



PUBLICATIONS



PATENTS



PRESENTATIONS



JULY 2001

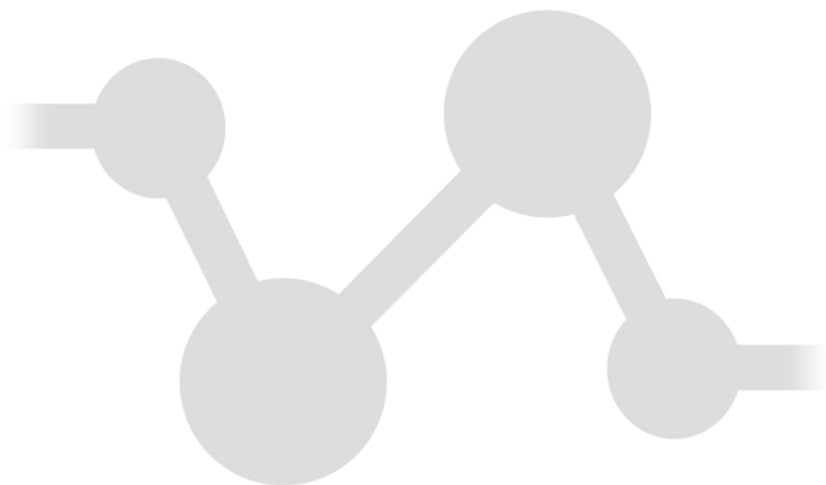


**Albany Molecular Research, Inc.<sup>®</sup>**



Albany Molecular Research, Inc.

## PUBLICATIONS





## HUDKINS, R. L.; D'AMBRA, T. E.

Recent Advances in the Medicinal Chemistry of Cannabinoids. *Curr. Opin. Med. Chem. Cannabinoids* 1993, 3, 403-416.

Preparations of *Cannabis sativa* (marijuana) have a long history as therapeutic agents. Marijuana contains numerous chemical constituents, but the active constituent (-)- $\Delta^9$ -6a, 10a-*trans*-tetrahydrocannabinol (1,  $\Delta^9$ -THC) is responsible for its main pharmacological effect.  $\Delta^9$ -THC is also referred to as  $\Delta^1$ -THC using numbering based on the monoterpene moiety (2).  $\Delta^8$ -THC ( $\Delta^6$ -THC) is also a physiologically active isomer which is formed in a few varieties of the plant. The (+)-isomers, in addition to the two other (-)-isomers with a 6a, 10a-*cis* ring junction (*cis*-(-)-  $\Delta^9$ -THC and *cis*-(-)-  $\Delta^8$ -THC), are pharmacologically weaker or inactive, demonstrating strict stereochemical requirements.  $\Delta^9$ -THC and the more potent metabolite, 11-hydroxy- $\Delta^9$ -THC (3), produce a wide variety of pharmacological effects in both laboratory animals and humans.

## KHMELNITSKY, Y. L.; WELCH, S. H.; CLARK, D. S.; DORDICK, J. S.

Salts Dramatically Enhance Activity of Enzymes Suspended in Organic Solvents. *Journal of the American Chemical Society* 1994, 116 (6), 2647-2648.

The catalytic efficiency ( $k_{\text{cat}}/K_m$ ) of subtilisin Carlsberg and  $\alpha$ -chymotrypsin in anhydrous organic solvents is dramatically increased when the enzyme is lyophilized in the presence of excess salts. For example, a biocatalyst powder containing 98% (w/w) KCl, and 1% (w/w) each of phosphate buffer and subtilisin was 3,750-times more active in *n*-hexane for the transesterification reaction between N-Ac-L-Phe-OEt and 1-propanol than the enzyme prepared without KCl. This activation was primarily due to a large increase in the catalytic turnover ( $k_{\text{cat}}$ ) and provides for catalytic efficiencies in *n*-hexane to within an order of magnitude of that obtained for hydrolytic reactions in aqueous solutions. The activation was also observed in other organic solvents as well as with other salts indicating that the effect is general. Enzymatic catalysis in the gas phase is not dramatically affected by the presence of salt in the lyophilized mixture and this suggests that salt does not act as a lyoprotectant, but rather protects the enzyme from direct interaction (and inactivation) by the organic solvent. We hypothesize that a highly polar salt matrix is formed that excludes direct solvent contact with the enzyme and helps to maintain the native structure of the enzyme.

**OPALKA, C. J.; D'AMBRA, T. E.; FACCONI, J. J.;  
BODSON, G.; COSSEMENT, E.**

A Novel Synthesis of the Enantiomers of an Antihistamine Drug by Piperazine Formation from a Primary Amine. *Synthesis* **1995**, 766-768.

An enantioselective synthesis of each enantiomer of the antihistamine drug 2-(2-{4-[(4-chlorophenyl) phenylmethyl]-1-piperazinyl}ethoxy)acetic acid dihydrochloride (**1**) is described, involving the preparation of the benzhydrylpiperazine portion of the molecule from reaction of each enantiomer of 4-chlorobenzhydrylamine with *N,N*-bis(2-chloroethyl)-4-methylbensenesulfonamide. A modification of standard toluenesulfonamide deprotection with hydrogen bromide in acetic acid was introduced, substituting 4-hydroxybenzoic acid for phenol.

**RICH, J.; BEDELL, B.; DORDICK, J.**

Controlling Enzyme-Catalyzed Regioselectivity in Sugar Ester Synthesis. *Biotechnol. and Bioeng.* **1995**, *45*, 426-434.

The rational control over enzyme-catalyzed regioselectivity has been studied using sucrose acylation by vinyl esters in organic media as a model. Subtilisins BPN' and Carlsberg preferentially acylate at the 1'-hydroxyl group of sucrose with some acylation observed at the 6-hydroxyl group. The preference for the 1'-hydroxyl group is strongly affected by the hydrophobicity of the organic solvent and the chain length of the vinyl ester. Increasingly hydrophobic solvents and longer chain lengths lower the favorable formation of the 1'-acylation and improve 6-acylation. Molecular modeling of sucrose in the binding pocket of subtilisin BPN' shows that the 1'-acylation is favored in solvents that can solvate sugars (such as pyridine) as the glucose moiety is exposed to the medium, whereas 6-acylation leaves the entire sucrose molecule buried within the enzyme's binding pocket. Thus, 1'-acylation is sterically more favorable than 6-acylation. Increasingly hydrophobic solvents affect regioselectivity by changing the degree of solvation of the glucose moiety in the medium and forcing the sucrose 1'-ester completely into the binding pocket. In a related modeling, the vinyl ester chain length was shown to modulate regioselectivity by controlling the bond angles between the resulting acyl-enzymes and the sucrose thereby affecting the positioning of the sucrose in the binding pocket of subtilisin BPN'. This study shows that control over enzymic regioselectivity can be achieved by rational choices of substrate and solvent.



## KHMELNITSKY, Y. L.; MICHELS, P. C.; DORDICK, J. S.; CLARK, D. S.

Generation of Solution-Phase Libraries of Organic Molecules by Combinatorial Biocatalysis. In *American Chemical Society Conference Proceedings Series. Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery*. Chapter 14, 144-157, 1996.

Combinatorial biocatalysis is a powerful methodology for synthesizing libraries of organic compounds in solution. Advantages of biocatalysis for generating organic libraries include the natural diversity of enzymatic reactions, the compatibility of reaction conditions and high-throughput screening techniques, ease of automation, and the ability to retrace synthetic pathways leading to active products. The integration of biocatalysis with high-speed robotics thus represents a new avenue of biotechnology for the discovery of new molecules and biotransformation by the synthesis of diverse libraries from small organic precursors, as well as the iterative derivatization of taxol, a complex natural product.

## TSAI, S.; DORDICK, J. S.

Extraordinary Enantiospecificity of Lipase Catalysis in Organic Media Induced by Purification and Catalyst Engineering. *Biotechnol. and Bioeng.* 1996, 52, 296-300.

A purified lipase preparation from *Candida rugosa* was compared to its crude counterpart in anhydrous and slightly hydrated hydrophobic organic solvents. The purified lipase preparation was less active than the crude enzyme in dry *n*-heptane, whereas the presence of small concentrations of added water dramatically activated the purified enzyme but not the crude enzyme. Thus, in the presence of as little as 0.25  $\mu\text{L}/\text{mL}$  of added water in *n*-heptane, the purified enzyme is over 230-fold more active and 6-fold more enantioselective than the dry enzyme suspension in the esterification of racemic 2-(4-chlorophenoxy)propionic acid with *n*-butanol. The reactivity and selectivity of this biocatalyst, however, was affected by coalescence of the enzyme preparation suspended in the wet organic solvent. Engineering the biocatalyst environment by dissolving the purified lipase in an aqueous buffer and then adding this solution to *n*-heptane resulted in a precipitated enzyme preparation with smaller particle sizes that did not coalesce severely. In the presence of 5  $\mu\text{L}/\text{mL}$  of water added with the enzyme, this pretreatment resulted in an activation over the dry, enantiospecific catalysis ( $E > 100$ ). Hence, by simply modifying the way enzymes are hydrated, dramatic activation of the catalytic competency can be achieved.

**ALVI, K. A.; NAIR, B.; GALLO, C.; BAKER, D.**

Screening of Microbial Extracts for Tyrosine Kinase Inhibitors. *J. Antibiot.* **1997**, *50* (3), 264-266.

The main aim of this study was to identify inhibitors of p56LCK tyrosine kinase of microbial origin. We screened thousands of microbial extracts (both fungal and actinomycete) and identified three fungal extracts which showed potent inhibitor activity in the p56LCK assay. Bioassay-directed fractionation of the crude extracts provided five p56LCK tyrosine kinase potent inhibitors: emodic acid, chartreusin, 3,5-dimethylbenzoic acid, 3,5-dihydroxybenzoic acid and 3,4-dihydroxy-6-aminobenzoic acid. Emodic acid also inhibited the EGF receptor and p56fynPTK and did not inhibit proliferation of human foreskin fibroblast cells.

**ALVI, K. A.; NAIR, B.; PU, H.; URSINO, R.; GALLO, C.; MOCEK, U.**

Phomacins: Three Novel Antitumor Cytochalasan Constituents Produced by a *Phoma* sp. *J. Org. Chem.* **1997**, *62* (7), 2148-2151.

Three novel cytochalasans, phomacins A, B, and C, were isolated from a fermentation broth of the fungus *Phoma* sp. and purified by HSCCC (high speed countercurrent chromatography) followed by HPLC. The structures were determined by 1D and 2D NMR techniques. All three compounds have shown potent inhibitory activity against the HT29 colonic adenocarcinoma cell line.

**D'AMBRA, T. E.; JAVITT, N.B.; NAKANISHI, K.; WARCHOL, T.**

Synthesis of (25*R*)-Cholest-5-ene-3 $\beta$ , 26-diol and its Radiolabeled Analog. *Tetrahedron Lett.* **1997**, *38* (22), 3801-3804.

A new, convenient and stereoselective route to the synthesis of (25*R*)-cholest-5-ene-3 $\beta$ , 26-diol (**1**) and its radiolabeled analog **4** is described. The key step is a Julia condensation of sulfone **6** with aldehyde **12** to furnish compound **13**. Further reduction of the  $\alpha$ -hydroxysulfone moiety afforded 22,23-unsaturated i-steroid **14**. The double bond was reduced by hydrogen or by tritium to provide substrates for the preparation of **1** and **4**, respectively.



**KHMELNITSKY, Y. L.; BUDDE, C. L.; ARNOLD, J. M.;  
USYANTINSKY, A.; CLARK, D. S.; DORDICK, J. S.**

Synthesis of Water-Soluble Paclitaxel Derivatives by Enzymatic Acylation. *J. Am. Chem. Soc.* **1997**, *119* (47), 11554-11555.

Taxol has been acylated using the protease thermolysin in nearly anhydrous organic solvents to give a series of esters and carbonates. The enzyme, activated via lyophilization in the presence of KCl, catalyzes the facile and regioselective acylation of the 2'-hydroxyl moiety of the phenylisoserine side chain to give overall yields of taxol acylation in some cases in excess of 90%. A two-step enzymatic process employing thermolysin to generate a taxol 2'-vinyl adipate intermediate and a lipase from *Candida antarctica* has been used to generate two taxol 2'-esters, taxol 2'-adipic acid and taxol 2'-(adipoyl)glucose, with aqueous solubilities up to 1625-fold higher than native taxol.

**ALVI, K. A.; CASEY, A.; NAIR, B. G.**

Pulchellalactam: A CD45 Protein Tyrosine Phosphatase Inhibitor from the Marine Fungus *Corollospora pulchella*. *J. Antibiot.* **1998**, *51*(5), 515-517.

The novel compound, pulchellalactam was isolated from a fermentation broth of the marine fungus *Corollospora pulchella* and purified by high speed countercurrent chromatography (HSCCC) followed by HPLC. The structure was determined by 1D and 2D NMR techniques. As a part of our effort to find enzyme inhibitors from microbial sources, we identified a fungal extract which exhibited very potent activity in our CD45 assay. Bioassay directed fractionation of the marine fungal (*Corollospora pulchella*) extract yielded a novel compound as the active component of this extract. Pulchellalactam exhibited a dose-dependent inhibition of CD45 activity with an IC<sub>50</sub> of 124g/mL. This was comparable with the inhibition of CD45 activity observed in the presence of a known tyrosine phosphatase inhibitor, sodium orthovanadate which inhibited the activity with an IC<sub>50</sub> of 91.9 g/mL. The inhibition was specific for CD45 as another protein tyrosine phosphatase, PTP1B, was not inhibited by pulchellalactam.

**COTTERILL, I. C.; USYATINSKY, A. Y.; ARNOLD, J. M.;  
CLARK, D. S.; DORDICK, J. S.; MICHELS, P. C.;  
KHMELNITSKY, Y. L.**

Microwave Assisted Combinatorial Chemistry Synthesis of Substituted Pyridines. *Tetrahedron Lett.* **1998**, *39*, 1117-1120.

A new highly efficient MICROCOS technology (Microwave-assisted Combinatorial Synthesis) for generating combinatorial libraries is described. The technology is applied to the high throughput, automated, one-step, parallel synthesis of diverse substituted pyridines using the Hantzsch synthesis. The advantages of microwave-assisted chemistry for combinatorial synthesis include a broad range of available chemistries, simple reaction setup and product recovery readily amenable to automation, extremely short reaction times, and high product yields.

**DAVIDSON, M. R.; GREGG, B. T.**

Improved Synthesis of Dihydrothebainone and its 14 $\beta$ -Epimer. *Synth. Commun.* **1998**, *28* (3), 547-588.

An improved synthesis for the preparation of a diastereomerically pure dihydrothebainone (**1**) from thebaine (**3**) is reported. A 41% overall yield was realized over three steps via direct transformation of dihydrothebain- $\Phi$  (**4**) to thebainone-A (**6**) with 6N HCL.

**DORDICK, J. S.; KHMELNITSKY, Y. L.;  
SERGEEVA, M. V.**

The Evolution of Biotransformation Technologies. *Curr. Opin. Microbiol.* **1998**, *1*, 311-318.

Biotransformation is a broad and growing field of biotechnology and encompasses both enzymatic and microbial biocatalysis. Progress has been made in research on the key drivers of biotransformations, including the isolation and characterization of microbes and their enzymes from, and their utilization in, extreme environments, the manipulation, alteration, and augmentation of metabolic pathways, and the use of combinatorial biosynthesis and biocatalytic methodologies for new compound development.



**MOZHAEV, V. V.; BUDDE, C. L.; RICH, J. O.;  
USYANTINSKY, A. Y.; MICHELS, P. C.; KHMELNITSKY,  
Y. L.; CLARK, D. S.; DORDICK, J. S.**

Regioselective Enzymatic Acylation as a Tool for Producing Solution-Phase Combinatorial Libraries. *Tetrahedron* 1998, 54, 3971-3982.

A simple combinatorial strategy for sequential regioselective enzymatic acylation of multifunctional lead compounds has been developed and demonstrated using a polyhydroxylated flavonoid, bergenin, as a model. The approach is based on the ability of different enzymes to regioselectively acylate different sites on a lead molecule without affecting other similar functional groups. In sharp contrast to enzymatic acylation, conventional chemical acylation methods showed almost complete lack of regioselectivity. The enzymatic strategy was applied successfully to produce a solution phase combinatorial library of 167 distinct selectively acylated derivatives of bergenin on a robotic workstation in a 96-well plate format. General applicability of the automated combinatorial biocatalysis strategy is discussed.

**SCHULTZ, L.; GARR, C. D.; CAMERON, L. M.;  
BUKOWSKI, J.**

High Throughput Purification of Combinatorial Libraries. *Bioorg. Med. Chem. Lett.* 1998, 8, 2409-2414.

This article summarizes recent advances at MDS Panlabs which provide for the high-throughput preparative HPLC purification of our Optiverse™ screening library. Topics covered include unique methods for the preparation, purification, characterization, and plating of purified screening compounds. Also discussed are procedures for data tracking as samples proceed through the purification process.

**ALVI, K. A.**

A Strategy for Rapid Identification of Novel Therapeutic Leads from Natural Products. In *Biologically Active Natural Products: Pharmaceuticals*; Cutler, S. J.; Cutler, H. G.; Eds. CRC Press: 1999, pp 185-195.

The discovery of novel, small molecules through screening secondary microbial metabolites is still an important and fruitful activity in pharmaceutical and biotech industries. However, the isolation and structure elucidation of lead compounds

is often a tedious and time-consuming process, especially when the compounds being sought may only be present in infinitesimal quantities. When one considers, for example, that microorganism extracts have thousands of constituents, the difficulties in separating out one particular component can be appreciated.

The nature of the separation problem varies considerably, from the isolation of small quantities for dereplication study (analytical scale, milligram or less) to the isolation of larger quantities for structure elucidation and comprehensive biological testing (semi-preparative scale, 5 mg or more). For these purposes, a good selection of different techniques and approaches is essential.

The isolation and purification of a bioactive compound is a rate-limiting step in a natural products chemistry project, and significant improvement in this area is urgently needed. We have approached this problem with a view toward exploring the application of sophisticated modern scientific instruments. Our goal was to develop a process that is not only capable of effective fractionation, but also yields sufficient quantities necessary for structure elucidation and extensive biological evaluation. In addition, the process would allow us to distinguish between known and new compounds (dereplication) at an earlier phase of the project.

We have developed a rapid and systematic process for isolation and identification of biologically active components from natural products. The process reduces time and cost through application of advanced chromatographic instrumentation. It generates important activity and chemical information and also provides advanced active fraction(s) to accelerate isolation studies. As a result, lead prioritization, project management, and the cycle time of natural product lead discovery have been significantly improved.

The system relies upon preliminary fractionation of the microbial crude extract by dual-mode countercurrent chromatography coupled with photodiode array detection (PDA). The ultraviolet-visible (UV-Vis) spectra and liquid chromatography-mass spectrometry (LC-MS) of biologically active peaks are used for identification. Confirmation of compound identity is accomplished by nuclear magnetic resonance (NMR). Use of an integrated system countercurrent chromatography (CCC) separation, PDA detection, and LC-MS rapidly provided profiles and structural information extremely useful for metabolite identification.



ALVI, K. A.; PU, H.; LUCHE, M.; RICE, A.; APP, H.;  
MCMAHON, G.; DARE, H.; MARGOLIS, B.

Asterriquinones Produced by *Aspergillus candidus* Inhibit Binding of the Grb-2 Adapter to Phosphorylated EGF Receptor Tyrosine Kinase. *J. Antibiot.* 1999, 52 (3), 215-223.

Five new asterriquinone analogs (2 - 4, 6, 7), together with previously identified neoasterriquinone (1) and isoasterriquinone (5), were isolated from a fermentation broth of the fungus *Aspergillus candidus* and purified by HSCCC (high speed counter current chromatography) followed by HPLC. The structures were determined by 1D and 2D NMR and MS/MS techniques. All seven showed inhibitory activity against the binding of a recombinant protein containing the SH2 protein domain of Grb-2 to the tyrosine phosphorylated form of the EGF receptor tyrosine kinase. Some of these asterriquinones exhibited specific inhibition of Grb-2 binding compared to Grb-7 and PLC- $\gamma$ .

BAJORATH, J.

A Molecular Model of Inducible Costimulator Protein and Three-Dimensional Analysis of its Relation to the CD28 Family of T Cell-Specific Costimulatory Receptors. *J. Mol. Model.* 1999, 5, 169-176.

Inducible costimulator protein (ICOS) has recently been identified as a new member of the CD28 family of T cell costimulatory molecules. A molecular model of the extracellular immunoglobulin-like domain of ICOS was built based on the structure of CD152, another member of the CD28 family. Despite low sequence identity, ICOS shares consensus residues characteristic of immunoglobulin variable-type domains with CD152 and CD28 and also some unique features, suggesting that their three-dimensional structures are more similar to each other than to other proteins belonging to the immunoglobulin superfamily. The ICOS model was used to study sequence conservation in three dimensions and to compare the distribution of N-linked glycosylation sites in the extended CD28 family. The limited number of residues outside consensus/core positions that are conserved in ICOS and CD28 and/or CD152 are widely distributed over the extracellular domain. A few residues in CD152 and CD28 that are critical for binding of CD80/CD86 are also conserved in ICOS. However, the region in ICOS that corresponds to the CD80/CD86 binding site is masked by N-linked glycosylation. This suggests that this site is not available for binding of CD80/CD86 or other ligands. ICOS has probably diverged early from CD28 and CD152 and developed the capacity to recognize ligand(s) other than CD80/CD86, very likely uti-

lizing a different molecular region and mechanism for binding.

This has made a thorough three-dimensional analysis of both mAb and ALCAM binding, consistent with the presence of structural perturbations. However, several residues whose mutation affected both mAbs overlap and provides an explanation for the finding that these mAbs effectively block ALCAM binding. This has made a thorough three-dimensional analysis of CD6 mutagenesis and mAb binding experiments possible. Mutation of buried residues comprised both mAb and ALCAM binding, consistent with the presence of structural perturbations. However, several residues whose mutation affected both mAb and ALCAM binding, or, alternatively, only ligand binding were found to map to the surface in the same region of the domain. This suggests that the CD6 ligand binding site and epitopes of tested mAbs overlap and provides an explanation for the finding that these mAbs effectively block ALCAM binding.

## **BAJORATH, J.**

Three-Dimensional Analysis of CD6 Site-Directed Mutagenesis and Monoclonal Antibody Binding Studies Using the X-ray Structure of Mac-2 Binding Protein and a Molecular Model of the CD6 Ligand Binding Domain. *J. Mol. Model.* 1999, 5, 263-270.

The extracellular region of CD6 consists of three scavenger receptor cysteine-rich (SRCR) domains and binds activated leukocyte cell adhesion molecule (ALCAM), a member of the immunoglobulin superfamily (IgSF). Residues important for the CD6-ALCAM interaction have previously been identified by mutagenesis. A total of 22 CD6 residues were classified according to their importance for anti-CD6 monoclonal antibody (mAb) and/or ALCAM binding. The three-dimensional structure of the SRCR domain of Mac-2 binding protein has recently been determined, providing a structural prototype for the SRCR protein superfamily. An approximate molecular model of CD6 was used to delineate the ALCAM binding site.

## **BAJORATH, J.**

Specificity of the Tumor Necrosis Factor Receptor Superfamily. *J. Mol. Graphics Mod.* 1999, 17, 220-222.

No abstract available.



## BAJORATH, J.

Identification of the Ligand Binding Site in Fas (CD95) and Analysis of Fas-Ligand Interactions. *Proteins: Struct., Func., Genet.* **1999**, *35*, 475-482.

Fas(CD95), a member of the tumor necrosis factor receptor superfamily, and its ligand (FasL), a tumor necrosis factor-like protein, are intensely studied because their interaction on the cell surface is critical for the induction of programmed cell death (apoptosis) and the regulation of immune responses. The structure and specificity of the extracellular binding domains of Fas and its ligand were studied, in different laboratories, by combining molecular modeling, mutagenesis, and a variety of binding and functional experiments. Residues critical for the receptor-ligand interaction were identified and, in the absence of experimentally determined structures, binding sites and details of the Fas-ligand interactions were predicted. These studies provide an instructive example for the close combination of prediction and experiment and illustrate how insights into the structure and binding characteristics of Fas and its ligand were gradually refined. Discussed methodological aspects are representative of structure-function studies on extracellular domains of other single-path transmembrane proteins.

## BOULANGER, W. A.

Improved Resolution of ( $\pm$ )-*cis*-2'-Hydroxy-5,9-dimethyl-6,7-benzomorphan and Preparation of Racemic, (+)-, and (-)-*cis*-2-methylcyclopropyl-2'-hydroxy-5,9-dimethyl-6,7-benzomorphan (cyclazocine). *Synth. Commun.* **1999**, *29* (12), 2201-2210.

We report novel procedures for the efficient resolution of ( $\pm$ )-*cis*-2'-hydroxy-5,9-dimethyl-6,7-benzomorphan *via* the (-)-*N*-acetyl-L-glutamate and (+)- $\alpha$ -bromocamphor- $\pi$ -sulfonate salts, the direct *N*-substitution of the (+)- $\alpha$ -bromocamphor- $\pi$ -sulfonate salts, and the use of these methods for the synthesis of racemic, (+)- and (-)-cyclazocine.

**BUDDE, C. L.; KHMELNITSKY, Y. L.**

Aldolase Stability in the Presence of Organic Solvents. *Biotechnol. Lett.* **1999**, *21*, 77-80.

Rabbit muscle aldolase (RAMA) and trout muscle aldolase (TRMA) retained 100% activity in the presence of hexane, cyclohexane and toluene. Both enzymes retained greater than 80% activity in the presence of 20% (v/v) methanol. In the presence of 20% (v/v) *N, N*-dimethylformamide, RAMA and TRMA were inactive, but at least 50% activity could be restored by returning the enzymes to an aqueous environment.

**GAO, H.; BAJORATH, J.**

Comparison of Binary and 2D QSAR Analyses Using Inhibitors of Human Carbonic Anhydrase II as a Test Case. *Mol. Diver.* **1999**, *4*, 115-130.

Binary and conventional 2D QSAR have been derived for a set of carbonic anhydrase II (CA II) inhibitors. An overall predictive accuracy of 94% was obtained by binary QSAR and 84% by 2D QSAR model. For both models, preferred molecular descriptor sets were identified, which were overlapping but not identical. Both binary and 2D QSAR captured important molecular features of CA II inhibitors, notably the presence of a sulfonamido group, which is critical for binding, but also hydrophobicity. Promising results were obtained when the derived QSAR models were used to test a set of CA II inhibitors not included in the training set. In binary QSAR, previously unobserved boundary effects were detected both in the analysis of known inhibitors and when screening a large combinatorial library for putative inhibitors. The complementary use of binary and conventional 2D QSAR is thought to increase the accuracy of the lead discovery process by QSAR techniques.

**GAO, H.; WILLIAMS, C.; LABUTE, P.; BAJORATH, J.**

Binary Quantitative Structure – Activity Relationship (QSAR) Analysis of Estrogen Receptor Ligands. *J. Chem. Info. Comput. Sci.* **1999**, *39*, 164-168.

The use of high throughput screening (HTS) to identify lead compounds has greatly challenged conventional quantitative structure-activity relationship (QSAR) techniques that typically correlate structural variations in similar compounds with continuous changes in biological activity. A new QSAR-like methodology that can correlate less quantitative assay data (i.e., “active” versus “inactive”), as ini-



tially generated by HTS, has been introduced. In the present study, we have, for the first time, applied this approach to a drug discovery problem; that is, the study of estrogen receptor ligands. The binding affinities of 463 estrogen analogues were transformed into a binary data format, and a predictive binary QSAR model was derived using 410 estrogen analogues as a training set. The model was applied to predict the activity of 53 estrogen analogues not included in the training set. An overall accuracy of 94% was obtained.

## GODDEN, J. W.; STAHURA, F. L.; BAJORATH, J.

Statistical Analysis of Computational Docking of Large Compound Databases to Distinct Protein Binding Sites. *J. of Comp. Chem.* **1999**, *20* (15), 1634-1643.

The results of 16 docking simulations receptor sites and flexible ligands (~60,000 compounds in each case) are statistically analyzed and compared. Different combinations of binding sites, scoring functions, and compound collections are used in these calculations. The docking scores are not randomly distributed over the scoring range; they follow Gaussian distributions (regardless of the binding sites), scoring functions, or screened compounds. If the docking sites are small, the Gaussian distributions are positively skewed. Peaks of the Gaussian distributions are populated with compounds having similar scores but different sizes and binding modes. These findings have implications for compound selection via computational docking.

## KHMELNITSKY, Y. L.; RICH, J. O.

Biocatalysis in Nonaqueous Solvents. *Curr. Opin. in Chem. Biol.* **1999**, *3*, 47-53.

Biotransformation technologies have enjoyed a renewed interest from researchers and industry because of the progress made in the discovery and design of new, efficient biocatalysts for synthetic applications. Biocatalysis in nonaqueous media, which offers unique capabilities and thus plays a major role in biotransformation technologies, has made tremendous progress in recent years. On average, at least one paper dealing with biocatalysis in organic solvents is published every day. New, remarkable developments have taken place in several key areas of this exciting field during the past year.

**KRSTENANSKY, J. L.; KHMELNITSKY, Y.**

Biocatalytic Combinatorial Synthesis. *Bioorg. Med. Chem.* **1999**, *7*, 2157-2162.

Combinatorial biocatalysis, based on a principle of the combinatorial use of biosynthetic steps rather than the combinatorial use of reagents, offers a complementary approach to combinatorial chemistry, which, used individually or in connection with synthetic organic transformations, provides access to analogues not readily accessible by chemical synthetic means alone. The issues and strategies particular to this approach are discussed. Examples are given demonstrating these principles as well as the unique advantages of achieving chemo-, regio- and stereoselectivity under mild reaction conditions that biocatalytic methods offer.

**STAHURA, F. L.; XUE, L.; GODDEN, J. W.;  
BAJORATH, J.**

Molecular Scaffold-Based Design and Comparison of Combinatorial Libraries Focused on the ATP-Binding Site of Protein Kinases. *J. Mol. Graphics Mod.* **1999**, *17*, 1-9.

Compound libraries were designed to target specifically the ATP cofactor-binding site in protein kinases by combining knowledge- and diversity-based design elements. A key aspect of the approach is the identification of molecular building blocks or scaffolds that are compatible with the binding site and therefore capture some aspects of target specificity. Scaffolds were selected on the basis of docking calculations and analysis of known inhibitors. We have generated 75 molecular scaffolds and applied different strategies to compute diverse compounds from scaffolds or, alternatively, to screen compound databases for molecules containing these scaffolds. The resulting libraries had a similar degree of molecular diversity, with at most 12% of the compounds being identical. However, their scaffold distributions differed significantly and a small number of scaffolds dominated the majority of compounds in each library.

**XUE, L.; BAJORATH, J.**

Distribution of Molecular Scaffolds and R-Groups Isolated from Large Compound Databases. *J. Mol. Model.* **1999**, *5*, 97-102.

We describe an approach to isolate molecular scaffolds and R-groups from known chemical compounds in order to generate scaffold and R-group databases from



two large compound collections, Optiverse™ and Maybridge™. The distributions of molecular scaffolds and R-groups in the parent databases were analysed and compared. We find that a limited number of scaffolds occur only once or twice in the compound databases. Diversity analysis suggests that the compound and scaffold databases have similar molecular diversity. Implications for library design are discussed.

## SCOTT, I. L.; MARKET, R. V.; DEORAZIO R. J.; MECKLER, H.; KOGAN, T. P.

Stereospecific  $\alpha$ -D-mannosylation. *Carbohydr. Res.* **1999**, *317*, 210-216.

The stereospecific formation of  $\alpha$ -D-mannosyl glycosidic linkages has been achieved in high yield using tetra-*O*-pivaloyl- $\alpha$ -D-mannopyranosyl fluoride and boron trifluoride diethyl etherate in dichloromethane. Examples of the  $\alpha$ -D-mannosylation of primary, secondary, benzylic and phenolic hydroxyl groups are described.

## XUE, L.; GODDEN, J. W.; BAJORATH, J.

Database Searching for Compounds with Similar Biological Activity Using Short Binary Bit String Representations of Molecules. *J. Chem. Info. Comput. Sci.* **1999**, *39*, 881-886.

In an effort to identify biologically active molecules in compound databases, we have investigated similarity searching using short binary bit strings with a maximum of 54 bit positions. These “minifingerprints” (MFPs) were designed to account for the presence or absence of structural fragments and/or aromatic character, flexibility, and hydrogen-bonding capacity of molecules. MFP design was based on an analysis of distributions of molecular descriptors and structural fragments in two large compound collections. The performance of different MFPs and a reference fingerprint was tested by systematic “one-against-all” similarity searches of molecules in a database containing 364 compounds with different biological activities. For each fingerprint, the most effective similarity cutoff value was determined. An MFP accounting for only 32 structural fragments showed less than 2% false positive similarity matches and correctly assigned on average ~40% of the compounds with the same biological activity to a query molecule. Inclusion of three numerical two-dimensional (2D) molecular descriptors increased the performance by 15%. This MFP performed better than a complex 2D fingerprint. At a similarity cutoff value of 0.85, the 2D fingerprint totally eliminated

false positives but recognized less than 10% of the compounds within the same activity class.

## XUE, L.; GODDEN, J.; GAO, H.; BAJORATH, J.

Identification of a Preferred Set of Molecular Descriptors for Compound Classification Based on Principal Component Analysis. *J. Chem. Info. Comput. Sci.* **1999**, *39*, 699-704.

An algorithm based on principal component analysis was investigated to classify molecules in a database consisting of 455 compounds with activities against seven different biological targets. Diversity profiles of these compound sets were calculated and compared. To effectively classify compounds with similar biological activity, all possible combinations of 17 molecular descriptors were tested by complete factorial analysis, and preferred descriptor combinations were identified. High efficiency was achieved for a combination of a limited set of structural keys and two or three additional 2D descriptors. The performance of the approach was compared to Jarvis-Patrick clustering.

## YET, L.

Free Radicals in the Synthesis of Medium-Sized Rings. *Tetrahedron* **1999**, *55*, 9349-9403.

Free radical cyclization reactions are important tools for the construction of various types of cyclic compounds including biologically active natural products and pharmaceuticals. The advantages these reactions offer to the synthetic organic chemist include mild reaction conditions with high levels of regio- and stereo-control along with significant functional group tolerance. Recent advances in radical chemistry have led to the development of some practical methods for the formation of seven-, eight-, and nine-membered rings via radical cyclization methods. In addition to formation of the usual five- and six-membered rings using carbon radicals, ring expansion via an oxy radical is increasingly becoming a useful tool in potential syntheses of medium- and large-sized rings. This review discusses recent applications of radical chemistry to the syntheses of medium-sized (seven- to nine-membered) rings that have occurred from 1980 to end of 1998.



**ALVI, K. A.; NAIR, B. G.; RABENSTEIN, J.; DAVIS, G.;  
BAKER, D.**

CD45 Tyrosine Phosphatase Inhibitory Components from *Aspergillus niger*. *J. Antibiot.* **2000**, *53* (2), 110-113.

Two inhibitors of CD45 tyrosine phosphatase, dihydrocarolic acid (1) and penitricin D (2), were isolated from a fermentation broth of the fungus *Aspergillus niger* and purified by HSCCC (high speed countercurrent chromatography) followed by HPLC. The structures were determined by NMR. The inhibitory activities of both compounds were specific to tyrosine phosphatases.

**ALVI, K. A.; BAKER, D. D.; STIENECKER, V.; HOSKEN,  
M.; NAIR, B. G.**

Identification of Inhibitors of Inducible Nitric Oxide Synthase from Microbial Extracts. *J. Antibiot.* **2000**, *53* (5), 496-501.

A new member of the angucycline family, vineomycin C (3), together with four known metabolites saquayamycin A1 (1), A-7884 (2), rabelomycin (5) and xanthomegnin (6) were isolated from microbial extracts. The structures were determined by 1D and 2D NMR techniques and chemical degradation. Compounds 1-3 and 5 were isolated from a fermentation of *Streptomyces* sp., while 6 was isolated from a fungal fermentation extract. All five compounds have shown potent inhibitory activity in the inducible nitric oxide synthase (iNOS) assay.

**BAJORATH, J.**

Understanding the Structural Basis of T-cell Costimulation. *J. Mol. Graphics Mod.* **2000**, *18*, 176-179.

No abstract available.

## BAJORATH, J.

Molecular Organization, Structural Features, and Ligand Binding Characteristics of CD44, A Highly Variable Cell Surface Glycoprotein with Multiple Functions. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 103-111.

CD44 is a type I transmembrane protein and member of the cartilage link protein family. It is involved in cell-cell and cell-matrix interactions and signal transduction. Several CD44 ligands have been identified. CD44 is a major cell surface receptor for hyaluronan, a component of the extracellular matrix. It is implicated in diseases such as cancer and inflammation and therefore intensely studied. A characteristic feature of CD44 is the occurrence of many isoforms that are expressed in a cell-specific manner and differentially glycosylated. Although a number of CD44 isoforms have been characterized, the structural diversity of CD44 makes it often challenging to study (isoforms-specific) CD44-ligand interactions at the molecular level of detail. The structural organization and ligand binding characteristics of CD44 are focal points of this review. On the basis of recent structural and mutagenesis studies, details of the CD44-hyaluronan interaction are beginning to be understood.

## BAKER, D.; MOCEK, U.; GARR, C.

Natural Products vs. Combinatorials: A Case Study. In *Biodiversity: New Leads for the Pharmaceutical and Agrochemical Industries*; Wrigley, S. K., Hayes, M. A., Thomas, R., Chrystal, E. J. T., Nicholson, N., Eds., Royal Society of Chemistry: Cambridge, 2000, pp 66-72.

Historically, pharmaceuticals have largely been derived from natural product sources. With the development of combinatorial approaches to chemical synthesis in the last decade, drug discovery programs have adapted to large-scale screening programs of combinatorial chemical libraries of up to several hundred thousands of wells. Natural product screening programs are perceived by some in the pharmaceutical industry as antiquated, inefficient or even unproductive, despite a steady flow of natural product-derived New Chemical Entities (NCE's) into the market. Natural products hold great potential for novel drug discovery. Newer, automated and high-throughput technologies, which make high-throughput combinatorial chemical synthesis possible, can be adapted to improve the efficiency of natural product discovery programs.

Libraries of synthetic combinatorial chemicals complement libraries of natural product metabolites, but do not duplicate or replace them. Examples of side-by-side screening of synthetic combinatorial chemicals and microbial fermentation



extracts indicate that both sets of chemical diversity can provide unique leads. The quantity of leads generated by each method is irrelevant, if the quality of the lead is not considered. Therefore, scientific research should be focused not on which type of library is better, but rather how to take advantage of both resources in a cost-effective and timely manner.

To position natural products to meet the current drug discovery paradigm of high to ultra-high throughput random screening, certain well-known technologies should be used. Biological characterization as well as separation chemistries for semi-purification or full purification can be employed prior to screening to reduce the number of compounds in the screening mixture and to create links from the physical entities to databases containing chemical and biological characteristics. Lead candidate compounds from both synthetic and natural product sources can both be used to commence computational approaches to analogue compound generation for optimal lead drug development.

## **BOWEN, M.A.; ARUFFO, A.A.; BAJORATH, J.**

Cell Surface Receptors and Their Ligands: In Vitro Analysis of CD6-CD166 Interactions. *Proteins: Struct., Funct., and Genet.* **2000**, *40*, 420-428.

CD6 is a cell surface receptor belonging to the scavenger receptor cysteine-rich (SRCR) protein superfamily (SRCRSF). It specifically binds activated leukocyte cell adhesion molecule (ALCAM, CD166), a member of the immunoglobulin (Ig) superfamily (IgSF). CD166 was among the first molecules identified as a ligand for an SRCRSF receptor, and the CD6-CD166 interaction was the first interaction characterized involving SRCRSF and IgSF proteins. We focus here on what has been learned about the specifics of the CD6-CD166 interaction from in vitro analysis. The studies are thought to provide an instructive example for the analysis of interactions between single-path transmembrane cell surface proteins. Using soluble recombinant forms, the extracellular binding domains of receptor and ligand have been identified and characterized in a variety of assay systems. Both CD6 and CD166 have been subjected to intense mutagenesis and monoclonal antibody (mAb) binding studies and residues critical for their interaction have been identified. The availability of structural prototypes of both superfamilies has made it possible to map the binding site in CD166 and, more recently, in CD6, and compare these regions to epitopes of mAbs that block, or do not block, the interaction. In addition, the molecular basis of observed cross-species receptor-ligand interactions could be rationalized. These studies illustrate the value of structural templates for the interpretation of sequence and mutagenesis analyses.

---

**CHEN, Z.; PABBA, C.; MULLIGAN, S.; LATURNER, S.;  
DOLAN, J.; GARR, C.**

New Strategies for Combinatorial Library from Natural Products. In *Biodiversity: New Leads for the Pharmaceutical and Agrochemical Industries*, Wrigley, S.K., Hayes, M. A., Thomas, R., Chrystal, E. J. T., Nicholson, N., Eds., Royal Society of Chemistry: Cambridge, 2000.

In the past decade, there has been tremendous impact made by the integration of combinatorial chemistry with drug discovery research, the effect being more rapid and efficient acceleration of the drug discovery process. We apply several strategies to increase the success of library synthesis, the goal being the efficient synthesis of a collection of drug-like molecules from natural products scaffolds. Some parallel synthesis approaches we use to address both target-focused and lead expansion natural products libraries are solution phase, gel phase, resin assisted and solid phase chemistry. Features of these processes are demonstrated, using well known natural product scaffolds as examples.

**GODDEN, J. W.; XUE, L.; STAHURA, F. L.;  
BAJORATH, J.**

Searching for Molecules with Similar Biological Activity: Analysis by Fingerprint Profiling. *Pa. Symp. Biocomput.* **2000**, 8, 566-575.

We have recently developed a mini-fingerprint (MFP) representation for small molecules that performs well in database searches for compounds with similar biological activity. The MFP consists of only 54 bit positions that account for numerical ranges of three two-dimensional (2D) descriptors or the presence or absence of defined structural fragments. Here we present an analysis method, termed fingerprint profiling, to systematically compare bit patterns of compounds belonging to different biological activity classes. Some but not all bit positions were variably occupied in seven different activity classes and responsible for the detection of structure-activity differences. The analysis has made it possible to rank bit positions and encoded molecular descriptors according to their importance for our similarity search calculations. Fingerprint profiling can be applied to any keyed bit string representation and should be helpful, for example, to analyze descriptor distributions in large compound databases.



## GODDEN, J. W.; BAJORATH, J.

Shannon Entropy – A Novel Concept in Molecular Descriptor and Diversity Analysis. *J. Mol. Graphics Mod.* **2000**, *18*, 73-76.

No abstract available.

## GODDEN, J. W.; XUE, L.; BAJORATH, J.

Combinatorial Preferences Affect Molecular Similarity/Diversity Calculations Using Binary Fingerprints and Tanimoto Coefficients. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 163-166.

A combinatorial method was developed to calculate complete distributions of the Tanimoto coefficient (Tc) for binary fingerprint (FP) representations of specified length, regardless of the chemical parameters they reflect. Theoretical Tc distributions were calculated for FPs consisting of up to 67 bit positions which revealed significant statistical preferences of certain Tc values. Calculation of Tc distributions in a large compound database using different FPs mirrored the effects identified by our general analysis. On the basis of these findings, an average Tc is biased by statistically preferred values.

## GODDEN, J. W.; STAHURA, F. L.; BAJORATH, J.

Variability of Molecular Descriptors in Compound Databases Revealed by Shannon Entropy Calculations. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 796-800.

A method is introduced to calculate and compare the variability of molecular descriptors in compound databases. Descriptor variability analysis is based on histograms recording the distribution of molecular descriptors and calculation of Shannon entropy (SE), a metric originally applied in digital communication. SE values reflect the variability of descriptor settings. We have calculated a total of 92 molecular descriptors in the ACD and NCI databases and ranked them according to their variability. Significant differences in entropy are observed for a number of descriptors. However, the most variable descriptors are similar in the ACD and NCI databases. Such high-entropy descriptors are preferred tools to discriminate between compounds or account for the diversity of chemical libraries.

**HERR, J.; MECKLER, H.; SCUDERI, JR.; F.**

Observed Acidities of Charcoals, Clays, and Common Laboratory Purification Reagents in Aqueous and Organic Solutions. *Org. Process Res. Dev.* **2000**, *4*, 43-45.

In studying the purification and clarification of organic compounds with mixtures of charcoals, clays, and chromatography supports in organic solution, we noted that the observed pH of the resulting slurry was an important factor in determining which of the materials was appropriate. To examine this aspect of these purification methods, we determined the observed acidity of several commercially available reagent-grade clays, decolorizing agents, chromatography supports, and several common laboratory drying reagents in water and in several organic solvents. As expected, we found that clays, decolorizing carbons, and filter aids cannot be assumed to be pH neutral and nonreactive with organic molecules. For the purification of organic compounds with pH-sensitive functionalities, the potential acidity or basicity of a clarification reagent in the chosen solvent should be considered.

**HERR, R. J.; KUHNER, J. L.; MECKLER, H.; OPALKA, C. J.**

A Convenient Method for the Preparation of Primary and Symmetrical *N,N'*-Disubstituted Thioureas. *Synthesis* **2000**, *11*, 1569-1574.

A convenient process for the preparation of both primary thioureas **2** and symmetrical *N,N'*-disubstituted thioureas **6** based on the condensation of amine hydrohalides **5** with potassium thiocyanate has been developed. This approach tolerates sterically bulky primary amine substrates (both chiral and achiral), and the products can usually be isolated by a simple filtration of the reaction mixture. This method is an especially attractive alternative for the synthesis of thioureas when the corresponding isothiocyanates are unavailable, or difficult to prepare. It is also worth noting that a wide variety of amine hydrohalides, which are used in this procedure, are commercially available.



## ITOV, Z.; MECKLER, H.

A Practical Procedure for the Resolution of (+)- and (-)- Tramadol. *Org. Process Res. Dev.* **2000**, *4* (4), 291-294.

A practical procedure for the efficient resolution of *cis*-tramadol [*cis*-2-(dimethylaminomethyl)-1-(3-methoxyphenyl)cyclohexanol] has been developed. This process was based on the observation that *cis*-tramadol free base selectivity formed mandelic acid salts at different rates, affording a readily scalable kinetic resolution of each enantiomer. The key to the process was the observation that the resolving salt needed to be broken and re-formed to ultimately improve the optical purity. The mandelate salt of each *cis*-enantiomer was found to be >99% optically pure after three cycles through the salt formation process. A sample of each mandelate salt enantiomer was successfully converted to the known, optically active hydrochloride salt.

## KING, C. R.; MECKLER, H.; HERR, R. J.; TROVA, M. P.; GLICK, S. D.; MAISONNEUVE, I. M.

Synthesis of Enantiomerically Pure (+)- and (-)-18-Methoxycoronaridine Hydrochloride and Their Preliminary Assessment as Anti-Addictive Agents. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 473-476.

Chemical resolution of racemic 18-methoxycoronaridine (**18-MC**) was achieved by the formulation of its diastereomeric sulfonamides with either (*R*)-(-)- or (*S*)-(+)-camphorsulfonyl chloride. Preliminary assessment of (+)-, (-)-, and (±)-**18-MC** HCl showed similar effects on morphine self-administration in a rat model, and similar affinities at the opioid receptors.

## KRSTENANSKY, J. L.; COTTERILL, I.

Recent Advances in Microwave-Assisted Organic Syntheses. *Curr. Opin. Drug Dis. Dev.* **2000**, *3*, 454-461.

The application of microwaves to organic synthesis is opening up new opportunities for the synthetic chemist by providing new routes not practical by traditional means, such as improved reaction yields, solvent-free reaction conditions or decreasing the time needed for transformations. This is an area of synthesis that is rapidly developing. It is becoming clear that microwave approaches can be developed for most reactions requiring heating. The most interesting developments are those demonstrating the enabling of microwave reactions where traditional

methods have failed or give only poor yields. This review intends to point out some of the more interesting papers in this area published over the last year or so, where the microwave method was demonstrated as clearly superior to previously available techniques or a new useful approach was presented.

## MECKLER, H.; SHULTIS, K.

A Formula for Outsourcing Success. *Contract Pharma* **2000**, 42-50.

A process of teamwork and communication between the client and contractor is important to good project management in all stages. The efficiency of communication is critical to the client's leverage of internal resources and utilization of external resources. The four factors of success are characteristics of a partnership between the client and the contractor.

- Take ownership of a client's project
- Clarify expectations at every stage of a project
- Communicate progress efficiently
- Communicate trouble without delay

Following these factors will lead to the formula for success for pharmaceutical and biotechnology companies looking for a competitive advantage in their chemistry research and development efforts.

## SERGEEVA, M. V.; MOZHAEV, V. V.; RICH, J. O.; KHMELNITSKY, Y. L.

Lipase-Catalyzed Transamidation of Non-Activated Amides in Organic Solvent. *Biotechnol. Lett.* **2000**, *22*, 1419-1422.

A novel biocatalytic reaction of transamidation of non-activated amides with amines is reported. Among 45 different lipolytic and proteolytic enzymes tested, only the lipase from *Candida antarctica* was able to catalyze this reaction. The reaction proceeded with up to ca. 80% conversion in anhydrous methyl *tert*-butyl ether and worked with both *N*-substituted and unsubstituted amides. The biocatalytic transamidation is an equilibrium process and, therefore, higher conversions to the desired amide were achieved by using increased concentrations of the amine nucleophile.



**STAHURA, F. L.; GODDEN, J. W.; XUE, L.;  
BAJORATH, J.**

Distinguishing between Natural Products and Synthetic Molecules by Descriptor: Shannon Entropy Analysis and Binary QSAR Calculations. *Chem. Inf. Comput. Sci.* **2000**, *40*, 1245-1252.

Molecular descriptors were identified by Shannon entropy analysis that correctly distinguished, in binary QSAR calculations, between naturally occurring molecules and synthetic compounds. The Shannon entropy concept was first used in digital communication theory and has only very recently been applied to descriptor analysis. Binary QSAR methodology was originally developed to correlate structural features and properties of compounds with a binary formulation of biological activity (i.e., active or inactive) and has here been adapted to correlate molecular features with chemical source (i.e., natural or synthetic). We have identified a number of molecular descriptors with significantly different Shannon entropy and/or "entropic separation" in natural and synthetic compound databases. Different combinations of such descriptors and variably distributed structural keys were applied to learning sets consisting of natural and synthetic molecules and used to derive predictive binary QSAR models. These models were then applied to predict the source of compounds in different test sets consisting of randomly collected natural and synthetic molecules, or, alternatively, sets of natural and synthetic molecules with specific biological activities. On average, greater than 80% prediction accuracy was achieved with our best models. For the test case consisting of molecules with specific activities, greater than 90% accuracy was achieved. From our analysis, some chemical features were identified that systematically differ in many naturally occurring versus synthetic molecules.

**USYANTINSKY, A. Y.; KHMELNITSKY, Y. L.**

Microwave-Assisted Synthesis of Substituted Imidazoles on a Solid Support Under Solvent-Free Conditions. *Tetrahedron Lett.* **2000**, *41*, 5031-5034.

The solvent-free microwave-assisted synthesis of 2,4,5-substituted and 1,2,4,5-substituted imidazoles is reported. Imidazoles are obtained as a result of the condensation of a 1,2-dicarbonyl compound with an aldehyde and an amine using acidic alumina impregnated with ammonium acetate as the solid support.

**XUE, L.; BAJORATH, J.**

Molecular Descriptors in Chemoinformatics, Computational Combinatorial Chemistry, and Virtual Screening. *Combin. Chem. High Throughput Screen.* **2000**, *3*, 363-372.

Many contemporary applications in computer-aided drug discovery and chemoinformatics depend on representations of molecules by descriptors that capture their structural characteristics and properties. Such applications include, among others, diversity analysis, library design, and virtual screening. Hundreds of molecular descriptors have been reported in the literature, ranging from simple bulk properties to elaborate three-dimensional formulations and complex molecular fingerprints, which sometimes consist of thousands of bit positions. Knowledge-based selection of descriptors that are suitable for specific applications is an important task in chemoinformatics research. If descriptors are to be selected on rational grounds, rather than guesses or chemical intuition, detailed evaluation of their performance is required. A number of studies have been reported that investigate the performance of molecular descriptors in specific applications and/or introduce novel types of descriptors in combinatorial chemistry and compound screening.

**XUE, L.; GODDEN, J. W.; BAJORATH, J.**

Evaluation of Descriptors and Mini-Fingerprints for the Identification of Molecules with Similar Activity. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1227-1234.

Combinations of 65 preferred 1D/2D molecular descriptors and 143 single structural keys were evaluated for their performance in compound classification focused on biological activity. The analysis was based on principal component analysis of descriptor combinations and facilitated by use of a generic algorithm and different scoring functions. In these calculations, several descriptor combinations with greater than 95% prediction accuracy were identified. A set of 40 preferred structural keys was incorporated into a small binary fingerprint designed to search databases for compounds with biological activity similar to query molecules. The performance of mini-fingerprints was tested by systematic similarity search calculations in a database consisting of compounds belonging to seven biological activity classes, which had not been used to select effective descriptors. In these blind test calculations, mini-fingerprints correctly identified approximately 54% of compounds sharing similar biological activity and with 1% false positives. Thus, although the design of mini-fingerprints is conceptually simple, they perform well in activity-oriented similarity searching.



## XUE, L.; BAJORATH, J.

Molecular Descriptors for Effective Classification of Biologically Active Compounds Based on Principal Component Analysis Identified by a Genetic Algorithm. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 801-809.

We have evaluated combinations of 111 descriptors that were calculated from two-dimensional representations of molecules to classify 455 compounds belonging to seven biological activity classes using a method based on principal component analysis. The analysis was facilitated by application of a genetic algorithm. Using scoring functions that related the number of compounds in pure classes (i.e., compounds with the same biological activity), singletons, and mixed classes, effective descriptor sets were identified. A combination of only four molecular descriptors accounting for aromatic character, hydrogen bond acceptors, estimated polar van der Waals surface area, and a single structural key gave overall best results. As this performance level, ~91% of the compounds occurred in pure classes and mixed classes were absent. The results indicate that combinations of only a few critical descriptors are preferred to partition compounds according to their biological activity, at least in the test cases studied here.

## YET, L.

Metal-Mediated Synthesis of Medium-Sized Rings. *Chemical Reviews* **2000**, *100* (8), 2963-3007.

The importance of medium- and large-sized rings in organic chemistry is exemplified by their being the structural core of a large number of biologically important natural products and their serving as target molecules for numerous synthetic studies. Several excellent reviews have been written on medium ring syntheses. One notable area is the vast amount of research over the past decade devoted to taxol and polycyclic ether antibiotics. Although synthetic approaches to five- and six-membered ring systems are common via cyclization and cycloaddition reactions, seven- and eight-membered ring formations are not as abundant. Cyclization strategies to medium-sized rings are often inhibited due to entropic factors and transannular interactions. In general, the number of methods for preparing medium-sized carbocycles by cyclization of cycloaddition reactions from acyclic substrates is relatively small. The past decade has witnessed a tremendous growth in the area of metal-mediated synthetic methodology, and this review will discuss recent applications of the use of metals to the syntheses of medium-sized (from seven- to nine-membered) rings that have occurred from 1990 to mid-1999. Any omissions on this wide topic are unintentional and should be brought to the attention of the author.

**ALBERT, D. K.; JONES, L. D.**

A New Paradigm in Chemical Contract Outsourcing. *Helix* 2000, 2, 7-8.

As the pharmaceutical and healthcare industry continues its rapid change, chemical outsourcing providers continue to grow in support of the industry's need for discovery chemistry services. The chemistry outsourcing market is rapidly increasing and is currently estimated to be worth \$500 million. The market is expected to outpace other markets, fueled by new technologies in genomics research, combinatorial chemistry and high-throughput biological screening.

Many value-added technologies are also applicable to areas outside traditional pharmaceutical requirements. The disciplines of combinatorial biocatalysis, combinatorial chemistry and computational chemistry combined with a strong organic chemistry expertise provide the basis for optimizing new catalysts, agrochemicals, and other commercially interesting intermediates.

**BAJORATH, J.**

Selected Concepts and Investigations in Compound Classification, Molecular Descriptor Analysis, and Virtual Screening. *J. Chem. Inf. and Comput. Sci.* 2001, 41, 233-245.

No abstract available.

**CHAPLIN, J. A.; BUDDE, C. L.; KHMELNITSKY, Y. L.**

Catalysis by Amine Oxidases in Nonaqueous Media. *J. Mol. Catal. B: Enzymatic* 2001, 13, 69-75.

The synthetic potential of amine oxidases was examined in different reaction systems, ranging from aqueous solutions to organic solvents with low water content. Substantial conversion was achieved in biphasic systems, which eliminated the product inhibition observed in the aqueous system. The conversion was particularly high in the more hydrophobic solvents. The use of low water systems was studied using amine oxidase immobilized on celite and pre-equilibrated in a salt hydrate environment to reach a constant water activity. Addition of water in the solvent was shown to be unnecessary, with significant conversion being attained through the water supplied by pre-equilibration of the immobilized enzyme at  $a_w = 0.55$ . The use of organic solvent-containing reaction systems thus presents a convenient method for oxidizing poorly water-soluble amines using amine oxidases.



**HERR, R. J.; ZHICHKIN, P.; HERNÁNDEZ ABAD, P. E.; MECKLER, H.; SCHOW, S. R.**

An Efficient Synthesis of 2-Hydroxyethyl *N,N,N',N'* – Tetrakis(2-chloroethyl) phosphorodiamidate. *Org. Process Res. Dev.* **2001**, *5*, 443-445.

A process for the multikilogram preparation of 2-hydroxyethyl *N,N,N',N'* – tetrakis(2-chloroethyl)phosphorodiamidate has been achieved in substantially pure form by a short synthetic sequence starting from phosphorus oxychloride and 2 equiv of bis(2-chloroethyl)amine. This process involves a two-step preparation of the intermediate mustard chloride in one pot, followed by the base-catalyzed reaction with excess ethylene glycol. This method has been carried out to provide 2.9 kg of this key drug substance intermediate in 52% overall yield.

**RICH, J. O.; KHMELNITSKY, Y. L.**

Phospholipase D-catalyzed Transphosphatidylation in Anhydrous Organic Solvents. *Biotechnol. Bioeng.* **2001**, *72*, 374-377.

A new reaction system suitable for phospholipase D (PLD)-catalyzed transphosphatidylation of alcohols with phosphatidylcholine under anhydrous conditions is reported. The key innovation of the reaction system is a cation-exchange resin serving as a scavenger for choline that forms as a byproduct in the transphosphatidylation reaction. Due to the absence of water in this system, the reaction path dramatically shifts in favor of the target transphosphatidylated product, whereas the undesirable side hydrolysis of phosphatidylcholine is completely suppressed, in contrast to commonly used biphasic water-organic systems. In addition, a salt activation technique is successfully applied to increase the catalytic activity of PLD in this anhydrous system. The new reaction system is successfully used for transphosphatidylation of a wide range of primary, secondary, and aromatic alcohols catalyzed by PLD from *Streptomyces sp.*

**VOGT, P. F.; MOLINO, B. F.; ROBICHAUD, A. J.**

A Regiospecific Synthesis of 3,3,6-Trimethylindan-1-one. *Synth. Commun.* **2001**, *31* (5), 679-684.

A novel, regiospecific synthesis of 3,3,6-trimethylindan-1-one (5) was achieved. The route to 5 was 6 steps and proceeded in 27% overall yield.

**XUE, L.; GODDEN, J. W.; STAHURA, F. L.;  
BAJORATH, J.**

A Dual Finger-Print Based Metric for the Design of Focused Compound Libraries and Analogs. *J. Mol. Model.* **2001**, *7*, 125-131.

A computational metric is introduced for the design of combinatorial libraries focused on small molecules with specific activity (e.g., enzyme inhibitors). The method follows a product-based design strategy and uses combinations of two binary molecular fingerprints to create chemical diversity around selected compounds and/or core structures. In the first step, compounds are sampled that are distinct from template molecules but likely to share similar biological activity. In the second step, designed compounds are accepted if they are not too similar to each other, as assessed by calculation of fingerprint overlap. Thus, it is possible to balance molecular "similarity" and "diversity" and control the degree of chemical diversity created in the vicinity of selected template molecules. In essence, the method aims to generate diverse arrays of compounds with a high probability of having activity similar to starting molecule(s) and is therefore well suited for the design of target-focused libraries to series of analogs. As an example, the method is applied to focus libraries on known protein kinase inhibitors.

**XUE, L.; STAHURA, F. L.; GODDEN, J. W.;  
BAJORATH, J.**

Mini-Fingerprints Detect Similar Activity of Receptor Ligands Previously Recognized Only by Three-Dimensional Pharmacophore-Based Methods. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 394-401.

Mini-Fingerprints (MFPs) are short binary bit string representations of molecular structure and properties, composed of few selected two-dimensional (2D) descriptors and a number of structural keys. MFPs were specifically designed to recognize compounds with similar activity. Here we report that MFPs are capable of detecting similar activities of some druglike molecules, including endothelin A antagonists and  $\alpha_1$ -adrenergic receptor ligands, the recognition of which was previously thought to depend on the use of multiple point three-dimensional (3D) pharmacophore methods. Thus, in these cases, MFPs and pharmacophore fingerprints produce similar results, although they define, in terms of their complexity, opposite ends of the spectrum of methods currently used to study molecular similarity or diversity. For each of the studied compounds classes, comparison of MFP bit settings identified a consensus or signature pattern. Scaling factors can be applied to these bits in order to increase the probability of finding compounds with similar activity by virtual screening.



**XUE, L.; STAHURA, F. L.; GODDEN, J. W.;  
BAJORATH, J.**

Fingerprint Scaling Increase the Probability of Identifying Molecules with Similar Activity in Virtual Screening Calculations. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 746-753.

Results of systematic virtual screening calculations using a structural key-type fingerprint are reported for compounds belonging to 14 activity classes added to randomly selected synthetic molecules. For each class, a fingerprint profile was calculated to monitor the relative occupancy of fingerprint bit positions. Consensus bit patterns were determined consisting of all bits that were always set on in compounds belonging to a specific activity class. In virtual screening calculations, scale factors were applied to each consensus bit position in fingerprints of query molecules. This technique, called "fingerprint scaling," effectively increases the weight of consensus bit positions in fingerprint comparisons. Although overall prediction accuracy was satisfactory using unscaled calculations, scaling significantly increased the number of correct predictions but only slightly increased the rate of false positives. These observations suggest that fingerprint scaling is an attractive approach to increase the probability of identifying molecules with similar activity by virtual screening. It requires the availability of a series of related compounds and can be easily applied to any keyed fingerprint representation that associates bit positions with specific molecular features.

**YET, L.**

Five-Membered Ring Systems: With More Than One N Atom. *Progress in Heterocyclic Chemistry*; Gribble, G. W.; Gilchrist, T. L.; Eds.; Pergamon: Oxford, 2001; In Press.

Major advancements in the chemistry of pyrazoles, imidazoles, triazoles, tetrazoles, and related fused heterocyclic derivatives appeared in 1999. Solid-phase combinatorial chemistry of benzimidazoles and triazoles has been particularly active. Synthetic routes to all areas continue to be pursued vigorously with improvements and applications. In medicinal chemistry, synthesis and structure-activity relationship (SAR) studies utilizing these core structures have been exploited heavily. The physical organic chemistry of pyrazoles and imidazoles continue.

## YET, L.

Recent Developments in Catalytic Asymmetric Strecker-Type Reactions. *Angew. Chem. Int. Ed.* **2001**, *40* (5), 875-877.

After 150 years, a truly efficient three-component asymmetric version of the Strecker reaction has finally been accomplished. There has also been significant progress in recent years in two-component Strecker-type systems: Enantiopure amino acids can now be obtained from enantioselective catalysis of the Strecker reaction with different cyanide sources and aldimines.



Albany Molecular Research, Inc.





Bothell Research Center  
Bothell, Washington



Syracuse Research Center  
Syracuse, New York



Albany Research Center  
Albany, New York



Bruno Biotechnology Center  
Rensselaer, New York



Biocatalysis Division  
Coralville, Iowa



**Albany Molecular Research, Inc.<sup>®</sup>**

21 Corporate Circle, P.O. Box 15098 Albany, NY 12212-5098 USA

Tel. 518-464-0279 Fax 518-464-0289 [www.albmolecular.com](http://www.albmolecular.com)