclic libraries.[3–6] Initial attempts to generate libraries of compounds focused on developing available chemistry. More recently, computational methods have been used to design libraries with optimized properties, such as diversity, bioavailability, and potency.[7–10]

Only recently has work focused on *in silico* construction *and* screening of libraries that meet the constraints of a particular pharmacophore for a given drug target. It is common that the specific interactions of a lead compound with its target are unknown. Despite this, the critical interactions can be inferred by comparison of common structural and functional motifs in a set of active molecules, a process termed *pharmacophore construction.* Once the set of common features (pharmacophore) has been

Journal of Molecular Graphics and Modelling 18, 497–511, 2000

655 Avenue of the Americas, New York, NY 10010

1093-3263/00/$–see front matter
PII S1093-3263(00)00070-X

specified, one can search large, real, or virtual databases for compounds that contain many or all of the salient features.

In recent years, several groups have developed pharmacophore models of the motilin interaction with the motilin receptor based on alignment of erythromycin derivatives with the motilin peptide or based solely on amino acid scans of the N-terminal portion of the peptide.[11] Khiat and Boulanger[12] used conformational analysis and three-dimensional quantitative structure-activity relationship (QSAR) methods to determine homology between a series of erythromycin A derivatives and the motilin peptide.[13] Beavers[14] effectively demonstrated the design and synthesis of a potent motilin receptor antagonist based on a pharmacophore derived solely from the motilin peptide.

We describe here the first example of the generation of a macrocyclic polyketide virtual library, as well as its screening against a unique pharmacophore based on the potent enol ether erythromycin motilide class.

Motilide Prokinetic Agents

Gastrointestinal (GI) motility regulates the orderly movement of ingested material through the gut to ensure adequate absorption of nutrients, electrolytes, and fluids. Appropriate transit through the esophagus, stomach, small intestine, and colon depends on regional control of intraluminal pressure and several sphincters that regulate forward movement and prevent backflow of GI contents. There are several significant human diseases of low gastric motility, the most prevalent of which are gastroparesis and gastroesophageal reflux disease.

The newest class of prokinetic agents are agonists of the motilin receptor. Motilin is a 22–amino-acid peptide that, on binding to the motilin receptor, stimulates gastric motility.[15,16] Erythromycin and certain erythromycin analogues also are potent agonists of the motilin receptor.[17,18] These macrolides that act as motilin agonists are called "motilides." Since its introduction in the 1950s, oral administration of erythromycin has been known to cause GI side effects such as diarrhea, flatulence, nausea, vomiting, and abdominal discomfort. Erythromycin subsequently was found to be a weak motilin agonist (ED_{50} ~1 μM). Importantly, erythromycin undergoes an acid-catalyzed rearrangement in the stomach, to form an 8,9-anhydro-6,9-hemiacetal ("enol ether," **II**), which is a 10-fold more potent motilin receptor agonist than erythromycin A and is devoid of antimicrobial effects.[19] Azithromycin and clarithromycin, two commonly used macrolactone antibiotics, are erythromycin analogues with structural modifications directed at eliminating this route of metabolism. Conversely, the most potent motilides also are analogues derived from the erythromycin enol ether scaffold (*vida infra*). Additional exposure of the erythromycin enol ether to acidic conditions causes further conversion to the inactive spiroketal (**III**), thus limiting the duration of prokinetic activity *in vivo*.

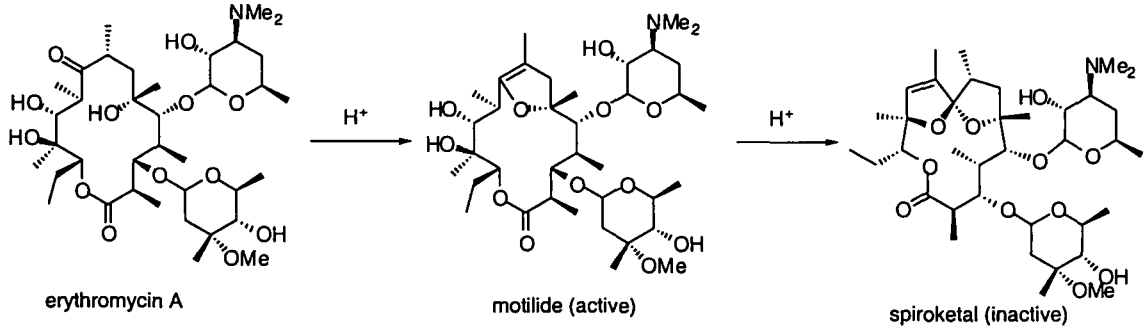
SARs of Motilides

Semisynthetic methods have been used to decrease acid instability and to increase potency of the enol ether (**II**) derived from erythromycin. Although such studies are not extensive, the combined information permits certain conclusions regarding regions of motilides that contribute and detract from activity.[20] Below is a synthesis of the concepts that contributed to our current SAR for motilide activity.

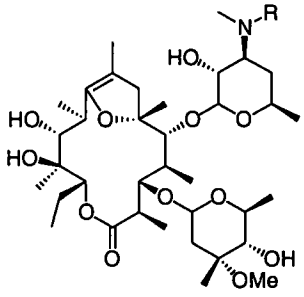
Even though acid stability is improved, saturation of the carbon-carbon double bond of the enol ether to give a less conformationally restricted molecule results in mixed effects depending on the stereochemistry at this position.[40] Macrolides containing only a desosamine residue and lacking the cladinose (e.g., 5-O-desosaminylerythronolide and picromycin) are inactive, as are those lacking an amine group on the desosamine.[19] Variation of the amine alkyl groups on the desosamine sugar provided two analogues having potential as prokinetic agents, EM-523 (**IV**) and EM-574 (**V**).[19,22] *In vitro*, EM-523 is 40-fold more potent than the erythromycin A enol ether (**II**) as a motilide and is devoid of antimicrobial activity. However, like its erythromycin parent, EM-523 shows poor oral bioavailability due to further acid degradation of the enol ether into inactive spiroketals (e.g. **II** → **III**).[23]

The acid-catalyzed conversion of enol ether motilides to spiroketals can be circumvented by blocking the reactive 12-hydroxyl group with an alkyl ether, as in the 11-keto-12-O-methyl motilide GM-611 (**VI**).[24] GM-611 shows about the same affinity for the motilin receptor as does EM-523; however, when incubated at pH 2.5 for 2 hours, it retains affinity for the motilin receptor, whereas EM-523 shows a 100-fold loss.

A second approach to the acid instability problem is to remove the 12-hydroxyl group, giving motilides derived from



I II III



IV, EM-523 R= ethyl

V, EM-574 R=isopropyl

erythromycin B rather than erythromycin A.[25] These analogues retain prokinetic activity and preclude acid-catalyzed spiroketal formation. These studies also found that removal of the cladinose 4″-hydroxyl group provided highly potent motilides, with the 4′-deoxy analogue of EM-523 a much more potent motilide *in vitro* than EM-523. Combined removal of the 12- and 4″-hydroxyl groups produced motilide, ABT-229 (**VII**), which shows ~750-fold higher *in vitro* activity than EM-523. ABT-229 also shows good oral bioavailability compared to EM-523, due to its acid stability (35% vs 1.4%).[21,26]

Finally, Greenwood et al.[27] reported a 12-membered motilide, LY267108 (**VIII**), but the free acyclic hydroxyl group makes it susceptible to acid degradation.

In summary, the limited SAR studies performed thus far have demonstrated the following: (1) 12- or 14-membered macrolides can be potent motilides; (2) the enol ether moiety contributes about 100-fold to binding compared to erythromycin A, although it is not clear whether this is due to specific interactions or a change in overall conformation; (3) the desosamine amino group is essential, and minor alkyl modifications of the amino group improve activity; (4) removal of the cladinose hydroxyl group can improve activity; and (5) stereospecific reduction of the enol ether to (9*R*),(8*S*) compound increases activity.

Historically, two assays have been used to assess motilide activity: one measures the strength of the interaction between a motilide and the motilin receptor, and the second measures physiological effects of motilide on duodenal tissues. A robust competitive binding assay using radiolabeled motilin and a membrane preparation from rabbit antral smooth muscle tissue has been used for the determination of SARs of motilides[18,28] and provides quantitative assessment of receptor binding. To demonstrate and quantitate the agonistic effect of potential motilides, contractile activity is tested by measuring the response of isolated segments of rabbit duodenum to exogenously added motilides.[17,20] Clinical evaluation of motilides can involve the introduction of force transducers into the digestive systems of experimental animals that are used to observe contractile activity *in situ*.[27,29]

MATERIALS AND METHODS

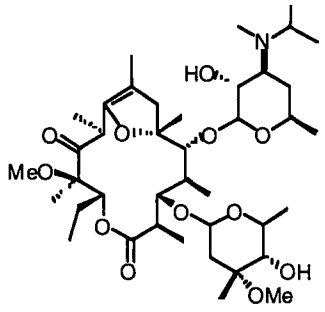
Our approach entails the design and generation of macrocyclic polyketide virtual libraries, generation of the motilide pharmacophore, and screening of the libraries against the pharmacophore model (Figure 1). The macrocycle polyketide libraries were generated using both the CHORTLES monomer approach and the reaction-based Diversify method. First, a set of large, generic 12- and 14-membered macrocyclic libraries were generated. Later, a more focused library of 2,000 compounds was generated to facilitate rapid screening against the pharmacophore. Several approaches were taken to create the pharmacophore, including the standard Comparative Molecular Field Analysis (CoMFA) hypothesis generation, as well as the conventional approach based on presentation of moieties of low-energy conformations of known leads. Finally, the library was screened against the pharmacophore, and compounds were sorted based on their pharmacophore match score.

Generation of Virtual Polyketide Libraries

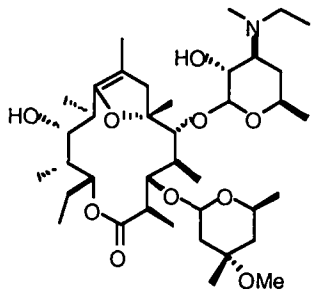
Our approach was to generate a simple set of monomers and attach them (or react them) to build up a virtual macrocyclic polyketide library of all possible combinations. Two methods were used to generate the members of these libraries. Because the goal was to mimic the constrained conformation of the enol ether analogues, macrocyclic scaffolds were sought that present the enol ether analogue moieties as they are oriented while bound to the motilin receptor. Our library, therefore, consisted of macrocyclic polyketides without modifications (i.e., C6-OH, C12-OH, C3-cladinosyl, C5-desaminyl). The basic macrocycle, biosynthesized by the DEBS polyketide synthase (PKS) complex, is known as 6-dEB.

The modular PKSs described to date are large multifunctional protein complexes with dedicated domains for each enzyme-catalyzed step in polyketide chain assembly and modification. Active sites are clustered into modules, each containing a full complement of the sites required for one cycle of condensation and the associated reduction cycle. Modular PKSs act in an assembly-line fashion in which the growing polyketide chain is passed from module to module as polymerization occurs. The erythromycin PKS from *Saccharopolyspora erythraea*, known as 6-deoxyerythronolide B synthase (DEBS), is the prototypical example of modular PKS (Figure 2). Technology is rapidly becoming available that allows the rational manipulation of DEBS and enables the generation of specific novel polyketide structures.[30–34] Because they are made modularly by a series of enzymes that may be genetically modified, macrocyclic polyketides lend themselves to compact library representation. Combinatorial rearrangement of the erythromycin PKS has been accomplished to generate a library of 6-dEB analogues with more than 50 members.[34] The virtual library presented here is an elaboration of this biosynthesized library.

Generic virtual libraries The generic macrocycle library consists of the set of all compounds with all combinations of natural monomers at each position in the macrocycle. The 14-membered macrocycle library contains 14 atoms, or seven monomers in the cycle (Figure 3). Each monomer contains two substituents, R1 and R2. R1 contributes three substituents (including both enantiomers for the methyl) at six positions in



VI, GM-611



VII, ABT-229

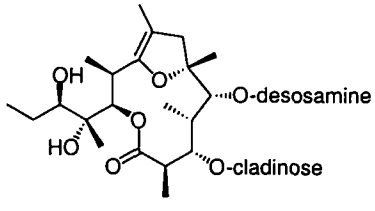
the macrocycle. R2 contributes five possible substituents (including the potential backbone double bond and two enantiomers for the hydroxyl substitution) at five positions in the macrocycle. R3 has just two choices at one position based on the erythromycin 6-dEB model. This generic library contains $3^6 \times 5^5 \times 2^1 = 4{,}556{,}250$ unique molecules.

The restricted motilide library (14-membered macrocycle) with C3-OH and C5-OH specified reduces the total contents of the library to $3^6 \times 5^3 \times 2^2 \times 2^1 = 729{,}000$ unique molecules. The additional 2^2 term is derived from the two enantiomers of the C3-OH and C5-OH.

The 12-membered macrocycle library contains 12 atoms, or six monomers in the cycle (Figure 3). R1 contributes three substituents (including both enantiomers for the methyl) at five positions. R2 contributes five possible substituents (including the potential backbone double bond and two enantiomers for the hydroxyl substitution) at four positions. R3 has just two choices at one position based on the erythromycin 6-dEB model. This library contains $3^5 \times 5^4 \times 2^1 = 303{,}750$ unique molecules.

Diversify approach Macrolactone libraries also were generated by modeling the PKS reactions with the chemical engine Diversify (A. Geoffrey Skillman, University of California, San Francisco, Dissertation). For a 14-membered macrolactone, such as erythromycin, the synthesis consists of six cycles of condensation and reduction followed by cyclization and two additional hydroxylations. For condensations, three reactions were considered: malonyl-CoA condensation, and methylmalonyl-CoA condensation (branched methyl with *R* or *S* stereochemistry). For reduction, five reactions were considered; no reduction (ketone preserved), keto-reduction (hydroxyl with *R* or *S* stereochemistry), dehydration (trans double bond), and enoyl-reduction (yields a methylene). Twelve transforms were generated to encode the basic macrolactone polyketide synthesis, including the transesterification ring closure and the two hydroxylations. Two additional transforms were generated to model the acid-catalyzed metabolism of the 14-membered macrolactones into enol ethers and spiroketals. When these transforms were applied sequentially to the propyl starting material, a single 6-dEB product is generated.

To generate libraries of macrolactones, the basic Diversify method was extended to include reaction sequence control. Reaction sequence control allows the user to determine which reactions are carried out at each step of the synthesis. In the degenerate form, the user simply specifies the synthetic scheme, for example, the series of reactions needed to generate the desired product. Thus, to generate a single compound, such as 6-dEB, a reaction sequence with only one reaction in each step of the sequence can be used. In the more complex form, the user can specify multiple possible reactions at each step, which can all be applied to generate a library of products. For instance, in order to form a library with all possible oxidation states at C9, all five reduction transforms would be applied in the third round of reduction; alternatively, in order to form an enol ether, the C9 must be left in the ketone oxidation state, so only the "no reduction" transform would be applied in the third round of reduction. In both cases, the number and type of transform applied at each step is controlled independently. The reaction sequence can contain as many reactions at each step as necessary, and the reactions at each step are applied exhaustively to the products from the previous reaction step to generate a "reaction-based" combinatorial library. In the library generation described here, we limited ourselves to 14-membered macrolactones with hydroxyl groups on the C3 and C5 carbons, which are the natural sites of glycosylation.



VIII, LY267108

CHORTLES approach Using the CHORTLES representation,[35] the molecular graphs of any library of polyketides are readily represented. The 6-dEB may be represented as a molecular graph (line drawing or SMILES, the equivalent text representation) or as a monomer-based graph or CHUCKLES.[36] All three representations of 6-dEB are shown in Figure 4. A library is a set of compounds, represented by a monomer-based set (CHORTLES).

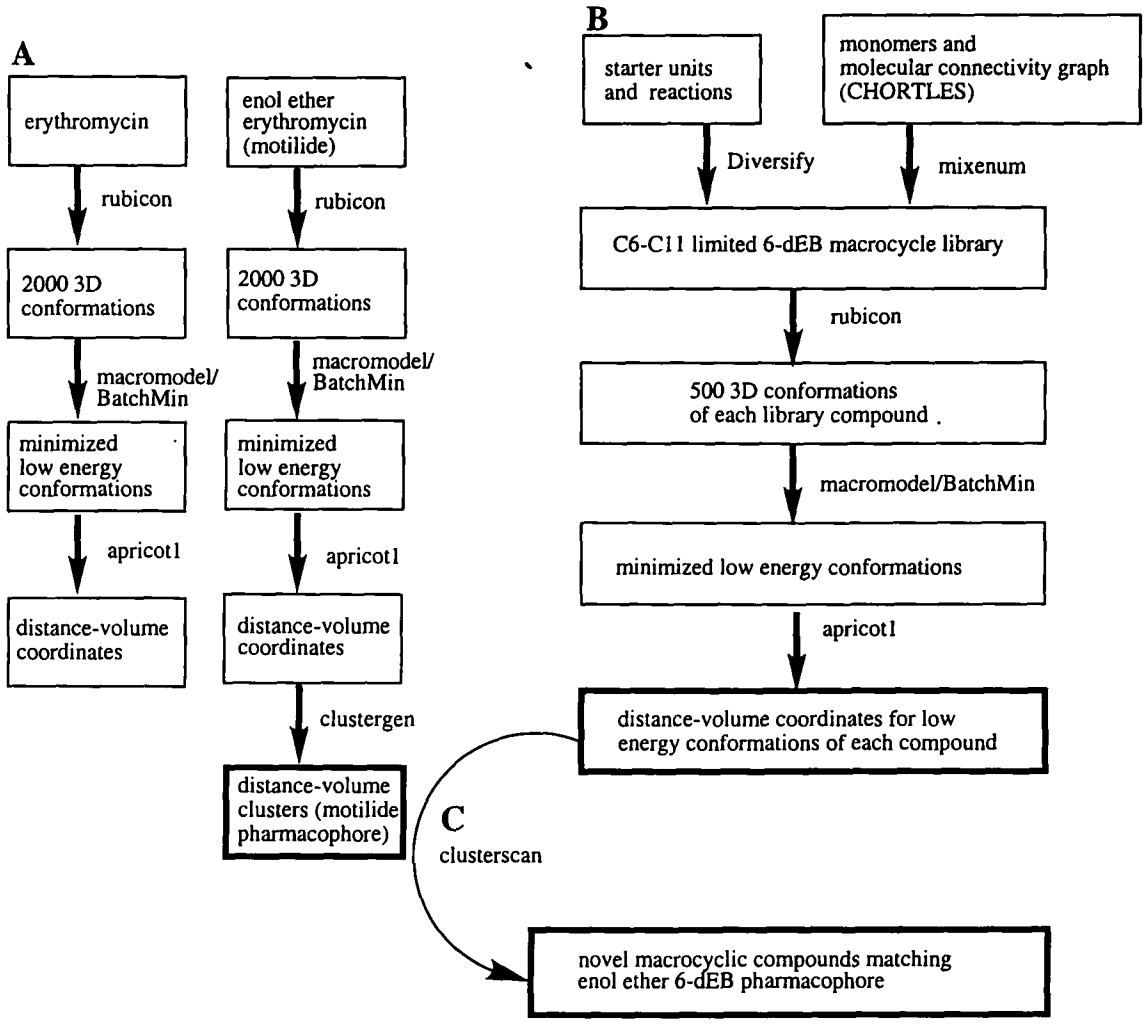


Figure 1. Flow chart detailing pharmacophore generation (A), virtual library generation (B), and screening of library components against the pharmacophore (C).

Just as the polyketide chain is built up by the addition of two-carbon units, the CHORTLES macrocyle may be represented by a series of subunits linked into a chain and then cyclized. We chose to represent each polyketide by a starter unit *Starta*, the carbonyl (C=O) unit. Attached to this starter unit are successive two-carbon backbone pieces with various branching atoms. The two-carbon subunit monomers (*Mono_a* through *Mono_n*) have the following properties: (1) different chirality at each of the two backbone positions, (2) no substitution or a methyl branching off the first carbon position, and (3) various degrees of reduction of an oxygen substitution at the second backbone carbon atom, including keto, hydroxyl, methylene (no oxygen), or a double bond connecting the two backbone carbons. Due to the nature of the enzymes that build these molecules, all double bonds connecting the two backbone carbons are *trans*. The closing units contain the same two-carbon units, plus an ester linkage to attach to the starter unit. All possible, naturally occurring monomers are shown in Figure 5.

The compounds in the CHORTLES representation were rapidly enumerated with the mixenum program (Kosan Biosciences, Hayward, CA), using the Daylight monomer toolkit (Daylight Chemical Information, Santa Fe, NM). Mixenum yields molecular graphs (SMILES) for all entries in the library. The SMILES,[37] containing defined atoms, bonds, and stereochemistry, are converted to multiple low-energy 3D conformations using distance geometry techniques by rubicon (Daylight Chemical Information Systems, Irvine, CA), then subjected to 100 steps of conjugate gradient minimization using the MM2* force field (Macromodel/BatchMin, Columbia University, New York, NY).

Focused library: C3-OH, C5-OH, variation between C6-C12, with fixed methyl chirality Because it was untenable to thoroughly explore conformational space of the millions of compounds in the generic, virtual macrocyle library, we focused on a library where only the region near the enol ether ring (C6-C9) is varied. The reason for this choice was that acid treatment of erythromycin, which yields the enol ether ring, gives the motilide activity and abolishes antibiotic activity. In addition, because the sugars have been shown to be required for activity against the receptor,[20] we did not alter this region of the macrocycle. The sugar-attachment enzymes or chemistry will be used to attach the appropriate sugars to all leads generated by this screening method. Thus, a library, named

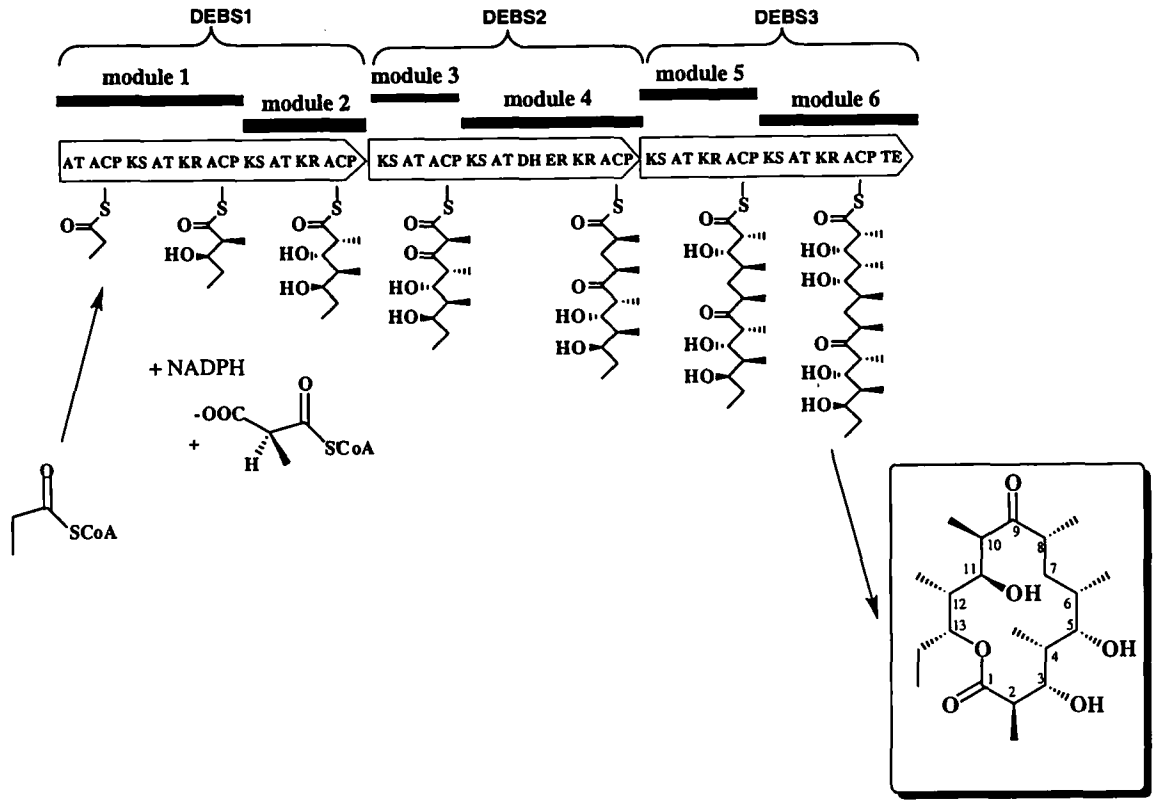


Figure 2. Modular organization of the 6-deB polyketide synthase. Functional domains of each of the three polypeptides from the DEBS PKS *gene cluster and modular organization are shown. Stepwise synthesis of 6-dEB begins at DEBS1 and ends with cyclization by the TE domain in DEBS3 to yield 6-dEB, which is functionalized further to yield erythromycin. KS, ketoacyl ACP synthase; AT, acyltransferase; DH, dehydratase; ER, enoylreductase; KR, ketoreductase; MT, methyltransferase, ACP, acyl carrier protein; TE, thioesterase.*

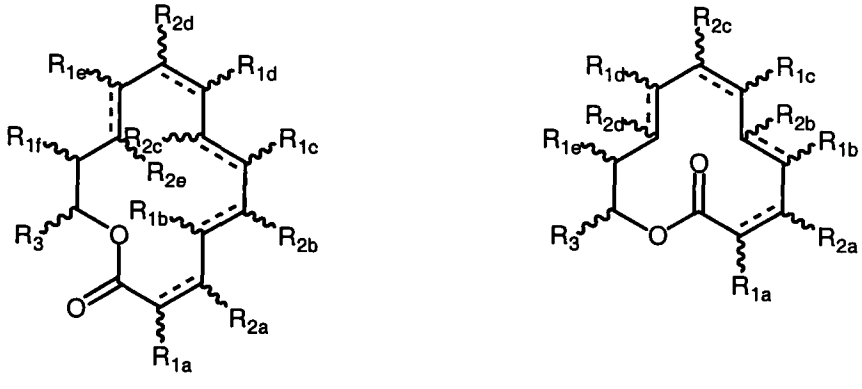


Figure 3. Generic 14-membered and 12-membered macrocycle libraries. $R1_{a\text{-}f}$ = {H,CH_3}; $R2_{a\text{-}e}$ = {H,OH,=O}; R3 = {CH_3,CH_2CH_3}. Although positions $R1_{a\text{-}f}$ are all derived from the same substituent set, they can vary independently. The same is true for $R2_{a\text{-}d}$. This yields the large number of entries in the library. Dashed lines represent either double or single bonds. All double bonds are trans.

C6-C12, of 14-membered macrolide rings with only the substituents between C6 and C12 varied from the 6-dEB was generated. The chirality at the methyl-substituted carbons was fixed to the enantiomer found in the native enol ether 6-dEB, because the current combinatorial biosynthetic toolbox does not include changing methyl stereochemistry.

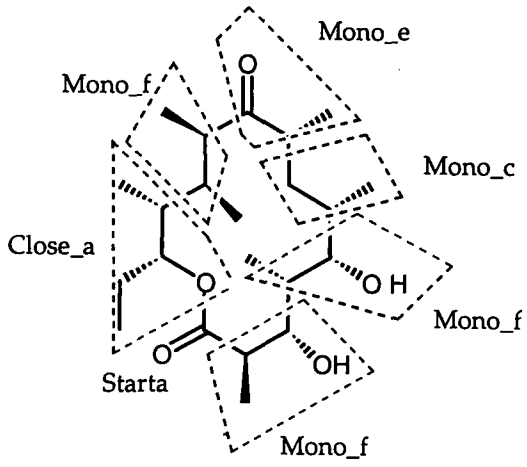
As in the CHORTLES representation of 6-dEB shown in Figure 4, the C6-C12 library has the Start and the next two monomers (Mono_f) fixed to match those of erythromycin. The Close monomer set contains two entries (Close_a and Close_c) because we have been interested in both the wild-type C12-methyl and C12-desmethyl substitutions. The final library contains 1 × 1 × 10 × 10 × 10 × 2 = 2,000 unique compounds, thus allowing conformational exploration of the region between C6-C12 (Figure 6).

The mixenum program was used to enumerate all 2,000 molecular graphs (SMILES) in this library.

Motilide Pharmacophore Model

Because the structure of the motilin receptor,[38] a G-protein–coupled receptor, is unknown, motilide agonists were used to

Figure 4. 6-dEB represented as a molecular graph in CHUCKLES monomer notation. CHUCKLES: Start_a12.Mono_f23.Mono_f34. Mono_c45.Mono_e56.Mono_f67.Close_a71 SMILES: C1(=O)-[C@H](C)[C@@H](OH)-[C@@H](C)[C@@H](OH)-[C@@H](C)C-[C@@H](C)C(=O)-[C@H](C)[C@@H](C)-[C@@H](C)[C@@H](CC)O1 The CHUCKLES monomers are attached via the indices that follow them. The attachment points within the monomers are specified in their definitions. In the SMILES version of the native erythromycin, adjacent monomers are attached via single (covalent) bonds as depicted by dashes. The cyclization bond is still represented by the index 1 adjacent to the Start and Close monomers.



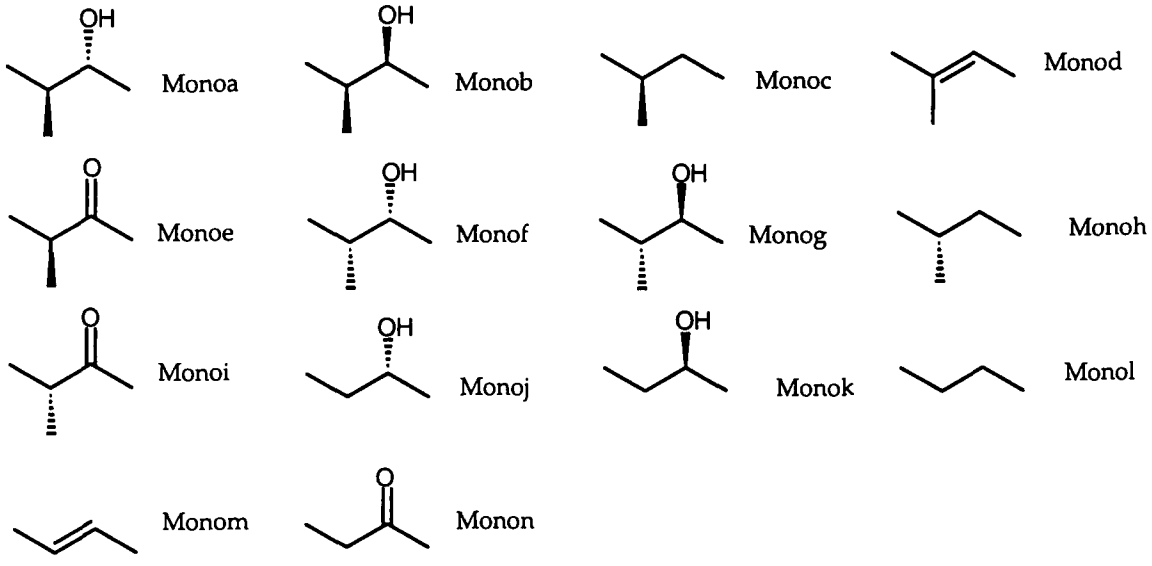


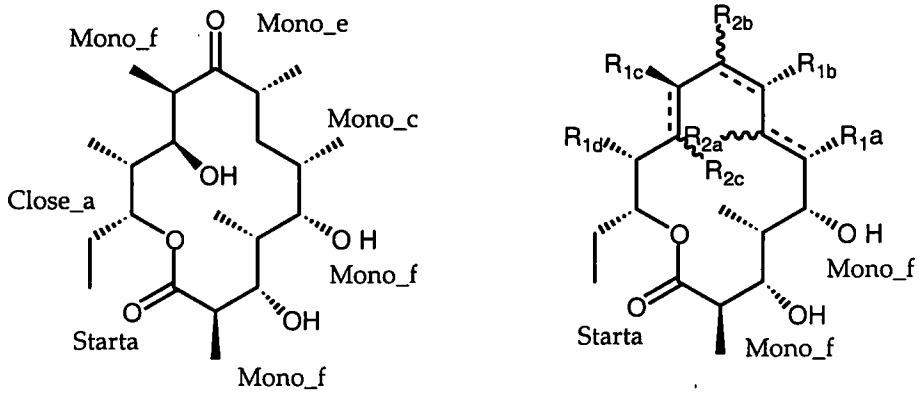
Figure 5. Two-carbon unit monomer representation for construction of virtual library.

develop our pharmacophore model. Two approaches were used to build the motilide pharmacophore based on known motilides. The first used literature compounds of various polyketides and erythromycin derivatives[25] in an attempt to build a CoMFA model. Given that the first approach failed to yield a viable model (see later), we attempted the second approach, which was designed to extract a simple 3D pharmacophore from the conformationally constrained enol ether erythromycin derivative.

Conventional pharmacophore based on enol ether analogue of erythromycin The five-membered enol ether ring introduces a significant conformational constraint on the flexible 14-membered erythromycin macrocycle. Because of the more limited conformational space, we decided to build a conventional pharmacophore based on presentation of certain constituents on the enol ether macrocycle. Based on agonist activity of erythromycin analogues in the literature, we know that the two sugars at the C3- and C5-positions are required for activity. In addition, the five-membered enol ether ring increases activity approximately 100-fold over the less constrained 14-membered macrocycle. Chemical modifications to other areas on the macrolide have not yielded dramatic changes in activity. Therefore, we incorporated into our pharmacophore model these three known features that contribute to binding activity.

Using the Daylight toolkits (Daylight Chemical Information Systems), we wrote a program to calculate the unit normal to the centroid of the enol ether five-membered ring and the vectors extending from C3 and C5 to their respective hydroxyl groups (as depicted in Color Plate 1). Our pharmacophore initially consisted of the distances between the centroid, and the C3 and C5 carbons. In addition, we calculated a unit volume as the cross-product between the three vectors. Together, these three distances and volume comprise our distance-volume space pharmacophore.

This model was refined further to accommodate the members of the focused virtual library that lack the enol ether ring and thus the centroid. Instead of the centroid and its normal, we calculated the unit vector to the C8 carbon, normal to C7-C8 and C9-C8 bonds. In the enol ether erythromycin analogue, this



R1a-d = {H, CH3} (chirality fixed to how it appears in native 6-dEB)
R2a-c = {OH, =O, H2, H (when double bond is present in macrocycle)}

Figure 6. A focused, virtual library, varied between the C6 and C12 carbons, represented in CHORTLES. The methyl substitutions between C6 and C12 are fixed to the original enantiomers in 6-dEB to limit the number of compounds in the library. The library was enumerated rapidly by the program mixenum *using Daylight SMILES and monomer toolkits. Start_a12. Mono_f23. Mono_f34. [Mono_a;Mono_b;Mono_c;Mono_d;Mono_e;Mono_j;Mono_k;Mono_l;Mono_m; Mono_n]45. [Mono_a;Mono_b;Mono_c;Mono_d;Mono_e;Mono_j;Mono_k;Mono_l;Mono_m;Mono_n]56. [Mono_d;Mono_f; Mono_g;Mono_h;Mono_i;Mono_j;Mono_k;Mono_l;Mono_m;Mono_n]67. [Close_a;Close_c]71.*

vector is parallel to the normal to the centroid because the ring is planar. In the focused virtual library screened against this pharmacophore, we expected a relatively rigid presentation of this vector to be consistent with the enol ether pharmacophore.

Ideally, the *distance-volume space* for the bound enol ether analogue would be derived from a crystal structure. However, there is no structural information for the compounds bound to the motilin receptor. Therefore, the pharamcophore comprises a set of points in the distance-volume space generated by the vectors described earlier for low-energy, minimized conformations of the enol ether analogue. After generating 2,000 random conformations using both rubicon (Daylight Chemical Information Systems) and the rubicon toolkit with its built-in conjugate gradient minimizer, we decided that the C3-C5 distance was not sufficiently interesting to include in the pharmacophore. Therefore, we defined the *distance-volume space* by the following three axes: C3-C8 distance, C5-C8 distance, and the three-vector unit volume.

Two thousand minimized random enol ether conformations fill a subset of the *distance-volume space* occupied by native 6-dEB, and this difference is most dramatic along the distance(C3,C8)-axis. However, there are many "outlier" coordinates. Additional minimization using Macromodel's BatchMin (Columbia University), yielded further, significant collapse of the points into more defined regions of *distance-volume space* (Color Plate 2). These coordinates were broken into a set of clusters that define the new pharmacophore. Some control calculations with the apricotl (Kosan Biosciences) program showed that four different enol ether analogues (randomly chosen from an enol ether 14-membered macrocycle virtual library) generated distance-volume coordinates that fell into the same space as the native enol ether analogue. An apricotl run on random 14-membered macrolide rings without an enol ether ring yielded distance-volume coordinates diffused throughout the space filled by the native erythromycin.

CoMFA pharmacophore based on overlaying known agonists The CoMFA (Comparative Molecular Field Analysis) program (Tripos Corporation, St. Louis, MO) was used as an alternative approach to build a pharmacophore from a set of erythromycin analogues. The goal was to derive a model that generated a hypothesis about what steric and electrostatic conditions in 3D space are contributing to activity. The CoMFA method consists of overlaying a series of molecules with a wide range of activities. The output is a three-dimensionsal contour map with polyhedra representing areas of desired charge or steric hindrance. This map allows one to predict the activity of novel molecules.

We selected 16 motilide analogues from an Abbott patent[25] with ED_{50} data for induction of *in vitro* rabbit duodenal smooth muscle contraction. Most modifications are to the 3′-N on the amino sugar, but these modifications are done consistently for three sets of macrocycle backbone modifications. The amino sugar modifications to 3′-$N(CH_3)_2$ include 3′-N-desmethyl, 3′-N-desmethyl-3′-N-ethyl, and 3′-N-propargyl. The backbone variations include the 8,9-anhydro-6,9-hemiketal (the standard enol ether), the 6-deoxy,9-deoxo-6,9-epoxy (which has an 8,9 single bond within the five-membered ring), the chiral enantiomer of the C8-methyl (8-epi), and the 12-OH (erythromycin A) and the 12-deoxy (erythromycin B). Activity data on all these variations are listed in Table 1.

The 3D models of each analogue were built based on the EACPAZ (8*R*,9*R*)-9-dihydro-6,8-anhydroerythromycin A crystal structure from the Cambridge Structure Database (CSD). Random conformers were minimized using standard *in vacuo* techniques using macromodel (Columbia University)[39] on an SGI O2 workstation. For the purpose of this calculation, the amino group on the C5-sugar is considered protonated, that is, tetrahedral. For minimization, all hydrogens were built on using Sybyl (Tripos Corporation), the Gasteiger partial charges were assigned, and Tripos force fields chosen. Minimization

Table 1. Lartey et al.[25] patent activity data on induction of *in vitro* rabbit duodenal smooth muscle contraction by erythromycin analogues

Lartey ID	pED_{50} (-log M)	Relative potency	Name[a]
1	8.41	363	8,9anhydro-erythromycin A-6,9hemiketal
2	11.26	257,039	8,9anhydro-erythromycin B-6,9hemiketal
3	9.74	7,762	8,9anhydro-3′-N-desmethyl erythromycin A 6,9hemiketal
4	11.5	188,526	8,9anhydro-3-N-desmethyl-3-N-ethylerythromycin A-6,9hemiketal
5	7.64	62	8,9anhydro-3′-N-propargylerythromycin A-6,9hemiketal
6	*b*	>8,000	8,9anhydro-3′-N-desmethylerythromycin B-6,9hemiketal
7	*b*	>8,000	8,9anhydro-3′-N-desmethyl-3′-N-ethylerythromycin B-6,9hemiketal
8	7.30	28	8,9anhydro-3′-N-propargylerythromycin B-6,9hemiketal
10	7.75	79	9-deoxo-3′-N-desmethyl-6-deoxy-8-epi-6,9-epoxyerythromycin A
11	8.54	490	9-deoxo-3′-N-desmethyl-6-deoxy-8-epi-3′-N-ethyl-6,9-epoxyerythromycin A
12	7.63	60	9-deoxo-6-deoxy-8-epi-6,9-epoxy-3′-N-propargylerythromycin A bromide
13	9.78	8,511	9-deoxo-6-deoxy-6,9-epoxyerythromycin A
14	7.15	60	9-deoxo-6-deoxy-3′-N-desmethyl-6,9-epoxyerythromycin A
15	7.63	35	9-deoxo-3′-N-desmethyl-6-deoxy-6,9-epoxy-3′-N-ethylerythromycin A
16	7.39	35	9-deoxo-6-deoxy-6,9-epoxy-3′-N-propargylerythromycin A bromide
	5.85	1	erythromycin A

[a]All Lartey samples are 4″-deoxy.
[b]Endpoints for samples 6 and 7 could not be obtained due to their extremely high potency.

consisted of 50 steps of steepest descent to remove van der Waals overlap, followed by a maximum of 1,000 steps of conjugate gradient. All structures were aligned in Sybyl6.4. The alignment/3D coordinates were read into CoMFA package (Tripos Corporation), and the activity data was read into a CoMFA table. The CoMFA hypothesis was built for predicting activity of unknowns.

Because of the limited number of analogue sites, the CoMFA software only generated a hypothesis around the amino sugar and the C-13 ethyl position. This approach cannot be used to test the virtual macrolide library because it tells us nothing about the pharmacophore contributions of the macrocycle. We conclude that the Lartey analogues are too similar and the wide range of activities (over five orders of magnitude) is not explained by the one or two atom differences among the analogues.

A further attempt involved incorporation of other macrocyclic compounds, such as roxithromycin, LY267,108, clarithromycin, oleandomycin, and native erythromycin, all with varying degrees of motilin agonist activity. Although the new CoMFA contour map was more detailed, the steric/charge hypothesis again was only generated around the amino sugar and the C13-ethyl. Because our molecules all contain similar macrocycles, we are unable to derive any hypotheses about this region. An additional attempt to incorporate the motilin peptide hormone nuclear magnetic resonance structure[13] into the CoMFA model was unsuccessful. An alignment of the peptide hormone with erythromycin-derivative motilides was published recently.[12]

Screening Virtual Polyketide Libraries Against Motilide Pharmacophore

Screening the C6-C12 virtual library against the motilide pharmacophore consisted of (1) generation of a set of minimized, random conformations, (2) calculating the *distance-volume* coordinates for the unique, low-energy conformations, then (3) screening against the enol ether erythromycin clusters in *distance-volume* space. For each SMILES entry in the library, 500 3D conformations with explicit hydrogens were generated by the distance geometry program rubicon (Daylight Chemical Information Systems). Each of these 500 conformations were processed further by Macromodel/BatchMin with 100 steps of conjugate gradient minimization, using the GB/SA solvation model and MM2* force field (Columbia University, New York, NY). The structures did not converge within the first 100 steps, but we wanted to "clean up" the less than optimal minimization provided by the rubicon program. Conformations in the bottom 10 kcal/mol (42 kJ/mol) were fed into apricot for calculation of distance-volume coordinates (Kosan Biosciences). These distance-volume coordinates were saved for each low-energy conformation and then matched against the motilide pharmacophore by the program clusterscan (Kosan Biosciences).

RESULTS

Descriptor Set that Distinguishes Between Enol Ether and Erythromycin

The *clusterscan* algorithm was able to group the distance-volume coordinates for the enol ether 6-dEB into a set of boxes containing adjacent coordinates in 3D space. In order to closely contain the coordinates, 33 clusters, with some overlap, were calculated. A visual examination of the distance-volume coordinates of the enol ether conformations confirmed the clustering. Furthermore, the distance range between the C5 and C8 carbons is more constrained in the enol ether analogue than in the erythromycin (Color Plate 2). The distance-volume coordinates of the enol ether erythromycin analogue cluster in the

distance-volume space, whereas the more flexible erythromycin macrocycle yields low-energy conformations that are scattered more evenly throughout the distance-volume space. We observed a band along the C5 to C8 distance axis for the enol ether, whereas the erythromycin coordinates cover a wide range, sandwiching in the enol ether band (Color Plate 2). Members of a separate library of enol ether analogues (all containing the five-membered ring) showed similar clustering based on the two distances and vector cross-product volume.

Verification of Pharmacophore Model

In order to verify the pharmacophore model, in spite of the lack of information about the bound, active conformation, three known, potent motilides were subjected to our analysis. Faghih et al.[40] developed a set of compounds in an attempt to increase acid stability. ABT-229 (compound **VII**), which lacks the C12-hydroxyl group, giving the macrocycle more flexibility than the native 6,9-enol ether motilide, yielded 42 hits out of 160 low-energy conformers, a 26.25% success against the distance-volume cluster pharmacophore. The single enantiomer of reduced erythromycin A, (9*R*),(8*S*)-9-deoxo-4″-deoxy-3′-N-desmethyl-3′-N-ethyl-6,9-epoxyerythromycin A (compound **IX**),[40] yielded 16 hits out of 89 low-energy conformers, for a 17.98% success against the distance-volume cluster pharmacophore. Compound **VII** has a pED_{50} = 7.22, whereas compound **IX** has a pED_{50} = 8.54. This increase presumably is due to the single enantiomer **IX**, reduced compound achieving a better conformation for binding to the receptor. GM-611 (compound **VI**) yielded 34 hits out of 92 low-energy conformers, for a 36.96% success against the distance-volume cluster pharmacophore. Aside from the unusual library compound no. 1323, these known motilides show significantly better rates of matches against the pharmacophore than any of the components of the library.

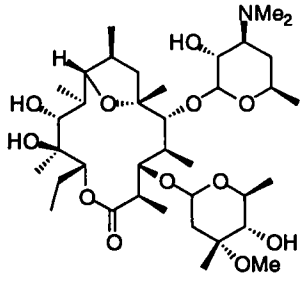
Although erythromycin and most of the C6-C12 library compound distance-volume coordinates were spread throughout the distance-volume space, it is possible to sort library compounds based on the number of conformations that fall into the motilide pharmacophore clusters. Certain compounds with more constrained rings (due to substitution at every position or introduction of a double bond into the macrocycle) exhibit clustering of their coordinates in distance-volume space. Several of these macrolactones show significant overlap with the enol ether distance-volume coordinate clusters *(vida infra)*.

Novel Leads that Fit the Pharmacophore

Because conformation generation and minimization of these compounds takes about 20 minutes/500 conformations of each compound on an SGI O2 R10000, all 2,000 unique entries were processed in 27.8 days. The criteria used to sort the processed compounds included (1) constrained ring systems (as determined by a small number of low-energy conformations generated by rubicon and minimized by macromodel; (2) high percentage (>10%) of conformations that match the distance-volume space of the enol ether motilide pharmacophore; and (3) only one or two mutations from the native erythromycin macrolide. Some of the combinations of monomers in the macrocyle yield fewer conformations and are therefore more rigid or constrained (Table 2). Furthermore, some of these more-constrained compounds match the enol ether erythromycin-based distance-volume pharmacophore.

Based on our criteria, three sets of novel compounds were found in the macrolactone virtual library. Interestingly, a total of 580 compounds of our 2000 member library did not overlap the enol ether distance-volume clusters at all.

Three classes of polyketide macrolactones were identified that are consistent with our criteria for overlapping the enol ether conformations in distance-volume space (Figure 7). These compounds are novel polyketide structures that have the potential to be biosynthesized by Kosan's directed PKS genetic modification procedures. The leads can be divided into three groups: (1) compounds that contain a C6-methylene, C7-keto, and C8-C9 *trans* double bond; (2) compounds that contain a C6-methyl, C7-keto, and C8-C9 *trans* double bond; and (3) a small set of compounds with various substitutions at all of the C6-C10 positions. An interesting compound that does not fit into any of the three groups is no. 1323, which, despite C6 and C8 methylenes, yields only four low-energy conformations by our methods. Because one of these conformations falls into the enol ether distance-volume space, it has the highest percentage (25%) of good lead conformations (Table 2).



IX, (9*R*),(8*S*)-9-deoxo-4"-deoxy-3'-N-desmethyl-3'-N-ethyl-6,9 epoxyerythromycin A

Table 2. Novel library components exhibiting the highest overlap with the enol ether 6-dEB pharmacophore

Compound ID (group ID)	Number of conformations in clusters	Total low-energy conformations	Percentage of conformations overlapping enol ether clusters	Monomers different from native 6-dEB
880 (2)	14	151	9.27%	4
8 (3)	10	107	9.35%	4
1868 (1)	8	84	9.52%	4
1872 (1)	21	218	9.63%	4
1863 (1)	18	186	9.68%	2
1880 (1)	18	178	10.11%	4
871 (2)	14	134	10.45%	3
863 (2)	9	84	10.71%	2
872 (2)	13	115	11.30%	4
1864 (1)	15	129	11.63%	3
865 (2)	9	76	11.84%	3
1866 (1)	23	178	12.92%	4
1873 (1)	15	112	13.39%	3
866 (2)	18	133	13.53%	4
868 (2)	12	74	16.22%	4
1323	1	4	25.00%	2

Group ID corresponds to depictions in Figure 7.

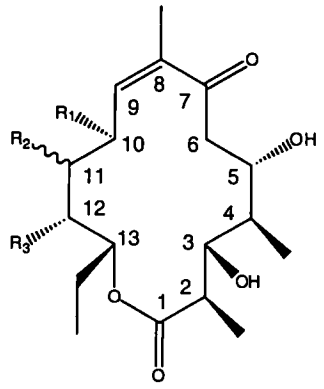
DISCUSSION

To our knowledge, this is the first published example of screening a large virtual library of macrolactone polyketides.

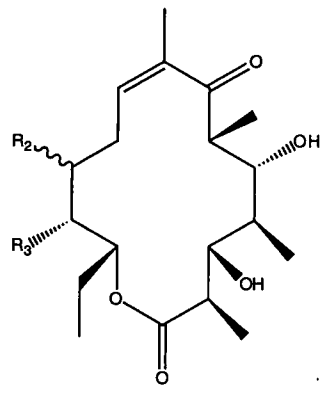
Although two methods were utilized to generate the virtual polyketide libraries, both yield the same set of compounds and, thus, the same set of conformations used to screen against the pharmacophore. The reaction-based method is true to the chemistry of the biosynthetic pathway, where each reaction step represents the action of an enzyme in the PKS modules. The starting materials and intermediates are actual compounds found during synthesis. Thus, generation of new libraries is accommodated rapidly by rearrangement of the appropriate reactions, corresponding to the rearrangement of enzymes in the biosynthetic pathway. Smaller macrocycles are generated simply by applying fewer rounds of the reaction methods. The alternative CHORTLES-based method is focused on the monomeric unit and its connectivity to other units. The library of macrocycles comprises sets of monomers in a specific connected graph. This representation compactly represents the entire library as a single expression depicting connectivity between a set of monomers. The monomers at each position can be shuffled quickly or substituted to generate a new library. Smaller macrocycles are generated by removing nodes in the connected graph. Both methods are efficient and reliable mechanisms for generating biopolymer virtual libraries. The reaction-based method is more intuitive from a synthetic or genetic viewpoint, whereas the subunit-based method is more intuitive from a polymer viewpoint.

The basic macrocyclic polyketide core consists mainly of monomers with simple chemical moieties (methyl, hydroxyl, olefin, and ketone functionalities) that form a structural as well as functional framework. We generally think of macrocyclic polyketides as combinatorial scaffolds with limited, although significant, ability to interact with the target (much like a peptide backbone). However, the macrocycle can form a rigid core, optimally directing attached moieties to make specific receptor interactions. Some of the chemical decorations commonly incorporated into polyketides include sugars, secondary hydroxyl groups, internal ring closures (e.g., 6,9-enol ether five-membered ring in the motilide erythromycin analogue), and methyl groups. Further, recent work showed that chemical handles can be incorporated into the macrocycle allowing combinatorial attachment of macrocycle decorations. These handles provide significant modifications yielding a wide variety of potential modes for interaction with a receptor. In general, the macrocycles are analogous to cyclic peptides with a structural backbone and functional side chains (sugars, etc.). By this same analogy, the specific set of polyketides we investigated is similar to a set of cyclic peptides with limited numbers of side chains, but in which we have a very high degree of control of the backbone conformation that controls the direction of side-chain interactions.

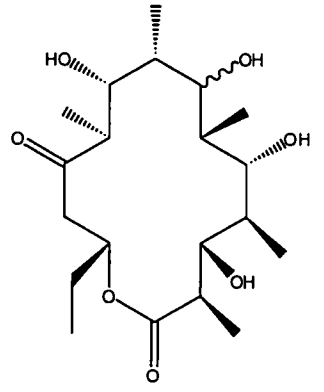
Alternatives to the computationally expensive enumeration and generation of multiple low-energy conformations were considered. For exploration of the low-energy conformations, we considered (1) piecewise construction, using minimized 3D conformations of monomers attached and then minimized in the context of the macrocycle; (2) systematic torsion searching about the bonds in the macrocycle; (3) molecular dynamics followed by extensive minimization; and (4) distance geometry methods to generate random conformations, which satisfy interatomic distances of a given seed structure, followed by short minimization. Approaches 1 and 2, hindered by the large number of atoms in the macrocycle, might be extremely inefficient when taking ring closure in 3D space into consideration. Method 3 would require extensive analysis to determine whether it is appropriate for exploring conformational space on 14-membered rings and would entail minimization to conver-



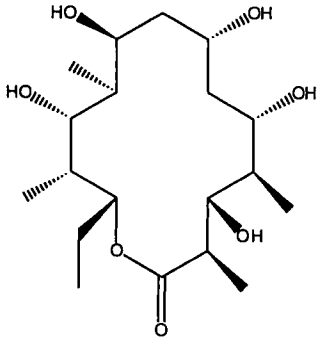
Group 1



Group 2



Group 3



Compound 1323

Figure 7. Compounds identified to be consistent with the enol ether conformations in distance-volume space. These compounds are novel polyketide structures that have the potential to be biosynthesized by Kosan's directed polyketide synthase genetic modification procedures. Macrocyclic double bonds are trans *(although they are drawn as cis).*

gence to ensure finding unique low-energy conformations. Based on our experience, distance geometry methods would lead to the most random sampling of conformational space. Although generation of 500 random conformations takes several minutes for each molecule, it must be followed by a reasonable minimization (see Methods) to yield low-energy conformations.

Consider more efficient methods such as genetic algorithms[41] to evolve a set of compounds that meet the criteria of a pharmacophore: the genetic algorithm would save us from enumerating all the compounds in the library. However, due to the need for multiple generations (sometimes hundreds) of crossover, mutation, and selection in the genetic algorithm, one needs a rapid screen. In the macrocyclic polyketides, the large flexible, undecorated rings require extensive examination of low-energy conformations to determine accurately the two vectors and volume that comprise the motilide pharmacophore. Due to the large number of potential conformations of these 14-membered rings, one cannot readily or accurately derive the values that define the pharmacophore, and one cannot infer these values from the sequence of monomers. The monomers, which further constrain the ring system (e.g., *trans* double bond in Monom) have somewhat predictable local effects, but they have more subtle effects translated about the ring. In addition, ring closure would be difficult to derive if one is representing torsion angles in a bit-string. For smaller ring systems and aromatic polyketides, a simpler approach to evaluating members of a virtual library might be feasible.

It is clear from the literature of active and inactive polyketide motilides that the three critical chemical motifs for activity are the two sugar rings and the region of C8-C9 of the macrocycle. Many of the compounds in the literature focus on modifications of the sugars, which are significantly easier to modify chemically than the macrocycle. In contrast, the biosynthetic technology we use has the unique ability to allow modifications of the entire macrocycle. In order to focus on the novel biosynthetic possibilities, we limited our potential library to include the native sugars. Nevertheless, if an active macrolide with native sugars were identified, chemical optimization of the sugar moieties could occur at a later stage.

In construction of the motilide pharmacophore for screening this unique biosynthetic library, we used two considerations: (1) What are the critical binding features?; and (2) What is the minimal necessary description of them? The three critical fea-

tures for motilide activity are the two sugars and the C8-C9 region of the macrocycle. In order to focus on the macrolide library, we assumed that the changes in the macrocycle could be varied independently from the variation of the sugar analogue series. Thus, the critical effect of the macrocycle variation on *any* attached sugar can be optimized by the direction in which the bridging ether oxygen orients the sugar to interact with the receptor. This orientation information was encoded in our pharmacophore with the C3-C5 distance and the C3-hydroxyl vector and the C5-hydroxyl vector. Although chemical modification of the macrocycle is difficult, current data point to the region of the C8 and C9 carbons. The actual chemical moiety in the C8-C9 region that optimizes binding has not been explored thoroughly with structural variations related to those in the macrocylic virtual library, so specific chemical specifications in this region could not be incorporated in the pharmacophore. Instead, we incorporated structural information about the relative position of the C8-C9 compared to the sugar orientations. To geometrically encode the relative location of these moieties, we included the C3-C8 and C5-C8 distances. To geometrically encode the relative orientation of the three moieties, we calculated a chiral volume formed by unit vectors perpendicular to the plane of C7-C8-C9 and along the C3-OH and C5-OH vectors. Thus, by considering the limited activity data pertinent to our macrocycle library and by focusing on the novel macrolactone variations feasible with current biosynthetic technology, it was possible to distill a macrolactone pharmacophore with three distances and one chiral volume. If we had planned to screen a large corporate database of structurally unrelated compounds, a significantly more detailed pharmacophore would have been necessary. However, by analyzing the nature of variation in the macrocycle library, it was possible to simplify our pharmacophore representation.

Careful analysis of the best hits is possible by examination of one of their 3D conformations overlaid on the corresponding enol ether 6-dEB conformation (Color Plate 3). Different conformations of each of the compounds correspond to different points in the motilide distance-volume space as matched by clusterscan. Interestingly, two compounds (no. 1873 from group 1 and no. 868 from group 2) show excellent alignment with the same low-energy motilide conformation. The C7-keto in combination with the C8-C9 *trans* double bond positions the normal to the C8 very similarly to that found in the low-energy conformation of the enol ether ring, as shown in Color Plate 3a and b. Although the group 1 compounds lack the C6-methyl substitution, they readily align along the entire C3-C6 portion of the ring. The compounds in group 2 contain the C6-methyl substituent, which aligns beautifully with the C6-methyl found in the enol ether motilide. These conformations illustrate the success with which the clusterscan algorithm matched virtual library compounds with the enol ether motilide distance-volume clusters. The figures illustrate the divergence of the C11-C13 and lactone portions of the ring that are not considered when screening conformations against the pharmacophore.

Compounds in group 3 and compound no. 1323 contain hydroxyl substituents at the tetrahedral C7 position, instead of the ketone seen in groups 1 and 2. Rather than rigidifying the C7-C9 portion of the ring (as seen for groups 1 and 2), the flexibility allows the alignment with structurally different low-energy enol ether motilide conformers. Despite the lack of methyl substituent on the C8 in compound no. 1323, the distance-volume calculation defining the normal to C8 as the normal to the C7-C8 and C8-C9 bonds is successful in matching the normal to C8 in the enol ether motilide.

It is noteworthy that the methodologies described herein have identified a structural change that previously was used to potentiate the motilide activity of erythromycin analogues. Both GM-611[24] and compounds in group 3 (Figure 7) contain an 11-ketone rather than the hydroxyl group found in 6-dEB and erythromycin. Whereas most of the structural alterations of 6-dEB identified in the target motilide compounds are not accessible by chemical transformation, oxidation of the 11-OH in the context of the 6,9-enol ether of erythromycin has been done.[24,42] Identification of this functionality through screening of the virtual 6-dEB analogue library against the motilide enol ether pharmacophore provides validation for pharmacophore model and screening methodology described here.

Additional analysis may be necessary to narrow down these candidates to a single candidate; for example, one could process a larger number of conformations (say 10,000) that have been minimized to convergence. One also could take into account conformational information involved in the reduced compounds (**IX**) synthesized by Faghih et al.,[40] which has been shown to be a more potent motilide than the native enol ether. In addition, one could further refine the pharmacophore with information about the bound conformation of the enol ether analogue or the motilin peptide.

The distance-volume pharmacophore was developed specifically for macrolides, because that is the library on which we are focusing our combinatorial efforts. Most of the literature compounds focus on modifications outside of the macrolide ring (e.g., on the desosamine nitrogen), and these modifications do not affect the values for the distance-volume calculation. For this reason, a CoMFA model could not be generated for these molecules. Developing statistical confidence limits or a chi-squared test of the pharmacophore based on a larger group of both active and inactive compounds would allow a more quantitative analysis of proposed compounds. However, there were a limited number of appropriate compounds tested against this receptor. In the future, when more is known about the bound conformation (including sugars) or when we have more binding information about novel macrocycles, a statistical model will be developed.

Synthesis of the molecules described in Figure 7 and Table 2 is, by design, amenable to polyketide combinatorial biosynthesis technology.[30,43,44] This technology uses genetic engineering to alter the assembly-line–like PKS enzymes, which serve both as templates and catalysts for polyketide chain assembly. There is significant precedent for the directed engineering of strains that express novel target molecules as fermentation products. In particular, the enzyme responsible for 6-dEB biosynthesis has been rationally engineered to create an extensive library of 6-dEB analogues, each of which was a member of the complete virtual library.[34] Although none of analogues suggested in this work have been reported, their generation via manipulation of the genes responsible for 6-dEB biosynthesis is a straightforward application of known technology.

Once the appropriate PKSs have been engineered to produce the desired novel 6-dEB analogues, there are several options for elaborating them to erythromycin analogues.[44,45]

CONCLUSION

Screening of a virtual library can rapidly yield conformationally constrained drug leads that satisfy a well-defined pharmacophore. Via engineering of the DEBS PKS genes, one can theoretically biosynthesize any of the leads.

The advantage of generating and screening an explicit library against a given pharmacophore model is that one is forced to consider all possible macrocycles in the library. Traditional medicinal chemistry approaches to developing potent motilides by synthesizing erythromycin analogues have been thoroughly explored.[24-27,40,46] However, these approaches are dramatically limited by the available macrolactone chemistry. In contrast, the biosynthetic technology described here offers access to nearly every position in the macrolactone ring. We chose to match the scope of our compound exploration to the potential scope of the biosynthesis by considering a large library of macrocycles, rather than a limited set of conservative analogues.

Further testing of the model will involve the generation of target molecules and their evaluation in assays that measure motilin receptor binding and agonism.

REFERENCES

1 Gallop, M.A., Barrett, R.W., Dower, W.J., Fodor, S.P.A., and Gordon, E.M. Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. *J. Med. Chem.* 1994, **37,** 1233–1251

2 Zuckermann, R.N., Martin, E.J., Spellmeyer, D.C., Stauber, G.B., Shoemaker, K.R., Kerr, J.M., Figliozzi, G.M., Goff, D., Siani, M.A., Simon, R.J., Banville, S.C., Brown, E.G., Wang, L., Richter, L.S., and Moos, W.H. Discovery of nanomolar ligands for 7-transmembrande g-protein coupled receptors from a diverse (N-substituted)-glycine peptoid library. *J. Med. Chem.* 1994, **37,** 2678–2685

3 Gordon, E.M., Barrett, R.W., Dower, W.J., Fodor, S.P.A., and Gallop, M.A. Applications of combinatorial technologies to drug discovery. 2. combinatorial organic synthesis, library screening strategies, and future directions. *J. Med. Chem.* 1994, **37,** 1385–1401

4 Dewitt, S. Diversomers: An approach to nonpeptide, nonoligomeric chemical diversity. *Proc. Natl. Acad. Sci. U.S.A.* 1993, **90,** 6909–6913

5 Bunin, B.A., Plunkett, M.J., and Ellmann, J.A. The combinatorial synthesis and chemical and biological evaluation of a 1,4-benzodiazepine library. *Proc. Natl. Acad. Sci. U.S.A.* 1994, **91,** 4708–4712

6 Thompson, L.A., and Ellman, J.A. Synthesis and applications of small molecule libraries. *Chem. Rev.* 1996, **96,** 555–600

7 Martin, E.J., Blaney, J.M., Siani, M.A., Spellmeyer, D.C., Wong, A.K., and Moos, W.H. Measuring diversity: Experimental design of combinatorial libraries for drug discovery. *J. Med. Chem.* 1995, **38,** 1431–1436

8 Warr, W.A. Combinatorial chemistry and molecular diversity: An overview. *J. Chem. Inf. Comput. Sci* 1997, **37,** 134–140

9 Brown, R.D., and Martin, Y.C. Designing combinatorial library mixtures using a genetic algorithm. *J. Med. Chem.* 1997, **40,** 2304–2313

10 Shemetulskis, N.E., Dunbar, J.B., Dunbar, B.W., Moreland, D.W., and Humblet, C. Enhancing the diversity of a corporate database using chemical database clustering and analysis. *J. Comput-Aid. Mol. Design* 1995, **9,** 407–416

11 Peeters, T.L., Macielag, M.J., Depoortere, I., Konteatis, Z.D., Florance, J.R., Lessor, R.A., and Galdes, A. D-amino acid and alanine scans of the bioactive portion of porcine motilin. *Peptides* 1992, **13,** 1103–1107

12 Khiat, A., and Booulanger, Y. Identification of the motilide pharmacophores using quantitative structure activity relationships. *J. Peptide Res.* 1998, **52,** 321–328

13 Jarvet, J., Zdunek, J., Damberg, P., and Graslund, A. Three-dimensional structure and position of porcine motilin in sodium dodecyl sulfate micelles determined by 1H NMR. *Biochemistry* 1997, **36,** 8153–8163

14 Beavers, M.P. Discovery of a small molecule motilin-antagonist through electronic database screening. In: *Innovative Computational Applications.* 1999, Institute for International Research (IIR)

15 Brown, J.C., Cook, M.A., and Dryburgh, J.R. Motilin, a gastric motor activity stimulating polypeptide: The complete amino acid sequence. *Can. J. Biochem.* 1972, **51,** 533–537

16 Miller, P., Gagnon, D., Dickner, M., Aubin, P., St.-Pierre, S., and Poitras, P. Structure-function studies of motilin analogs. *Peptides* 1995, **16,** 11–18

17 Peeters, T. Motilide web site: http://www.med.kuleuven.ac.be/med/gih/motilid.htm

18 Peeters, T.L. Erythromycin and other macrolides as prokinetic agents. *Gastroenterology* 1993, **105,** 1886–1899

19 Omura, S., Tsuzuki, K., Sunazuka, T., Marui, S., Toyoda, H., Inatomi, N., and Itoh, Z. Macrolides with gastrointestinal motor stimulating activity. *J. Med. Chem.* 1987, **30,** 1941–1943

20 Depoortere, I., Peeters, T.L., Matthijs, G., Cachet, T., Hoogmartens, J., and Vantrappen, G. Structure-activity relation of erythromycin-related macrolides in inducing contractions and in displacing bound motilin in rabbit duodenum. *Gastrointestinal Motility* 1989, **1,** 150–159

21 Lartey, P.A., Nellans, H.N., Faghih, R., Petersen, A., Edwards, C.M., Frieberg, L.A., Quigley, S., Marsh, K., Klein, L.L., and Plattner, J.J. Synthesis of 4″-deoxy motilides: Identification of a potent and orally active prokinetic drug candidate. *J. Med. Chem.* 1995, **38,** 1793–1798

22 Tsuzuki, K., Sunazuka, T., Marui, S., Toyoda, H., Omura, S., Inatomi, N., and Itoh, Z. Motilides, macrolides with gastrointestinal motor stimulating activity. 1. O-substituted and tertiary N-substituted derivatives of 8,9-anhydroerythromycin A hemiacetal. *Chem. Pharm. Bull.* 1989, **37,** 2687–2700

23 Funabashi, Y., Maeshiba, Y., Inatomi, N., Tanayama, S., Harada, S., Itoh, Z., and Omura, S. Bioactive metabolites of EM574 and EM523, erythromycin derivatives having strong gastrointestinal motor stimulating activity. *J. Antibiotics* 1996, **49,** 794–801

24 Koga, H.T., Sato, K., Tsuzuki, H., Onoda, H., Kuboniwa, H., and Takanashi, H. Potent, acid-stable and orally active macrolide-type motilin receptor agonists, GM-611 and other derivatives. *Bioorg. Med. Chem. Lett.* 1994, **4,** 1347–1352

25 Lartey, P.A., Klein, L.L., Faghih, R., and Nellans, H.N., Abbott Laboratories, 4″-deoxyerythromycin derivatives, USA 5,578,579, November 26, 1996

26 Verhagen, M., Samsom, M., Maes, B., Geypens, B.J., Ghoos, Y.F., and Smout, A. Effects of a new motilide, ABT-229, on gastric emptying and postprandial antroduodenal motility in healthy volunteers. *Aliment. Pharmacol. Ther.* 1997, **11,** 1077–1086
27 Greenwood, B., Dieckman, D., Kirst, H.A., and Gidda, J.S. Effects of LY267108, an erythromycin analogue derivative, on lower esophageal sphincter function in the cat. *Gastroenterology* 1994, **106,** 624–628
28 Bormans, V., Peeters, T.L., and Vantrappen, G. Motilin receptors in rabbit stomach and small intestine. *Regul. Pept.* 1988, **15,** 143–153
29 Suzuki, H., Mochiki, E., Haga, N., Satoh, M., Mizumoto, A., and Itoh, Z. Motilin controls cyclic release of insulin through vagal cholinergic muscarinic pathways in fasted dogs. *Am. J. Physiol.* 1998, **274,** G87–G95
30 Hutchinson, C.R. Microbial polyketide synthases: More and more prolific. *Proc. Natl. Acad. Sci. U.S.A.* 1999, **96,** 3336–3338
31 Carreras, C.W., and Santi, D.V. Engineering of modular polyketide synthases to produce novel polyketides. *Curr. Opin. Biotechnol.* 1998, **9,** 403–411
32 Verdine, G.L. The combinatorial chemistry of nature. *Nature* 1996, **384,** 11–13
33 Tsoi, C.J., and Khosla, C. Combinatorial biosynthesis of "unnatural" natural products: The polyketide example. *Curr. Biol.* 1995, **2,** 355–362
34 McDaniel, R.M., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., Betlach, M., and Ashley, G. Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc. Natl. Acad. Sci. U.S.A.* 1999, **96,** 1846–1851
35 Siani, M.A., James, C., Weininger, D., and Blaney, J.M. CHORTLES: A method for representing oligomeric and template-based mixtures. *J. Chem. Inf. Comput. Sci.* 1995, **35,** 1026–1033
36 Siani, M.A., Weininger, D., and Blaney, J.M. CHUCKLES: A method for representing and searching peptide and peptoid sequences. *J. Chem. Inf. Comp. Sci.* 1994, **34,** 588–593
37 Weininger, D. SMILES, a chemical language and information system 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* 1988, **28,** 31–36
38 Feighner, S.D., Tan, C.P., McKee, K.K., Palyha, O.C., Hreniuk, D.L., Pong, S.S., Austin, C.P., Figueroa, D., MacNeil, D., Cascieri, M.A., Nargund, R., Bakshi, R., Abramovitz, M., Stocco, R., Kargman, S., O'Neill, G., Van Der Ploeg, L.H., Evans, J., Patchett, A.A., Smith, R.G., and Howard, A.D. Receptor for motilin identified in the human gastrointestinal system. *Science* 1999, **284,** 2184–2188
39 Allinger, N.L. Conformational analysis 130 MM2 a hydrocarbon force field utilizing V1 and V2 torsional terms. *J. Am. Chem. Soc.* 1977, **99,** 8127
40 Faghih, R., Lartey, P.A., Nellans, H.N., Seif, L., Burnell-Curty, C., Klwin, L:, Thomas, P., Petersen, A., Borre, A., Pagano, T., Kim, K.H., Heindel, M., Bennani, Y.L., and Plattner, J.J. Synthesis of 9-deoxo-4″-deoxy-6,9-epoxyerythromycin derivatives: Novel and acid-stable motilides. *J. Med. Chem.* 1998, **41,** 3402–3408
41 Weber, L. Applications of genetic algorithms in molecular diversity. *Curr. Opin. Chem. Biol.* 1998, **2,** 381–385
42 Koga, H., Sato, T., and Takanashi, H., Chugai Seiyaku Kabushiki Kaisha, Erythromycin derivatives, USA 5,658,888, August 19, 1997
43 Katz, L., and McDaniel, R. Novel macrolides through genetic engineering. *Med. Res. Rev.* 1999, **19,** 543–558
44 Carreras, C.W., Ou, S., Frykman, S., Cadapan, L., Zavala, S., Woo, E., Leaf, T., Carney, J., Burlingame, M., Patel, S., Ashley, G., and Licari, P. Saccharopolyspora erythraea-catalyzed conversion of 6-deoxyerythronylide B analogs for production of novel erythromycins. 2000 (submitted)
45 Donadio, S., McAlpine, J.B., Sheldon, P.J., Jackson, M., and Katz, L. An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 1993, **90,** 7119–7123
46 Depoortere, I., Peeters, T.L., and Vantrappen, G. The erythromycin derivative EM-523 is a potent motilin agonist in man and in rabbit. *Peptides* 1990, **11,** 515–519