Accelerated Discovery Chemistry Using Fluoroorganic Ligands

228th ACS National Meeting

Philadelphia, PA

Andrei A. Gakh, PhD
Oak Ridge National Laboratory

August 22, 2004

Discovery Chemistry Project

an Overview

Objective

- **Synthesis and utilization of novel organic compounds by U.S. National Laboratories, chemical companies, and other government and non-government organizations.**

Approach

- **An international project offering the combination of expertise and resources of the participants in a team effort under the guidance and financial support of the Initiatives for Proliferation Prevention.**

Project-at-Glance

Resources

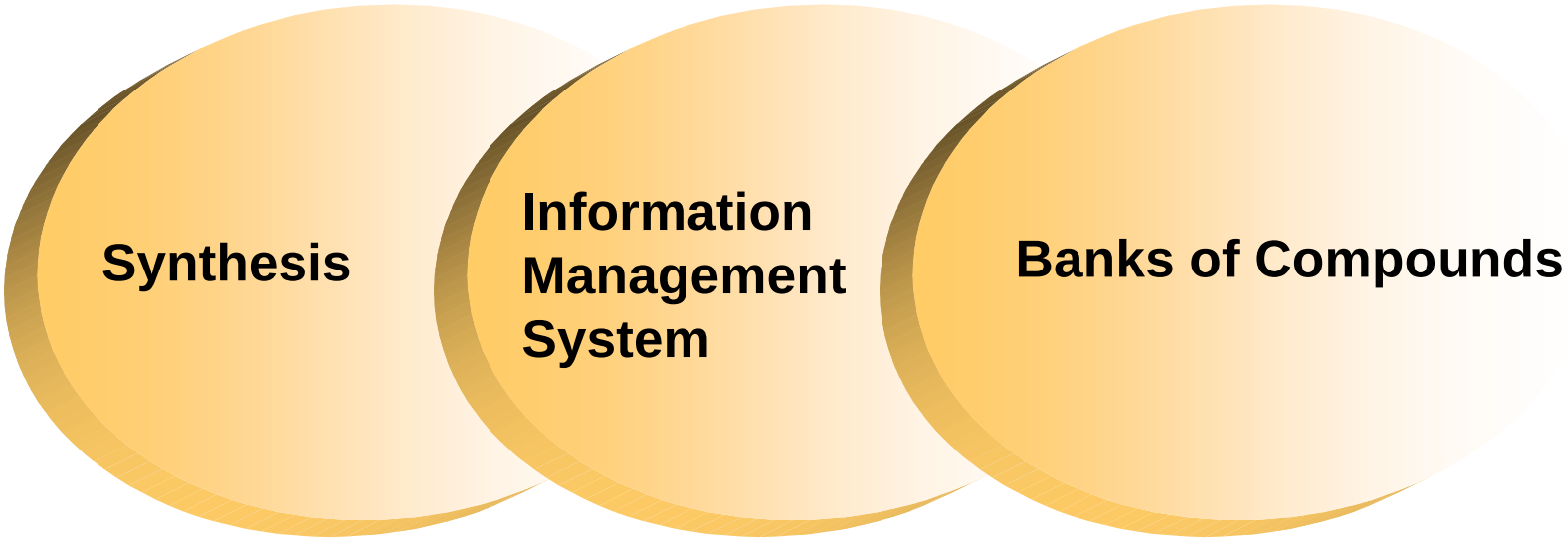
Resources:

- **Industry: 52%**
- **National Laboratories: 13%**
- **Universities: 4%**
- **International Collaborators: 31%**

Major USG Funding:

- **US DOE, National Nuclear Security Administration**
  - **Initiatives for Proliferation Prevention**

Major Research Directions



More than 60% of the resources are allocated to biological targets (biological discovery)

Biological Targets

The World of Big Numbers and Low Yields

Recent example:

- **39,647 structures proposed**
  - **11,248 structures accepted**
    - **6,217 compounds prepared**
    - **5,981 passed quality control**
    - **5,624 screened**
      - **127 active compounds identified**
        - **32 active leads selected**

Yield: 0.081% (active leads)

Human cost: >70 FTE (or > 2 FTE to develop an active lead)

Complexity of Biological Targets

GABA Receptors

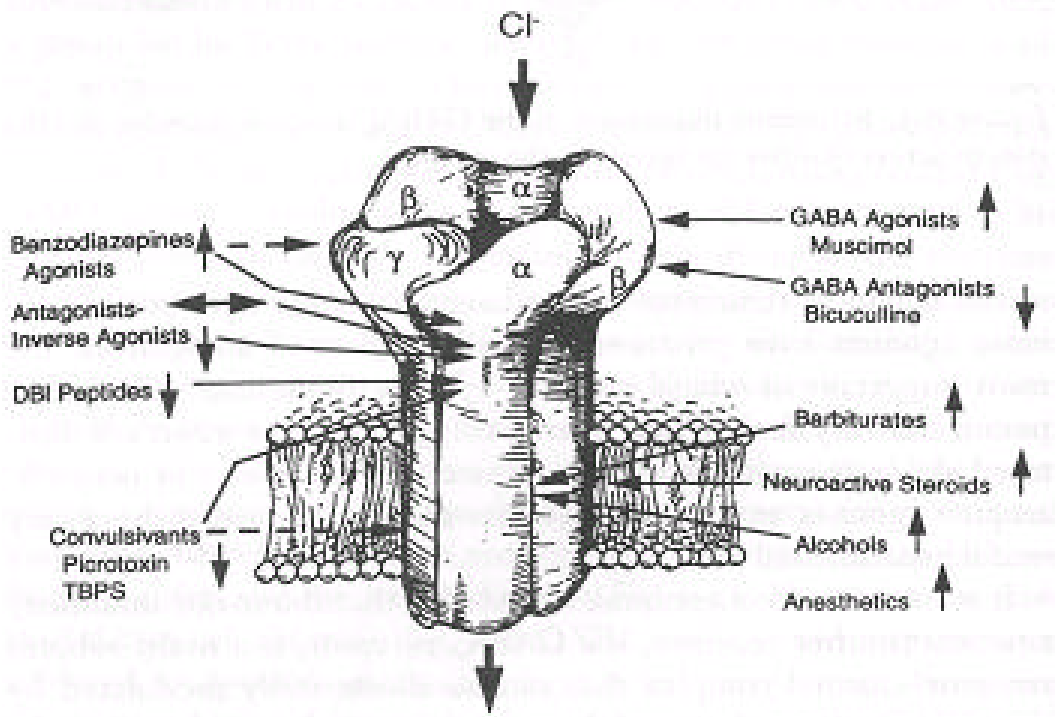
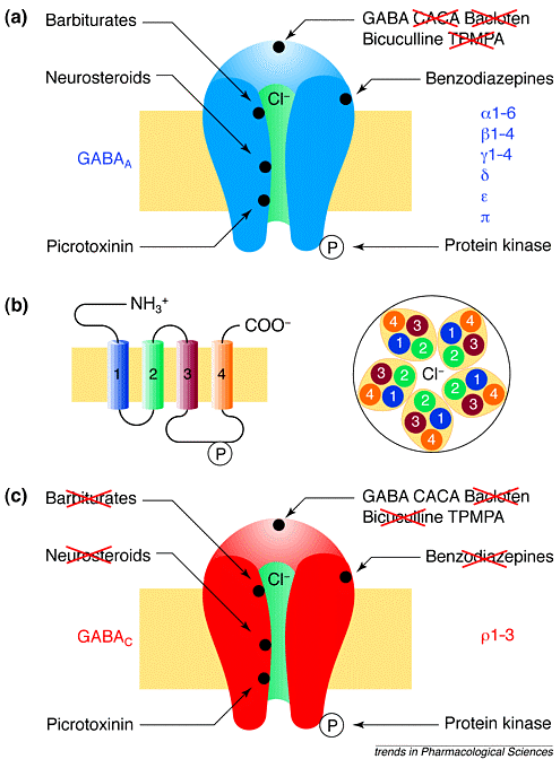
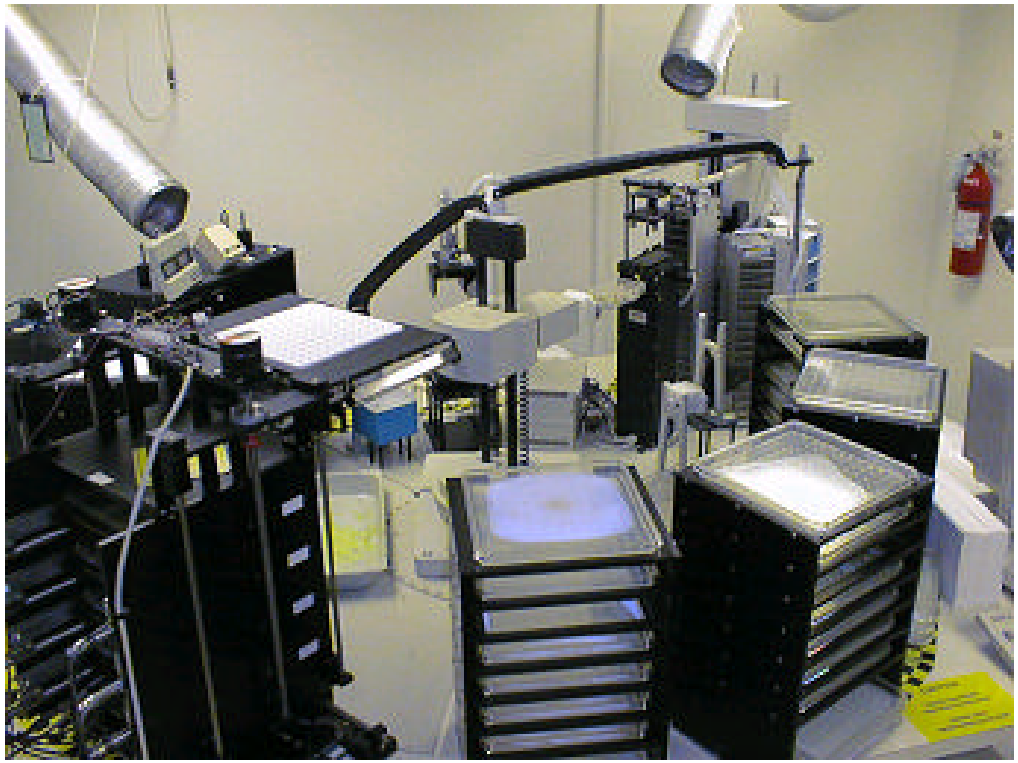


FIGURE 6-4. Schematic illustration of the $GABA_A$ receptor structure con-



The $GABA_A$ receptor is a pentamer consisting of various combinations of alpha, beta, and gamma subunits. Together these subunits form the chloride (Cl^-) ionophore (channel). The receptor also contains other binding sites (e.g., for benzodiazepines and barbiturates). These binding sites influence the frequency at which the channel opens and the duration of time the channel remains open in response to $GABA_A$, respectively. $GABA_B$ and $GABA_C$ receptors have different binding sites compared to $GABA_A$.

High Throughput Screening



The robot photo (courtesy of American Cyanamid Co.) shows the well plates with recombinant yeast containing a gene expressing a receptor and a marker gene. If a compound binds to the receptor the marker gene results in a color change so the actives can be easily picked out. The robot is capable of testing several hundred compounds per week.

Accelerated Discovery

- **Objective: increasing odds of finding an active ligand**
- **Approach: utilization of fluoroorganic compounds***

Screened Compounds	Active Compounds
➢**Total Number:** • **6,217** ➢**With Fluorine:** • **421** • **6.7% (S_T)**	➢**Total Number:** • **127** ➢**With Fluorine:** • **15** • **11.8%(S_F)**

"Acceleration Factor" a = S_F/S_T = 1.76

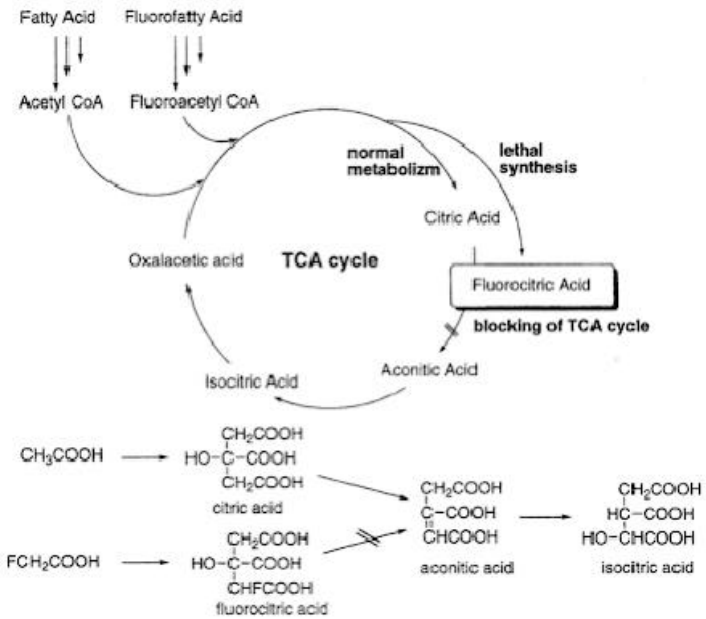
*Can be accomplished without a significant loss of diversity

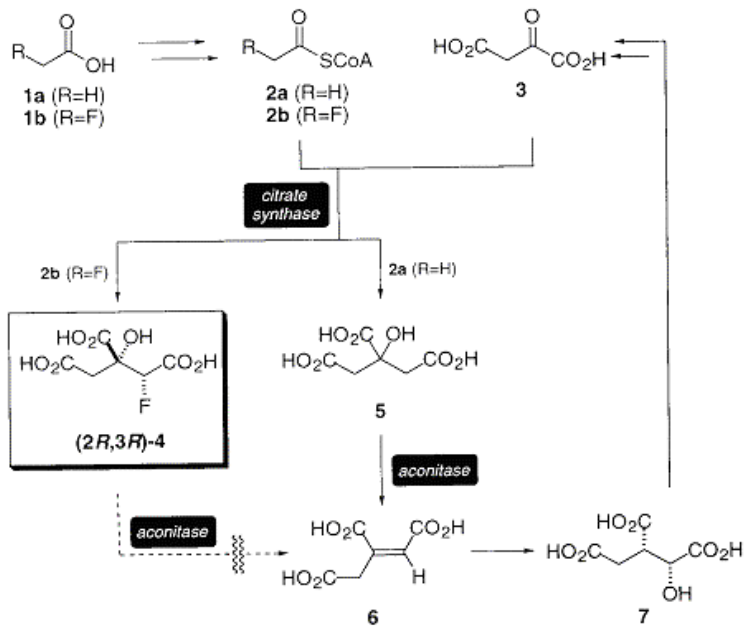
Fluorine in Biological Systems

- **General considerations**
  - Naturally-occurring fluoroorganic molecules are very rare (less than 0.1% of all halogenated compounds); no common biological defense mechanisms exist to neutralize these xenobiotics
  - The sizes of fluorine atoms and other biogenic second-row elements (O/N/C) are similar; enzymatic recognition is not always effective
  - C-F bond strength is comparable to C-C bond strength; fluorine atoms are usually retained in molecules during biochemical transformations
  - Polar character of many fluorinated fragments may facilitate interactions with active sites of some protein-based receptors
- **Fluorine effects**
  - Unique; other halogens did not improve performance in our tests
  - Appear to be statistically non-specific in respect to a molecular structure or a mode of bioaction; toxicity usually increases in tandem with the desired biological activity

Fluorine Biochemistry

Metabolic Transformations of Fluoroacetic Acid

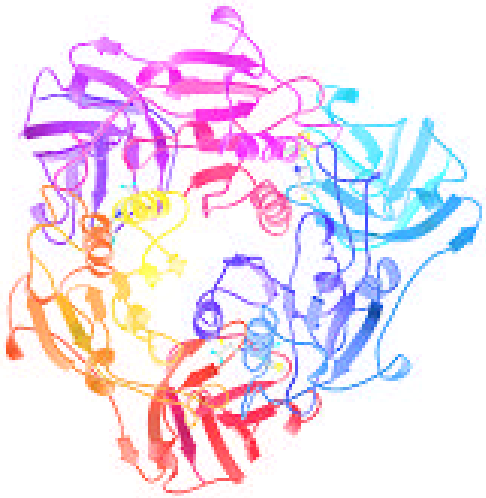


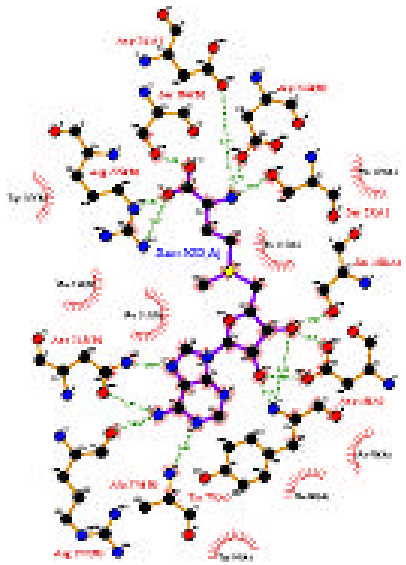


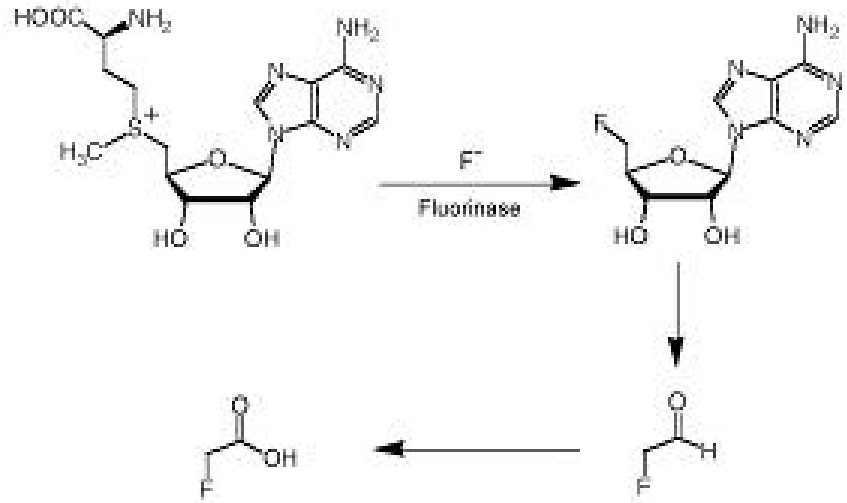
The fluoroacetic acid is “mis-recognized” as acetic acid and converted into fluorocitric acid via combination of fluoroacetyl-CoA and oxaloacetic acid with citrate synthase. Normally the citric acid undergoes the aconitase catalyzed dehydration/ hydration to yield isocitric acid. Fluorocitric acid does not isomerize leading to an inhibition of the citric acid cycle.

Fluorine Biosynthesis

Adenosyl-Fluoride Synthase (Fluorinase)







Fluorinase (ca. 32000 Dalton)

1) O'Hagan, D., Schaffrath, C., Cobb, S.L., Hamilton, J.T. and Murphy, C.D. Biosynthesis of an organofluorine molecule. Nature 416 (2002) 279.
2) Dong, C., Huang, F., Deng, H., Schaffrath, C., Specner, J.B., O'Hagan, D. and Naismith, J.H. Crystal structure and mechanism of a bacterial fluorinating enzyme. Nature 427 (2004) 561-565.

Practical Synthesis of Fluoroorganic Compounds

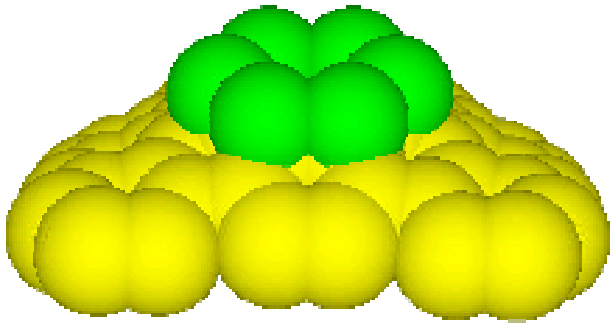
- **Direct electrophilic fluorination**
  - "CF_x" and related compounds, such as fluorofullerenes ($C_{60}F_{48}$)
  - Xenon fluorides (XeF_2, FXeOR)
  - Fluoroxysulphates ($FOSO_2O^-$)
- **Nucleophilic fluorination via quaternary salts**
  - Synthesis of quaternary C-heteroaromatic salts
  - Nucleophilic displacement (F^- or $(HF)_x \bullet F^-$)
- **Fluorodenitration**
  - Synthesis of activated nitroaromatic intermediates
  - Nucleophilic substitution with fluorides followed by derivatization of the remaining nitro groups
- **From fluorinated blocks**
  - Preparation of key fluorinated substrates
  - Incorporation of fluorinated fragments into target molecules

Fluorinating Agents: CF_x and $C_{60}F_{48}$

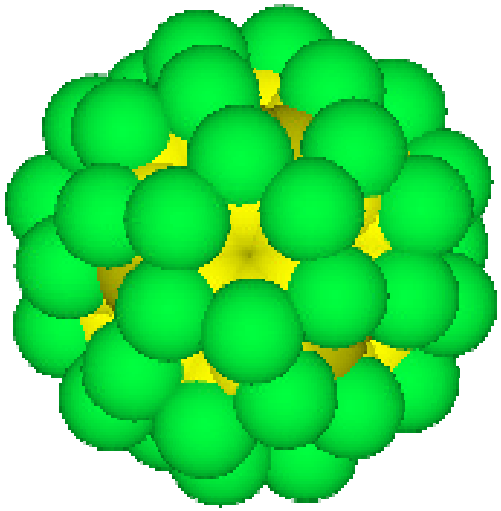
- **Advantages**
  - Solid, thermally stable, with unlimited shelf life and relatively low corrosive and toxicity profile
  - Excess of a reagent and related by-products can be separated from reaction mixtures by filtration
- **Disadvantages:**
  - Expensive (several $$ per gram).
  - Low level of fluorination; utilization of available fluorine is less than 10%
  - Typically require the use of strong catalysts (such as BF_3)
  - Strong oxidants; ineffective for fluorination of easily oxidizable compounds
  - Can be used only in a limited set of organic solvents (such as chlorinated hydrocarbons or acetonitrile)

Steric Effects and Mode of Action

CF_x and $C_{60}F_{48}$



CF_x



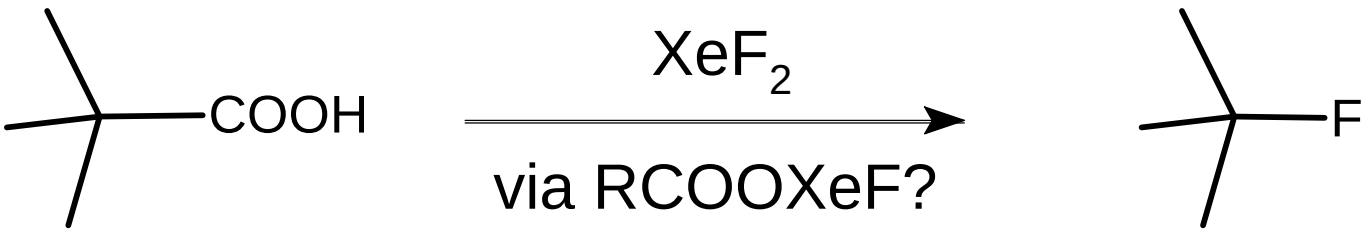
$C_{60}F_{48}$

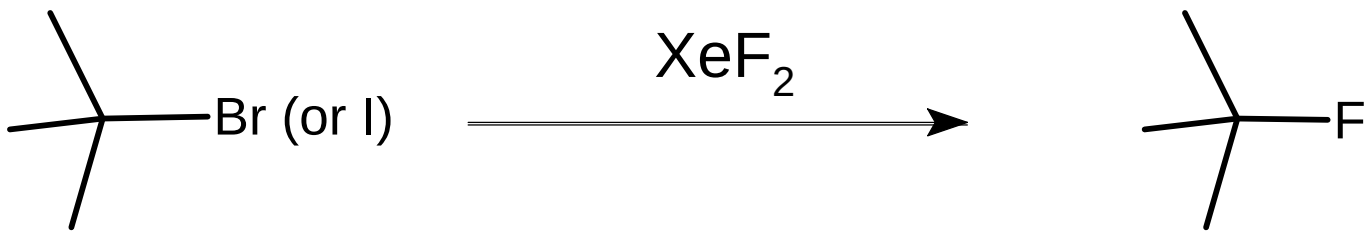
Xenon Fluorides (XeF_2, FXeOR)

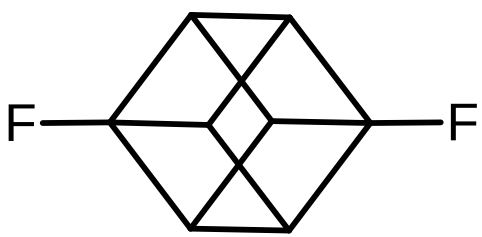
- **Advantages**
  - "Clean" reagents – the only by-product is Xe
  - Available from one of our International partners on a kilogram scale
  - Unique type of action (such as formation of Xe-organic intermediates)
  - The only convenient reagents for the fluorodecarboxylation and oxidative halogen exchange reactions
- **Disadvantages:**
  - Expensive ($3 per g)
  - Volatile (XeF_2), moisture-sensitive and corrosive
  - Require specially coated equipment to avoid autocatalytic degradation (glass equipment cannot be used)
  - React with many common solvents

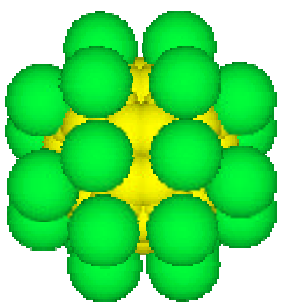
Electrophilic Fluorination

Xenon Fluorides (XeF_2, FXeOR)









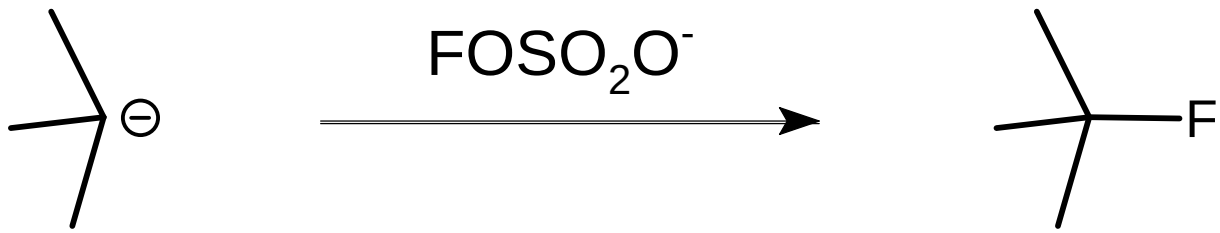
$C_{60}F_{24}$

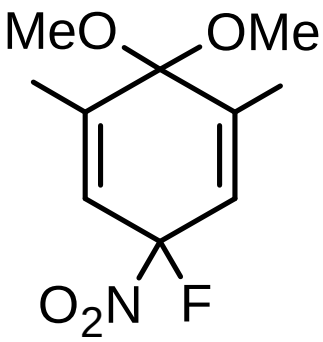
Fluoroxysulphates, $FOSO_2O^- M^+$

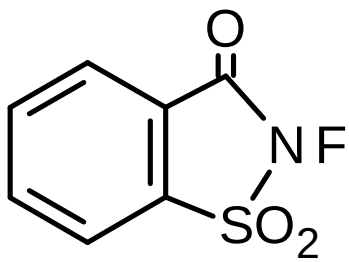
- **Advantages**
  - Can be produced *in situ* from inexpensive chemicals (elemental fluorine and sulfates) on a multi-kilogram scale using modified industrial equipment
  - Very potent electrophilic fluorinating agent (many reactions can be completed in seconds in 95-97% yield)
- **Disadvantages:**
  - Thermally unstable (require cryogenic equipment for solution storage)
  - Can be used (almost exclusively) only in water or acetonitirle
  - “Ambident” electropilic fluorinating agent: oxygen in O-F fragment can serve as an electrophilic center thus negatively affecting regioselectivity of fluorination process

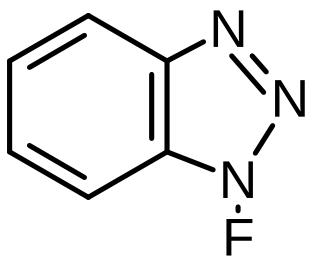
Electrophilic Fluorination

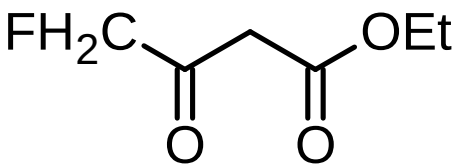
$FOSO_2O^- M^+$





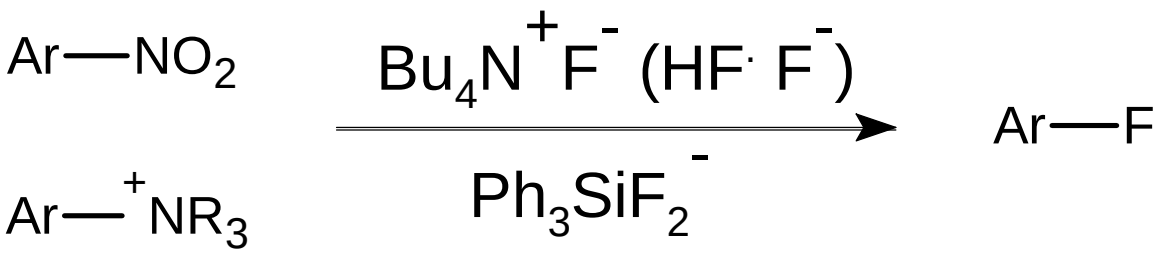


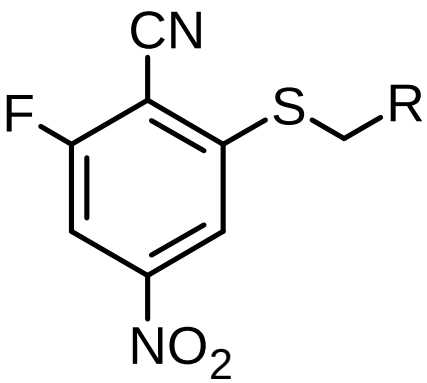


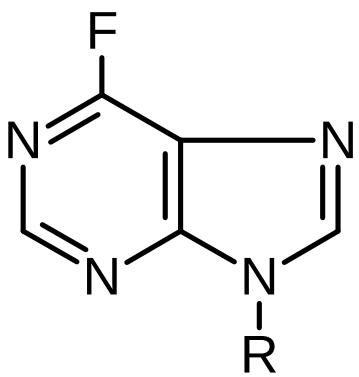


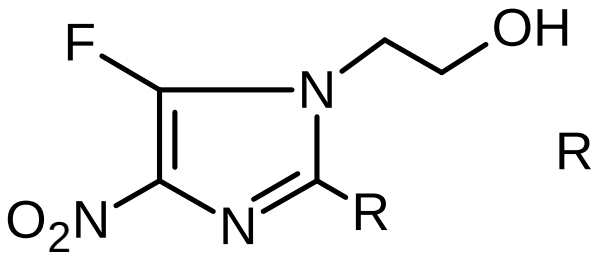
Nuclephilic Fluorination

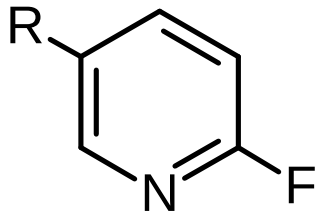
Via Quaternary Salts and Fluorodenitration





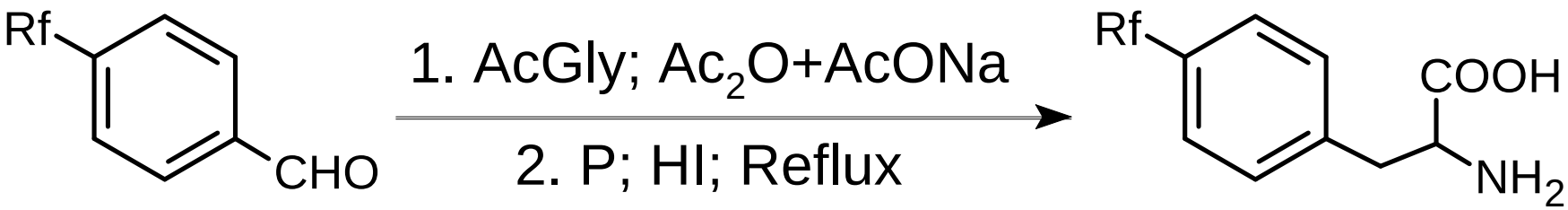






Synthesis from Fluorinated Blocks

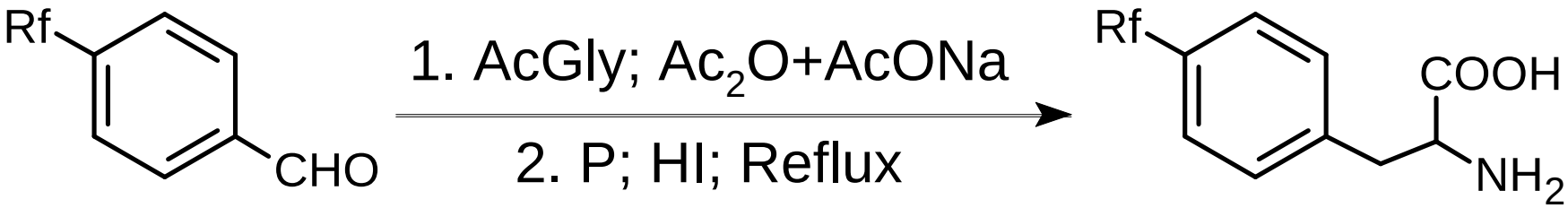
General Procedure



- **General Procedure:**
  - A mixture of N-acetylglycine, an aromatic aldehyde (10% molar excess), and NaOAc in Ac_2O was heated for 30 min. at 110-115 ^{o}C. Red phosphorus and more Ac_2O were added to the reaction mixture, followed by a dropwise addition of 47% HI. The resulting suspension was refluxed for 3.5 h with stirring, then filtered. The filtrate was evaporated to dryness in vacuo, and the residue was taken into a mixture of water and ether. The water layer was separated, extracted once more with ether and evaporated to dryness again. The residue was dissolved in water, and pH was adjusted to 5-6 with aqueous NH_3. The resulting white precipitate of amino acid was filtered off, washed with ice-cold water, and dried.

Synthesis from Fluorinated Blocks

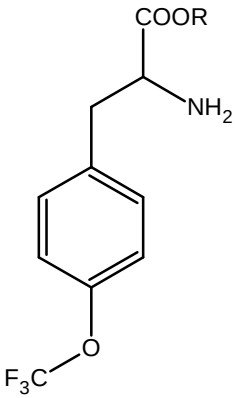
High Throughput Protocol

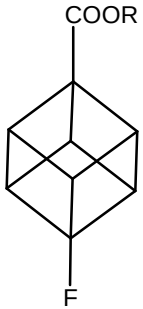


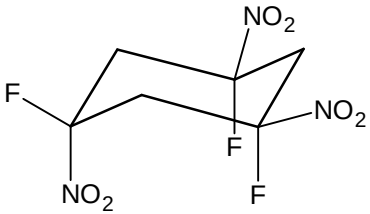
- **Protocol description:**
  - **General: Temperature: Heating and cooling. Pressure: No control (normal). Inert gas: No. Moisture protection: No.**
    - **Mixing of N-acetylglycine, an aromatic aldehyde, NaOAc and Ac_2O**
    - **Heating of the reaction mixture at 110-115 ^oC for 30 min**
    - **Addition of red phosphorus, more Ac_2O and 47% HI to the reaction mixture**
    - **Heating of the reaction mixture at 115-120 ^oC for 3.5 h …..**
- **Average success rate of the synthesis: 73%**
- **Average yield: 47% (32-62%)**
- **Average Purity: 86% (81-95%)**

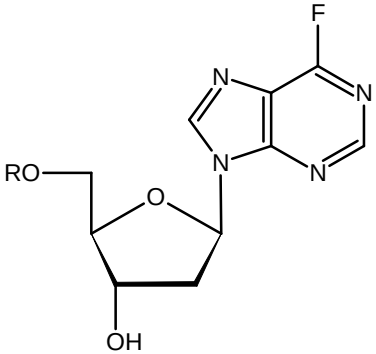
Additional Examples

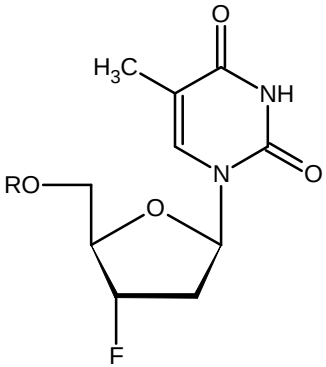
Fluoroorganic Compounds

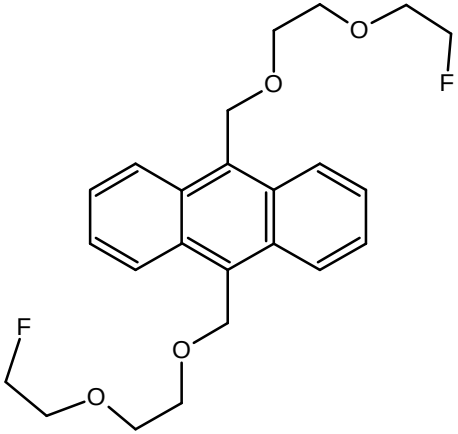


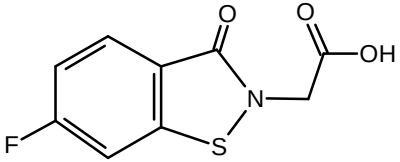






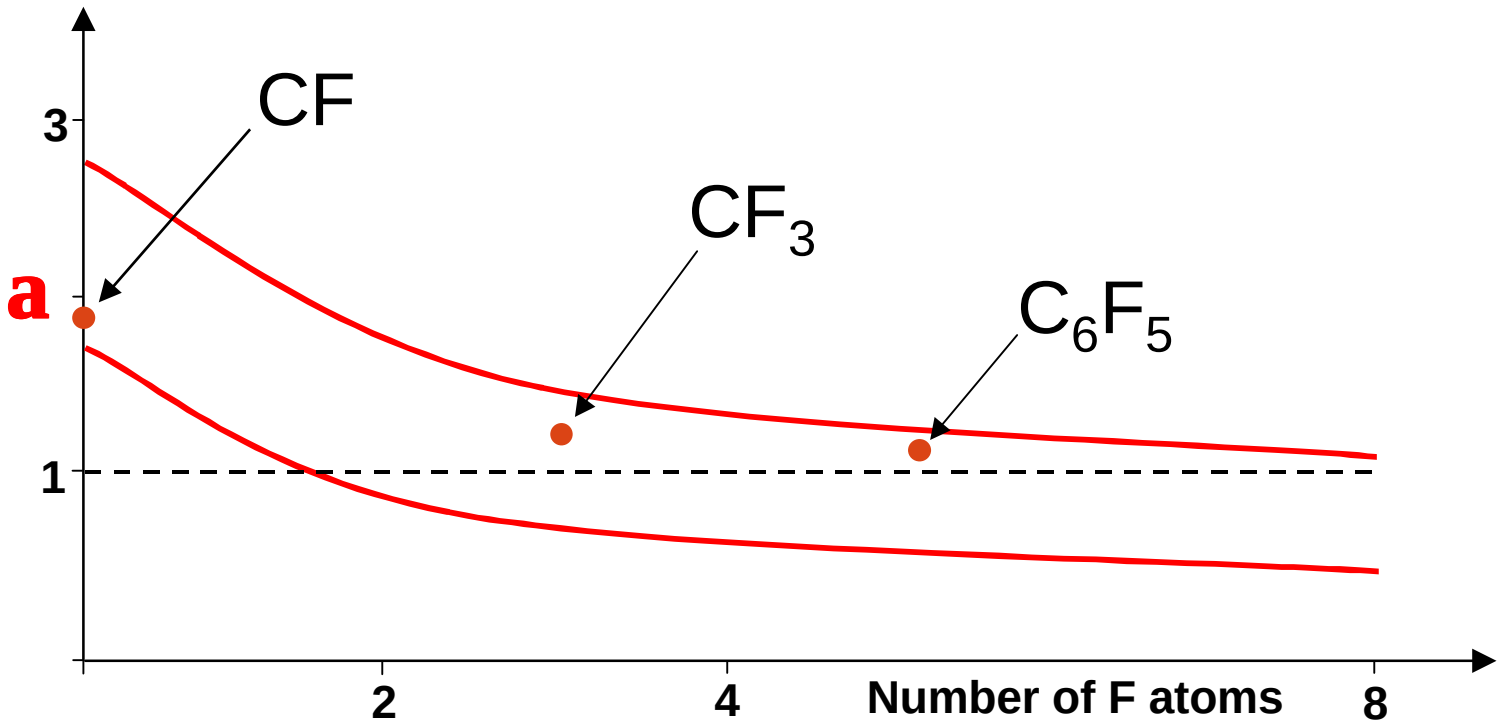






Further improvements

Acceleration Factor

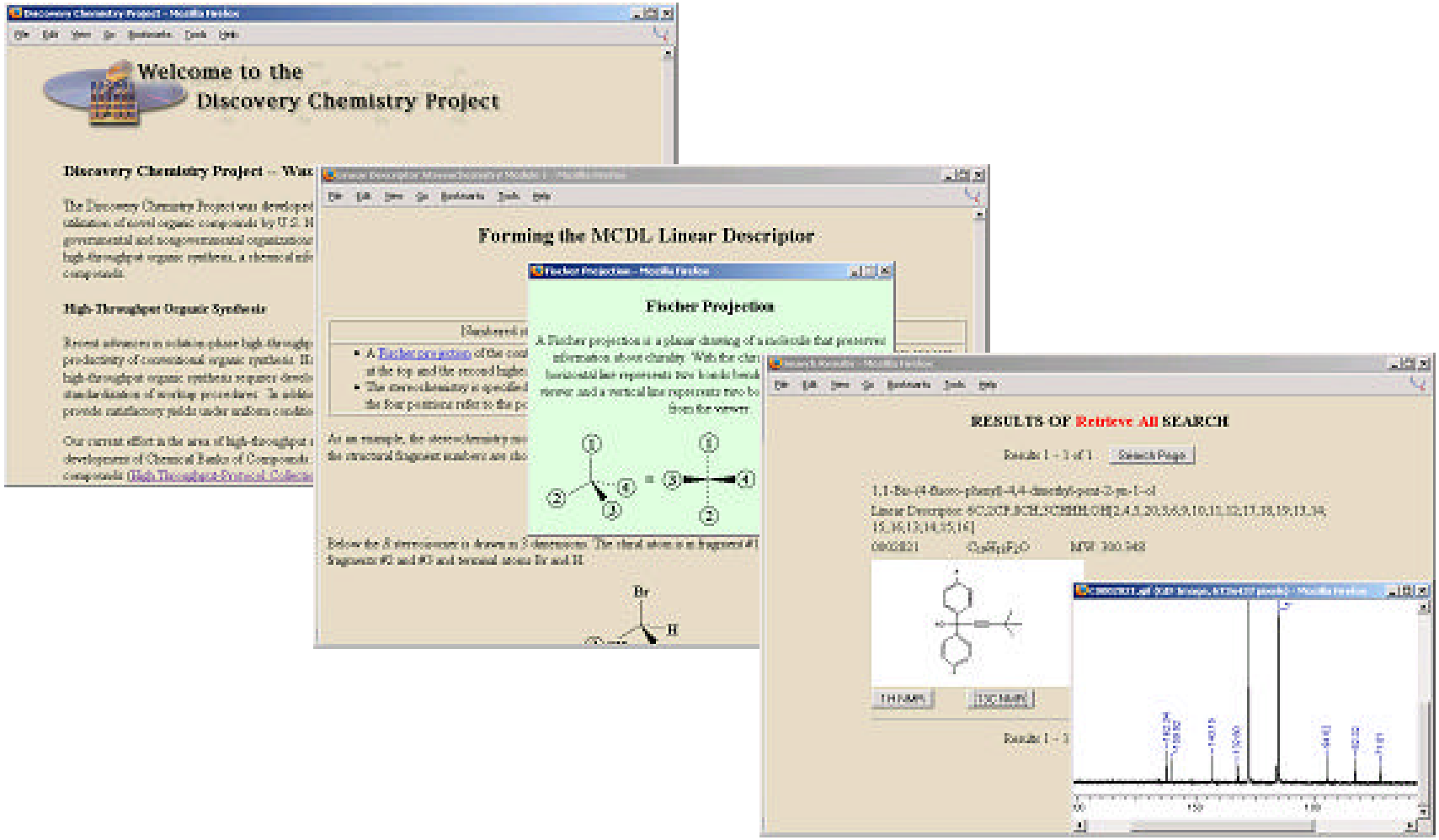


Acceleration coefficient **a** is affected by the number of fluorine atoms in a molecule. In this particular study the use of more than 3 fluorine atoms is counter-productive and actually reduces chances to find an active molecule.

Information Management System

- **Key features:**
  - An open source web-based solution (Linux, Perl, MySql, MCDL)
  - Internet distributed node architecture
  - Browser-based user interface (search, update, upload)
- **Current status**
  - Four fully functional independent nodes
- **Future developments:**
  - Automation of input/output protocols
  - Integration of Java MCDL editor
  - Improvement of the sub-structure search function
  - Standardization of new property-related modules (MCDL)
  - Implementation of XML-MCDL interconversion

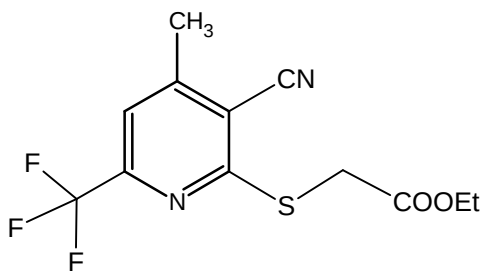
Information System Snapshots



Modular Chemical Descriptor Language (MCDL)

- **Basic features:**
  - Unique representation of a molecular structure (composition and connectivity modules) as well as atom- and bond-based stereochemistry (R,S and E,Z stereoisomers)
  - Prioritization algorithm is based on simple ASCII ordering (dictionary order in English)
  - Standardized as well as user-defined supplementary modules for compound-related information (spectra, physical-chemical data, references, activity)
- **Software implementation**
  - Current version of MCDL software (LINDES) is 1.5
  - Beta versions are available up to 2.4 (handles simple stereochemistry and salts)
  - C++ open code; compiled executables available for Unix (Linux), Windows and Power Mac OS

MCDL Example



• **4C;CFFF;CH;2CHH;2CHHH;CN;CO;N;O;S**	*Composition module (unique)*
[2,3,11;6,9;13,15; 5,6,13;;;10,14;12,15;;;;14]	*Connectivity module (unique)*
• **{CN:AUTONOM;(3-Cyano-4-methyl-6-trifluoromethyl -pyridin-2-ylsulfanyl)-Acetic-acid-Ethyl-ester}**	*Name*
• **{RE:J.Comb.Chem.,2000,2,24-28}**	*Reference*
• **{SM:EI;304(+M);231(+M-OEt);231(+M-OCOEt);M/Z}**	*Mass-spectrum*
• **{SI:KBr;1740(C=0);2226(CN);CM-1}**	*IR-Spectrum*
• **{SN:1H,DMSO-D6;1.3(T,J=7Hz,3H,CH3);2.61(S,3H,CH3); 4.0(s,2H,CH2S);4.2(q,J=7Hz,2H,CH2);7.31(S,1H,CH);PPM}**	*^{1}H NMR Spectrum*
• **{EA:C46.0;H3.6;N9.2;C12H11F3N2O2S;C47.3;H3.6;N9.2}**	*Elemental Analysis*
• **{YD:68%}**	*Yield*

Future Developments

Post-Combinatorial and Evolutional Chemistry

- **Background:**
  - Existing chemical banks of compounds are quite adequate for the preponderance of the Discovery Chemistry Program needs
  - Many biological targets have reliable 3D structures of active sites
- **Objective**
  - Speed-up discovery process (currently 2-3 years for lead identification and development)
  - Reduce cost (currently >2 FTE per active compound)
- **Approach**
  - Map active sites of biological targets using chemical banks of fluoroorganic compounds
  - Develop precise ligand match using natural selection/competition methodology

Acknowledgements

Discovery Chemistry Project Team and Collaborators

Dr. A.S. Astrat'ev
K.T. Barry
Dr. P.V. Bonnesen
Prof. J.C. Bryan
Dr. AC Buchanan
M.N. Burnett
Dr. D.J. Coughlin
V.P. Chernyshov
Dr. E.W. Hagaman
D.A. Harkins
Dr. W. Heilman
M.A. Janney
Dr. P.G. Kislitsin
Prof. A.L. Klibanov
T. Laudeman
Dr. D.W. Noid
Dr. M. Raikhstadt
Dr. A. Yu. Romanovich
Dr. R.S. Sachleben
Dr. A.V. Samet
Dr. V.V. Semenov
Dr. A.M. Shestopalov
Dr. M.A. Sigman
Dr. V.B. Sokolov
Dr. T. Spraggins
Dr. Y. Sterlelenko
Dr. B.G. Sumpter
Dr. S. Trepalin
L.D. Trowbridge
Prof. A.A.Tuinman
Dr. S.G. Zlotin

Major Funding for the Discovery Chemistry Project is provided by the IPP program. Oak Ridge National Laboratory is managed and operated by UT-Battelle, LLC, under contract DE-AC05-00OR22725 for the U.S. Department of Energy.