Design of Alternative Refrigerants via Global Optimization

Nikolaos V. Sahinidis
University of Illinois at Urbana-Champaign, Dept. of Chemical and Biomolecular Engineering, Urbana, IL 61801

Mohit Tawarmalani
Purdue University, Krannert School of Management, West Lafayette, IN 47907

Minrui Yu
Marconi Communications, Irving, TX 75063

Accurate and fast enumeration of large, combinatorial search spaces presents a central conceptual challenge in molecular design. To address this challenge, an algorithm is developed that guarantees globally optimal solutions to a mixed-integer nonlinear programming formulation for molecular design. The formulation includes novel structural feasibility constraints, while the algorithm provides all feasible solutions to this formulation through the implicit enumeration of a single branch-and-reduce tree. This algorithm is used to provide the complete solution set to a refrigerant design problem posed elsewhere. In addition to rediscovering CFCs, the proposed methodology identifies a number of novel potential replacements of Freon 12.

Introduction

Due to recent growing concerns regarding the depletion of the ozone layer by chlorofluorocarbons (CFCs), several actions have been taken to regulate the use of CFCs, including the 1985 Vienna Convention, the 1987 Montreal Protocol, the Clean Air Act of 1990, and the 1997 Kyoto Protocol. In response, industry ended the production of CFCs in developed countries in 1995 and introduced transitional substances, such as hydrofluorocarbons (HFCs) and hydrochlorofluorocarbons (HCFCs) to replace CFCs. However, it is now clear that these are not ideal alternatives. These replacements are not as energy efficient as earlier refrigerants. What is worse though, HCFCs still destroy the ozone layer and, although HFCs might cause no damage to the ozone layer, they have strong greenhouse effects.

The problem addressed in this work calls for the design of molecules with desired properties. Specifically, we are interested in the design of alternative refrigerants. Our goal is to develop a methodology that can serve as an aid to the design of replacement refrigerants by identifying candidates satisfying physical properties and that can be subsequently screened for high levels of energy efficiency and low levels of ozone depletion and global warming potential.

Traditional techniques in the design of new molecules with desired properties are based on the generate-and-test paradigm; a molecule is synthesized in the lab and its properties are measured through experiments. As these experiments are very time-consuming and expensive, computer-aided techniques have been developed in order to supplement experimental techniques by prescreening large numbers of possible molecules to select fewer to generate and test.

Conceptually, the problem of prescreening potential new molecules is very difficult due to its combinatorial nature. Consider a molecular design problem in which one seeks to synthesize a new compound by selecting up to N groups of atoms from a set of K potential groups. The total number of potential molecules in this case is known to be (Joback and Stephanopoulos 1995)

$$\sum_{i=2}^{N} \frac{(K+i-1)!}{i!(K-1)!}$$

This number of potential molecules explodes combinatorially. For example, with $K = 40$ groups, the total number of possible molecules is 12,300 for $N = 3$, and becomes 9,366,778 for $N = 6$ and 2,054,455,593 for $N = 9$. Coping with this combinatorial complexity presents a major conceptual challenge in computer-aided molecular design and requires the development of clever algorithms that select the optimal molecular design, while avoiding the explicit enumeration of all candidate designs.

In addition to the combinatorial aspect of the problem, the nature of current property estimation techniques also makes difficult the development of computer techniques for molecular design. Several formulae are available for estimating thermal, mechanical, and other physical properties of compounds (cf. Reid, et al. (1987); Horvath (1992)). Property prediction methods are typically nonlinear functions of atomic or group compositions. These nonlinearities lead to multiple local optima that may trap conventional search techniques in suboptimal parts of the search space. As a result, if standard optimization methods are used to rate the potential new compounds, one is faced with the multiple local optima difficulty.

Motivated by the above discussion, we ask:

• Is it possible to develop efficient techniques for globally searching the astronomically large space of combinatorial molecular designs without explicitly enumerating all possible designs?

• How should one deal with the nonlinearities and local optima traps that are inherent in property prediction techniques?

This article addresses the above issues through the development of a mixed-integer nonlinear programming model and a branch-and-reduce global optimization algorithm. Even though our basic modeling and algorithmic framework for molecular design is not restricted to any particular type of molecules, we demonstrate it in the context of a refrigerant design case study that was initially formulated in the seminal work of Joback and Stephanopoulos (1990). Numerous algorithms have been developed for this problem since it was posed. The main contributions of this article are two: (a) the development of a nonconvex mixed-integer nonlinear mathematical programming model that includes new structural feasibility constraints and captures the complexity of the model proposed by Joback and Stephanopoulos (1990, 1995), and (b) the development of an optimization algorithm that succeeds in finding not only a globally optimal solution, but all solutions that are feasible to this complex model. Finding all feasible solutions facilitates the systematic comparison of often competing objectives, including economic and environmental ones.

A formal statement of the problem is provided next, followed by a review of the current state-of-the-art in the field. Then, the optimization formulation is developed. Finally, our global optimization algorithm is described and computational results are presented.

Problem Statement

Qualitatively, the general problem we would like to address can be simply stated as: "Choose the constituent parts of a molecule in a way that satisfies chemical constraints and optimizes economic and environmental objectives."

To derive a mathematical abstraction of the problem, one can view the molecule as a collection of atoms or submolecular groups that are interlinked in a way that satisfies chemical connectivity and property constraints. The problem can be stated in the following general form of an optimization model

$$\max f(x,n)$$
$$\text{s.t. } g(x,n) \leq 0$$
$$x \in \mathfrak{R}^m, n \in Z^N$$

In this model, n is a vector of integer variables that are used here for the modeling of discrete decisions. In particular, n_j denotes the number of occurrences of the jth type of atom (or submolecular group) in the synthesized compound. The continuous variables x are used to model thermodynamic or other properties (heat capacity, enthalpy of vaporization, viscosity, and so on). It is the atomic/group composition n that dictates properties x through relationships captured in the constraint set $g(x,n) \leq 0$. Additional continuous variables might be used to model the operating conditions of the process in which the synthesized molecule will be used. This would allow for simultaneous product and process design. Finally, the design objective f is also expressed in terms of the discrete and continuous variables.

The formulation may incorporate explicit environmental requirements as $g(x,n) \leq 0$ can include constraints such as bounds on ozone depletion potential and global warming potential (for definitions of these terms and related discussion see Wuebbles (1981) and Wuebbles and Edmonds (1991)). At the same time, the objective function f may optimize traditional engineering objectives such as cost, profit, or energy use. Environmental objectives can also be optimized.

State-of-the-Art in the Field

The era of modern refrigeration began in the 1830s when the first refrigerants and refrigeration systems were proposed, developed, and patented. It took a period of a century before the first systematic approach to refrigerant design was developed by Midgely, Henne, and McNary. Starting with the periodic table of the elements, Midgely and co-workers eliminated those elements yielding unacceptable levels of volatility, instability, or toxicity. They were eventually left with the eight elements shown in Table 1. This approach quickly drew their attention to halogens and led to the discovery of CFCs. Nagengast (1995) provides an historical account of these developments. An enormous amount of work has been done, especially by industry, on issues related to refrigeration and alternative refrigerants since the development of CFCs. This work has relied mostly on careful experiments with potential refrigerants and is summarized in recent reviews by Calm

Table 1. Midgely's List of Candidate Elements from the Periodic Table

			H
C	N	O	F
		S	Cl
			Br

1994), Kuijpers (1994), Nagengast (1995), Smithart (1995), Thornton (1995), and Calm and Didion (1998). Here, we review only works that approach the problem from a theoretical angle, closely related to our work.

Gani and co-workers (Brignole et al., 1986; Gani et al., 1991; Gani and Fredenslund, 1993; Constantinou and Gani, 1994; Constantinou et al., 1996; Pretel et al., 1994) proposed a group contribution method for molecular design. Their approach starts by preselecting structural groups. It subsequently generates large numbers of combinations of the preselected groups. Properties of the generated molecules are then predicted using group contribution methods. Whereas these works relied mostly on explicit enumeration methods for exploring the space of combinatorial designs, several approaches have dealt with the development of efficient search techniques.

In a mathematical programming approach, Macchietto and co-workers (Macchietto et al., 1990; Odele and Macchietto, 1993), Naser and Fournier (1991), and Achenie and co-workers (Churi and Achenie, 1996, 1997a,b; Duvedi and Achenie, 1996) proposed the use of mixed-integer nonlinear programming techniques. Due to the nonconvex nature of their formulations, these approaches could at best guarantee locally optimal solutions.

The approach developed by Joback and Stephanopoulos (1990, 1995) overcomes the multiple local optima difficulty through the clever use of interval arithmetic techniques to construct bounds on the properties of "aggregate molecules," which are equivalent to constructing relaxations in the context of branch-and-bound search. Interval relaxations are weaker than mathematical programming bounds and, in general, require significant computational effort. To overcome this difficulty, Joback and Stephanopoulos (1990, 1995) developed an interactive heuristic search technique.

For polymer design, Vaidyanathan and El-Halwagi (1996) proposed the use of interval arithmetic techniques to solve a mixed-integer nonlinear formulation. For the same problem, Maranas (1996) has shown that the formulation can be linearized for certain fractional property prediction functions. For these types of prediction functions, a formulation based on chance-constraint programming can be used to model the probability of satisfying property constraints (Maranas, 1997a, b).

Recognizing the limitations of standard optimization methods, recent works have also proposed the use of guided stochastic search techniques (Venkatasubramanian et al., 1994, 1995, 1996; Devillers and Putavy, 1996; Ourique and Telles, 1998; Marcoulaki and Kokossis, 2000a,b). These approaches sample the search space in ways that increase the likelihood of locating good solutions.

Recent approaches to and applications of molecular design have included the design of mass separating agents for environmental impact minimization (Pistikopoulos and Stefanis, 1998), optimization with properties correlated with topological indices (Raman and Maranas, 1998; Camarda and Maranas, 1999; Siddhaye et al., 2000), the design of fiber reinforced polymer composite products (Vaidyanathan et al., 1998), benign solvent design (Sinha, et al., 1999), and a combinatorial approach based on linear (or linearized) group contribution techniques for generating candidate molecules (Friedler, et al., 1998). A recent collection of works on computer-aided molecular design is provided in Achenie, et al. (2002).

The above approaches represent major contributions towards automated molecular design. Issues that are still to be resolved include the following:

(1) The effect of inaccuracies in property prediction methods is, in general, not well-understood.

(2) All methods require preselecting the structural groups; there is no known theoretical justification for limiting the number and type of potential groups in the molecule.

(3) Rating synthesized compounds requires choosing an objective function. However, it is not well-understood what an appropriate objective function is. Current approaches use mostly thermodynamic objectives in the hope that the designed molecules will possess performance indices that are similar to those of existing compounds. However, this ignores environmental concerns.

(4) Even though physical and mechanical properties of new compounds can be predicted, environmental considerations such as flammability, toxicity, reactivity with the ozone layer, and, in general, any reactivity issues are much more difficult to predict. The current state-of-the-art in the modeling of such issues precludes them from being included in mathematical formulations.

(5) The search algorithms do not deal in a satisfactory way with the presence of multiple local optima in general nonlinear property prediction techniques.

The model/algorithm combination proposed in this article resolves the multiple local optima difficulty in the context of this problem. Furthermore, it is capable of providing all feasible solutions to the formulation. This makes it easier to deal with issues 3 and 4 above as only those compounds whose physical properties satisfy certain targets will need to be evaluated under reactivity, toxicity, flammability, and other environmental objectives. Finally, the proposed approach is so efficient computationally that it makes it routine to solve the problem for different property values without having to impose tight bounds on the number and type of potential groups to be considered.

Optimization Formulation

Previous attempts at this problem differ in the preselected groups that are allowed in the design, the property prediction techniques used, and the methodology employed in searching for new compounds. Our objective is to develop an approach that can handle a large number of preselected groups, while successfully overcoming the combinatorial search issue and local optima traps. Even though our approach is fairly general, to clearly illustrate its salient features, we detail it in the context of a case study that was posed initially by Joback and Stephanopoulos (1990) and addressed by many others in several variants (Gani et al., 1991; Pretel et al., 1994; Churi and Achenie, 1996; Duvedi, 1997a,b; and Achenie, 1996; Venkatasubramanian et al., 1995; Marcoulaki and Kokossis, 1998; Ourique and Telles, 1998).

The part of our model dealing with physical properties is identical to that of Joback and Stephanopoulos (1990). In order to model these authors' structural constraints, we will introduce integer and 0—1 variables so that mathematical programming bounds can be developed for the model. While do-
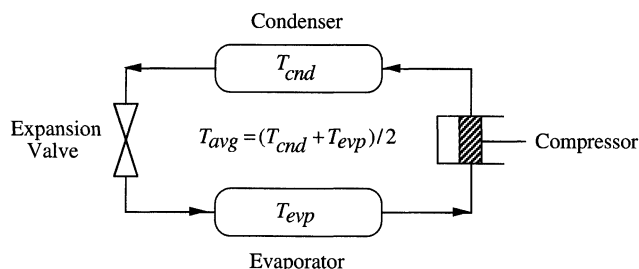
Figure 1. Refrigeration system considered.

ing so, we propose new structural constraints that eliminate some chemically infeasible solutions from these authors' original formulation, thus facilitating the solution process.

Design of an alternative to R12

The goal of the case study is to design an alternative refrigerant to dichlorodifluoromethane (CCl_2F_2), also known as *R*12. *R*12 was widely used in home refrigerators and automotive air conditioners. It is assumed that the refrigerant will be used in a typical automotive refrigeration cycle (see Figure 1) with an evaporation temperature $T_{evp} = 272.04$ K($= 30°$F), and an average operating temperature of $T_{avg} = 316.48$ K($= 70°$F).

The particular groups considered are shown in Table 2. In this table, "r" is used to denote bonds that are to be part of a ring only. We consider a total of $N = 44$ groups. In developing the model, the indices i and j will be used to denote groups ($i, j = 1, \ldots, N$). The formulation seeks to determine n_i, the number of groups of type i to be included in the synthesized compound. Other variables will be defined on an as-needed basis. A complete listing of variables and parameters is provided in the notation section.

Modeling of physical properties

Equations 1 through 4 reflect a group contribution method due to Joback and Reid (1987) to calculate the boiling temperature (T_b), critical temperature (T_c), critical pressure (P_c), and ideal gas heat capacity at average temperature (C_{p0a}), respectively, of the newly synthesized compound. These thermal properties will be used to derive other thermal properties in the sequel.

$$T_b = 198.2 + \sum_{i=1}^{N} n_i T_{bi} \quad (1)$$

$$T_c = \frac{T_b}{0.584 + 0.965 \sum_{i=1}^{N} n_i T_{ci} - \left(\sum_{i=1}^{N} n_i T_{ci} \right)^2} \quad (2)$$

$$P_c = \frac{1}{\left(0.113 + 0.0032 \sum_{i=1}^{N} n_i a_i - \sum_{i=1}^{N} n_i P_{ci} \right)^2} \quad (3)$$

$$C_{p0a} = \sum_{i=1}^{N} n_i C_{p0ai} - 37.93$$

$$+ \left(\sum_{i=1}^{N} n_i C_{p0bi} + 0.21 \right) T_{avg}$$

$$+ \left(\sum_{i=1}^{N} n_i C_{p0ci} - 3.91 \times 10^{-4} \right) T_{avg}^2$$

$$+ \left(\sum_{i=1}^{N} n_i C_{p0di} + 2.06 \times 10^{-7} \right) T_{avg}^3 \quad (4)$$

In addition to n_i, T_b is the only other variable appearing in the righthandside of Eqs. 1–4. Recall that T_{avg} is a given constant. All other parameters in these equations are constants whose values are taken from Joback and Reid (1987).

Using the above properties and the given operating temperatures of the cycle, Eqs. 5, 6, 7, and 8 calculate the reduced boiling temperature (T_{br}), the reduced average temperature (T_{avgr}), the reduced condensing temperature (T_{cndr}), and the reduced evaporating temperature (T_{evpr}), respectively

$$T_{br} = \frac{T_b}{T_c} \quad (5)$$

$$T_{avgr} = \frac{T_{avg}}{T_c} \quad (6)$$

$$T_{cndr} = \frac{T_{cnd}}{T_c} \quad (7)$$

$$T_{evpr} = \frac{T_{evp}}{T_c} \quad (8)$$

The next set of equations is used to calculate the acentric factor (ω)

$$\alpha = -5.97214$$

$$- \ln \left(\frac{P_c}{1.013} \right) + \frac{6.09648}{T_{br}} + 1.28862 \ln (T_{br}) - 0.169347 T_{br}^6$$

Table 2. Functional Groups Considered

Acyclic Groups	Cyclic Groups	Halogen Groups	Oxygen Groups	Nitrogen Groups	Sulfur Groups
$-CH_3$	$^r-CH_2-^r$	$-F$	$-OH$	$-NH_2$	$-SH$
$-CH_2-$	$_r>CH-^r$	$-Cl$	$-O-$	$>NH$	$-S-$
$>CH-$	$_r^r>CH-^r$	$-Br$	$^r-O-^r$	$_r^r>NH$	$^r-S-^r$
$>C<$	$_r^r>C<_r^r$	$-I$	$>CO$	$>N-$	
$=CH_2$	$_r>C<_r^r$		$_r^r>CO$	$=N-$	
$=CH-$	$>C<_r^r$		$-CHO$	$^r=N-^r$	
$=C<$	$^r=CH-^r$		$-COOH$	$-CN$	
$=C=$	$^r=C<_r^r$		$-COO-$	$-NO_2$	
$\equiv CH$	$^r=C<_r$		$=O$		
$\equiv C-$	$=C<_r^r$				

$$\beta = 15.2518 - \frac{15.6875}{T_{\mathrm{br}}} - 13.4721\ln\left(T_{\mathrm{br}}\right) + 0.43577 T_{\mathrm{br}}^{6}$$

$$\omega = \frac{\alpha}{\beta}$$

Using the acentric factor and previously calculated quantities, the liquid heat capacity (C_{pla}) is derived from the ideal gas heat capacity using Rowlinson's modification of Bondi's equation (Rowlinson, 1969)

$$C_{pla} = \frac{1}{4.1868}\left\{C_{p0a} + 8.314\left[1.45 + \frac{0.45}{1 - T_{\mathrm{avgr}}} + 0.25\,\omega \left(17.11 + 25.2\frac{(1 - T_{\mathrm{avgr}})^{1/3}}{T_{\mathrm{avgr}}} + \frac{1.742}{1 - T_{\mathrm{avgr}}}\right)\right]\right\}$$

The group contribution method of Joback and Reid (1987) is used to estimate the enthalpy of vaporization at boiling temperature (ΔH_{vb}), which is subsequently employed in Watson's method (Watson, 1943) to estimate the enthalpy of vaporization (ΔH_{ve}), at evaporating temperature

$$\Delta H_{vb} = 15.3 + \sum_{i=1}^{N} n_i \Delta H_{vb,i}$$

$$\Delta H_{ve} = \Delta H_{vb}\left(\frac{1 - T_{\mathrm{evp}}/T_c}{1 - T_b/T_c}\right)^{0.38}$$

We use the Riedel-Plank-Miller equation (Riedel, 1954) to obtain the reduced vapor pressure (P_{vpcr}) at condensing temperature, and that at evaporating temperature (P_{vper})

$$h = \frac{T_{br}\ln(P_c/1.013)}{1 - T_{br}}$$

$$G = 0.4835 + 0.4605h$$

$$k = \frac{h/G - (1 + T_{\mathrm{br}})}{(3 + T_{br})(1 - T_{br})^2}$$

$$\ln P_{\mathrm{vpcr}} = \frac{-G}{T_{\mathrm{cndr}}}\left[1 - T_{\mathrm{cndr}}^2 + k(3 + T_{\mathrm{cndr}})(1 - T_{\mathrm{cndr}})^3\right]$$

$$\ln P_{\mathrm{vper}} = \frac{-G}{T_{\mathrm{evpr}}}\left[1 - T_{\mathrm{evpr}}^2 + k(3 + T_{\mathrm{evpr}})(1 - T_{\mathrm{evpr}})^3\right]$$

Finally, the vapor pressure, (P_{vpc}) at condensing temperature and the vapor pressure (P_{vpe}) at evaporating temperature can be calculated from the corresponding reduced vapor pressures P_{vpcr} and P_{vper}

$$P_{\mathrm{vpc}} = P_{\mathrm{vpcr}} P_c$$

$$P_{\mathrm{vpe}} = P_{\mathrm{vper}} P_c$$

The physical property estimation relationships are summarized in Figure 2. Here, T_b, T_c, P_c, C_{p0a}, and ΔH_{vb} are the
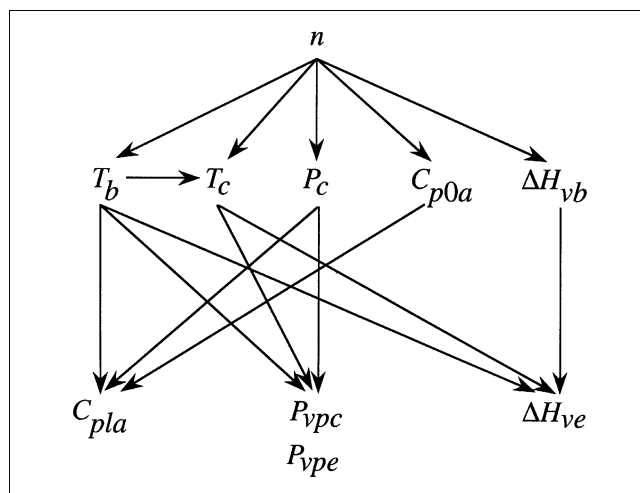

Figure 2. Physical property estimation relationship.

primary physical properties whose values are obtained by group contribution methods as a function of the molecular structure n. Values of primary physical properties are used to determine other physical properties of interest such as C_{pla}, ΔH_{ve}, P_{vpc}, and P_{vpe}.

Structural constraints

The purpose of the structural constraints is to eliminate combinations of groups that do not satisfy bonding requirements. In this section, we begin by developing an integer programming formulation of the structural constraints proposed by Joback and Stephanopoulos (1995). Then, we discuss the structural constraints proposed by Odele and Macchietto (1993) and present examples that demonstrate that the structural constraints proposed by Joback and Stephanopoulos (1995) and Odele and Macchietto (1993) are not sufficient for eliminating all infeasible group combinations. Finally, we develop new structural constraints that tighten the structural feasibility formulations of Joback and Stephanopoulos (1995) and Odele and Macchietto (1993).

Integer programming formulation of the Joback-Stephanopoulos structural constraints

The numbering of the structural constraints below follows Table II of Joback and Stephanopoulos (1995). The validity of these constraints follows from graph connectivity arguments: they are necessary constraints to ensure that the graph formed by the selected groups is connected.

(1) The first structural constraint requires that at least two groups be present in the molecule

$$\sum_{i=1}^{N} n_i \geq 2$$

(2) If groups with acyclic bonds and groups with cyclic bonds are to be combined in a molecule, there must exist groups with both bond types

$$\sum_{i \in \mathfrak{M}} n_i \geq 1 \quad \text{if} \quad \sum_{i \in \mathfrak{A}} n_i \geq 1 \quad \text{and} \quad \sum_{i \in C} n_i \geq 1$$

where $\mathfrak{M}$, $\mathfrak{A}$, and $\mathfrak{C}$, respectively, denote the union of groups having both acyclic and cyclic bonds, the union of groups having only acyclic bonds, and the union of groups having only cyclic bonds. To model the above logical condition, we define binary variables Y_M, Y_A, and Y_C, where:

$$Y_M = \begin{cases} 1, & \text{if groups having both acyclic bonds are selected} \\ 0, & \text{otherwise} \end{cases}$$

$$Y_A = \begin{cases} 1, & \text{if groups having only acyclic bonds are selected} \\ 0 & \text{otherwise} \end{cases}$$

$$Y_C = \begin{cases} 1, & \text{if group having only cyclic bonds are selected} \\ 0, & \text{otherwise}. \end{cases}$$

Then, we introduce the constraints

$$Y_A \leq \sum_{i \in \mathfrak{A}} n_i \leq N_{\max} Y_A \| \mathfrak{A} \| \tag{9}$$

$$Y_C \leq \sum_{i \in \mathfrak{C}} n_i \leq N_{\max} Y_C \| \mathfrak{C} \|$$

$$Y_M \leq \sum_{i \in \mathfrak{M}} n_i \leq N_{\max} Y_M \| \mathfrak{M} \|$$

where $\| \mathfrak{A} \|$, $\| \mathfrak{C} \|$, and $\| \mathfrak{M} \|$ are the cardinalities of $\mathfrak{A}$, $\mathfrak{C}$, and $\mathfrak{M}$, respectively. The binary variables should further satisfy

$$Y_M \geq Y_A + Y_C - 1 \tag{10}$$

which forces $Y_M = 1$ whenever $Y_A = Y_C = 1$.

(3) Structural stability also requires that, if one selects groups with both acyclic and cyclic bonds, there must be then at least one acyclic group or one cyclic group

$$\sum_{i \in \mathfrak{A}} n_i + \sum_{i \in \mathfrak{C}} n_i \geq 1 \text{ if } \sum_{i \in \mathfrak{M}} n_i \geq 1$$

This requirement is captured by

$$Y_A + Y_C \geq Y_M \tag{11}$$

which forces $Y_A + Y_C \geq 1$ whenever $Y_M = 0$. The combined effect of Eqs. 10 and 11 is that

$$\sum_{i \in \mathfrak{M}} n_i \geq 1 \text{ if and only if } \sum_{i \in \mathfrak{A}} n_i \geq 1 \text{ and } \sum_{i \in \mathfrak{C}} n_i \geq 1$$

(4) The number of groups with cyclic bonds must be at least three for cyclic molecules

$$3 \leq \sum_{i \in \mathfrak{R}} n_i \leq N_{\max} \| \mathfrak{R} \| \quad \text{if} \quad \sum_{i \in \mathfrak{R}} n_i \neq 0$$

where $\mathfrak{R}$ is the set of groups having at least one cyclic bond (that is, the union of $\mathfrak{M}$ and $\mathfrak{C}$), and $\| \mathfrak{R} \|$ is the cardinal-

ity of $\mathfrak{R}$. The above condition can be modeled through

$$3Y_R \leq \sum_{i \in \mathfrak{R}} n_i \leq N_{\max} Y_R \| \mathfrak{R} \|$$

where

$$Y_R = \begin{cases} 1, & \text{if groups with cyclic bonds are selected} \\ 0, & \text{otherwise} \end{cases}$$

(5) The next equation states that the number of groups having an odd number of free bonds must be even

$$\sum_{i \in \mathfrak{B}} n_i = 2 Z_B$$

where $\mathfrak{B}$ is the set of groups having an odd number of free bonds and Z_B is a nonegative integer variable. For different integer values of Z_B, this equation forces the number of groups having an odd number of free bonds to be an integer multiple of two.

(6) The number of bonds (one-half the number of available groups bonds) must be at least equal to the number of groups minus one

$$\sum_{i=1}^{N} n_i b_i \geq 2 \left(\sum_{i=1}^{N} n_i - 1 \right)$$

(7) The number of bonds formed cannot exceed that of a complete graph in which all groups are connected.

$$\sum_{i=1}^{N} n_i b_i, \leq \left(\sum_{i=1}^{N} n_i \right) \left(\sum_{i=1}^{N} n_i - 1 \right)$$

(8) In the absence of rings, if there exist groups with acyclic single bonds, but no double bonds and other groups with acyclic double bonds but no single bonds, then there must exist transition groups containing both acyclic single and acyclic double bonds. In addition, groups with both acyclic single and triple bonds must exist if groups with acyclic single but no triple bonds are selected along with groups with triple but no single bonds. These requirements can be stated as follows

$$\sum_{i \in SD} n_i \geq 1 \quad \text{if} \quad \sum_{i \in S/D} n_i \geq 1 \quad \text{and} \quad \sum_{i \in D/S} n_i \geq 1 \tag{12}$$

$$\sum_{i \in ST} n_i \geq 1 \quad \text{if} \quad \sum_{i \in S/T} n_i \geq 1 \quad \text{and} \quad \sum_{i \in T/S} n_i \geq 1 \tag{13}$$

with the additional requirement that Eq. 12 should be enforced only in the absence of rings. Both constraints are of the form: "*Condition x must be enforced if conditions y and z are met.*" To model Eq. 12, we begin by defining binary variables Y_{SDx}, Y_{SDy}, and Y_{SDz}, where

$$Y_{SDx} = \begin{cases} 1, & \text{if } \Sigma_{i \in SD} n_i \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$Y_{SDy} = \begin{cases} 1, & \text{if } \Sigma_{i \in S/D} n_i \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$Y_{SDz} = \begin{cases} 1, & \text{if } \Sigma_{i \in D/S} n_i \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

Then, we introduce the constraints

$$Y_{SDx} \leq \sum_{i \in \mathcal{S}\mathcal{D}} n_i \leq N_{\max} Y_{SDx} \| SD \| \tag{14}$$

$$Y_{SDy} \leq \sum_{i \in \mathcal{S}/\mathcal{D}} n_i \leq N_{\max} Y_{SDy} \| S/D \| \tag{15}$$

$$Y_{SDz} \leq \sum_{i \in \mathcal{D}/\mathcal{S}} n_i \leq N_{\max} Y_{SDz} \| D/S \| \tag{16}$$

$$Y_{SDy} + Y_{SDz} - 1 \leq Y_{SDx} + Y_R \tag{17}$$

where $\| \mathcal{S}\mathcal{D} \|$, $\| \mathcal{S}/\mathcal{D} \|$, and $\| \mathcal{D}/\mathcal{S} \|$ are the cardinalities of $\mathcal{S}\mathcal{D}$, $\mathcal{S}/\mathcal{D}$, and $\mathcal{D}/\mathcal{S}$, respectively. Constraints 14–16 force variables Y_{SDx}, $Y_{\mathcal{S}\mathcal{D}y}$ and $Y_{\mathcal{S}Dz}$ to a value of one when the corresponding condition is met. Constraint 17 enforces the logical requirement that $Y_{SDz} = 1$ if $Y_{SDy} = Y_{SDz} = 1$ and $Y_R = 0$. The modeling of Eq. 13, similarly, requires the introduction of a triplet of binary variables (Y_{STx}, Y_{STy}, Y_{STz}) along with a constraint set analogous to Eqs. 14–17 above, although the binary Y_R should not be used in this case.

(9) If a collection of groups contains mixed groups, then the number of acyclic groups with a single bond can be no more than the number of free bonds of all other groups

$$\sum_{i \in \mathcal{Q}|b_i=1} n_i \leq \sum_{i \in \mathfrak{M}} (b_i - 2) n_i + \sum_{i \in \mathcal{Q}|b_i \geq 3} (b_i - 2) n_i$$

$$\text{if } \sum_{i \in \mathfrak{M}} n_i > 0$$

Using the binary variable Y_M defined above, the last requirement can be expressed as

$$\sum_{i \in \mathcal{Q}|b_i=1} n_i \leq \sum_{i \in \mathfrak{M}} (b_i - 2) n_i + \sum_{i \in \mathcal{Q}|b_i \geq 3} (b_i - 2) n_i$$
$$+ N_{\max}(1 - Y_M) \| \mathcal{Q}_1 \|$$

where $\| \mathcal{Q}_1 \|$ is the number of acyclic groups with only one bond.

(10) The following relationship needs to be enforced between the number of groups with one, three, and four bonds for pure acyclic molecules

$$\sum_{i \in \mathcal{Q}, b_i=1} n_i - \sum_{i \in \mathcal{Q}, b_i=3} n_i - 2 \sum_{i \in \mathcal{Q}, b_i=4} n_i$$

$$= 2 \quad \text{if } \sum_{i \in \mathcal{R}} n_i = 0$$

This is modeled as

$$lY_R \leq \sum_{i \in \mathcal{Q}, b_i=1} n_i - 2 - \sum_{i \in \mathcal{Q}, b_i=3} n_i - 2 \sum_{i \in \mathcal{Q}, b_i=4} n_i \leq uY_R \tag{18}$$

where

$$l = -N_{\max}(\| \mathcal{Q}_3 \| + 2 \| \mathcal{Q}_4 \|)$$
$$u = N_{\max} \| \mathcal{Q}_1 \|$$

with $\| \mathcal{Q}_1 \|$, $\| \mathcal{Q}_3 \|$, and $\| \mathcal{Q}_4 \|$ being the cardinalities of $\mathcal{Q}_1 = \{i | i \in \mathcal{Q}, b_i = 1\}$, $\mathcal{Q}_3 = \{i | i \in \mathcal{Q}, b_i = 3\}$, and $\mathcal{Q}_4 = \{i | i \in \mathcal{Q}, b_i = 4\}$, respectively.

(11) The number of available bond types (acyclic single, cyclic single, acyclic double, cyclic double, and acyclic triple) must be even

$$\sum_{i \in \mathcal{S}} n_i \mathcal{S}_i = 2Z_S \tag{19}$$

$$\sum_{i \in \mathcal{S}\mathcal{R}} n_i S_i^R = 2Z_{SR} \tag{20}$$

$$\sum_{i \in \mathcal{D}} n_i D_i = 2Z_D \tag{21}$$

$$\sum_{i \in \mathcal{D}\mathcal{R}} n_i D_i^R = 2Z_{DR} \tag{22}$$

$$\sum_{i \in T} n_i T_i = 2Z_T \tag{23}$$

where Z_S, Z_{SR}, Z_D, Z_{DR}, and Z_T are nonnegative integer variables effectively counting the number of acyclic single, cyclic single, acyclic double, cyclic double, and acyclic triple bonds in the designed compound.

Odele-Macchietto structural constraints

Although clearly necessary, the above structural constraints are not sufficient for structural feasibility. For example, consider the following collection of groups: 2 ($_r^r > C <_r^r$), 1($= C <_r^r$), and 1 ($= O$). Although this collection of groups satisfies all the above constraints, it does not form a feasible molecule.

Odele and Macchietto (1993) presented two constraints in a different approach to structural feasibility. Their first constraint (also used earlier by Macchietto et al. (1990)) employs the octet rule to prohibit free attachments

$$\sum_i n_i(2 - b_i) = 2m$$

where

$$m = \begin{cases} 1, & \text{for acylic compounds} \\ 0, & \text{for monocyclic compounds} \\ -1, & \text{for bicyclic compounds} \end{cases}$$

The case of $m = 1$ is exactly the structural constraint Eq. 18. The case of $m \neq 1$ eliminates the combination 2 ($_r^r > C <_r^r$), 1

$(=C<^r_r)$, and 1 ($=O$). We prefer to not use the latter constraint, though, as it does not allow for more than two cycles.

The second constraint used in Odele and Macchietto (1993) requires that two adjacent groups not be linked by more than one bond

$$\sum_{i=1}^{N} n_i \ge n_j(b_j - 1) + 2 \quad j \in \mathfrak{A}$$

While the Odele-Macchietto (O—M) constraints eliminate some group contributions that are feasible to the Joback-Stephanopoulos (J—S) constraints, the reverse is also true. For instance, the selection of 2 ($-Br$) and 1 ($=\mathfrak{O}$) is feasible to the O—M constraints, but not to J—S constraints.

Even though valid, the above structural feasibility constraints do not eliminate all infeasible group combinations from consideration. For instance, consider the following collection of groups: 1 ($>CH-$), 3 ($=O$), and 1 ($=N-$). While chemically infeasible, this set of groups satisfies the J—S, as well as the O—M structural constraints, and motivates the development of additional structural feasibility constraints as follows.

New structural constraints

Constraints between One-Bonded and Higher-Bonded Groups in the Absence of Rings. In order to enforce the obvious requirement that two acyclic bonds of the same group may not be connected, an additional set of constraints needs to be added to the above structural constraint sets. We first define a "one-bonded group" to be a group with only one acyclic (single, double, triple) bond. Similarly, a group having more than one acyclic bond will be referred to as a "higher-bonded group." It is clear that a one-bonded group must be connected either to a higher-bonded group or to another one-bonded group.

Clearly, the number of one-bonded groups connected to higher-bonded groups may not exceed the number of free attachments of higher-order groups

$$\sum_{\substack{i \in \mathfrak{O}, \\ \text{single-bonded}}} n_i \le \sum_{i \in \mathfrak{IC}} n_i S_i^H \quad \text{if} \quad \sum_{i \in \mathfrak{IC}} n_i \ge 1$$
$$\text{and} \quad \sum_{i \in \mathfrak{R}} n_i = 0 \quad (24)$$

$$\sum_{\substack{i \in \mathfrak{O}, \\ \text{double-bonded}}} n_i \le \sum_{i \in \mathfrak{IC}} n_i D_i^H \quad \text{if} \quad \sum_{i \in \mathfrak{IC}} n_i \ge 1$$
$$\text{and} \quad \sum_{i \in \mathfrak{R}} n_i = 0 \quad (25)$$

$$\sum_{\substack{i \in \mathfrak{O}, \\ \text{triple-bonded}}} n_i \le \sum_{i \in \mathfrak{IC}} n_i T_i^H \quad \text{if} \quad \sum_{i \in \mathfrak{IC}} n_i \ge 1$$
$$\text{and} \quad \sum_{i \in \mathfrak{R}} n_i = 0 \quad (26)$$

where $\mathfrak{O}$ and $\mathfrak{IC}$, respectively, denote the sets of one-bonded and higher-bonded groups, while S_i^H, D_i^H, and T_i^H denote

the available acyclic single, double, and triple bonds of the ith higher-bonded group. The modeling of Eqs. 24–26 requires the introduction of a binary variable to monitor the selection of higher-bonded groups

$$Y_H = \begin{cases} 1, & \text{if } \Sigma_{i \in \mathfrak{IC}} n_i \ge 1 \\ 0, & \text{otherwise} \end{cases}$$

and the following constraints

$$Y_H \le \sum_{i \in \mathfrak{IC}} n_i \le N_{\max} Y_H \parallel \mathfrak{IC} \parallel \tag{27}$$

$$\sum_{\substack{i \in \mathfrak{O}, \\ \text{single-bonded}}} n_i \le \sum_{i \in \mathfrak{IC}} n_i S_i^H + N_{\max}(1 - Y_H + Y_R) \parallel \mathfrak{S}/\mathfrak{D} \parallel$$
$$\tag{28}$$

$$\sum_{\substack{i \in \mathfrak{O}, \\ \text{double-bonded}}} n_i \le \sum_{i \in \mathfrak{IC}} n_i D_i^H + N_{\max}(1 - Y_H + Y_R) \parallel \mathfrak{S}/\mathfrak{D} \parallel$$
$$\tag{28}$$

$$\sum_{\substack{i \in \mathfrak{O}, \\ \text{triple-bonded}}} n_i \le \sum_{i \in \mathfrak{IC}} n_i T_i^H + N_{\max}(1 - Y_H + Y_R) \parallel \mathfrak{S}/\mathfrak{D} \parallel$$
$$\tag{30}$$

where $\parallel \mathfrak{S} \mathfrak{D} \parallel$, $\parallel \mathfrak{S}/\mathfrak{D} \parallel$, and $\parallel \mathfrak{D}/\mathfrak{S} \parallel$ are the cardinalities of $\mathfrak{S} \mathfrak{D}$, $\mathfrak{S}/\mathfrak{D}$, and $\mathfrak{D}/\mathfrak{S}$, respectively. Constraint 27 enforces the definition of variable Y_H, while constraints 28–30 enforce conditions 24–26.

When a one-bonded group is connected to another one-bonded group, the two one-bonded groups will simply form a molecule even in the presence of additional groups. This situation can be avoided through the following relationship between the number of higher-bonded groups and their free attachments left after being connected to those one-bonded groups

$$\sum_{i \in \mathfrak{IC}} n_i(S_i^H + D_i^H + T_i^H) - \sum_{i \in \mathfrak{O}} n_i = 2\left(\sum_{i \in \mathfrak{IC}} n_i - 1\right)$$

The last constraint is applicable only if there are no rings. Thus, it should be included in the problem formulation as

$$-\Omega Y_R \le \sum_{i \in \mathfrak{IC}} n_i(S_i^H + D_i^H + T_i^H) - \sum_{i \in \mathfrak{O}} n_i - 2\left(\sum_{i \in \mathfrak{IC}} n_i - 1\right)$$
$$\le \Omega Y_R \tag{31}$$

with a sufficiently large value for Ω.

Note that the last constraint eliminates the group combination: 1 ($>CH-$), 3 ($=O$), and 1 ($=N-$), which is feasible to the J—S and O—M structural constraints.

Clique Constraints for Rings. Any two groups can be bonded through a single bond—a ring bond or an acyclic bond. Let n_R be the number of molecular groups that have ring bonds. Then, the number of ring bonds in the final design is bounded from above by $n_R(n_R - 1)/2$. If the number of ring bonds equals this upper bound, then all ring-bonded

groups form a clique and, thus, cannot share an acyclic bond. More generally, the constraint we would like to enforce is that the total number of bonds involving ring groups cannot exceed the number of bonds in the case of a clique, that is, $n_R(n_R - 1)/2$. Define

$$Y_{\text{clique}}^R = \begin{cases} 1, & \text{if the ring groups form a clique} \\ 0, & \text{otherwise} \end{cases}$$

and

$$n_R = \sum_{i \in \mathfrak{R}} n_i$$

Then

$$Z_{SR} + Z_{DR} + 1 - Y_{\text{clique}}^R \leq \frac{n_R(n_R - 1)}{2} \tag{32}$$

where Z_{SR} and Z_{DR} are used in Eqs. 20 and 22 and define the number of single and double ring bonds. Clearly, the sum of single and double ring bonds cannot exceed the maximum possible number of bonds, which is given by the righthand-side of Eq. 32. In the case of a clique, the righthand-side must be reduced by at least one. Thus, when $1 - Y_{\text{clique}}^R$ is added to the lefthand-side, constraint 32 simply forces Y_{clique}^R to be 1 if there is a clique. Note that constraint 32 by itself eliminates the group combination: 2 ($_\text{r}$ > C < $_\text{r}$), 1 (= C < $_\text{r}$), and 1 (= O), which was feasible to the J—S as well as the O—M structural constraints.

We further note that, if $Y_{\text{clique}}^R = 1$, the number of single acyclic bonds in the acyclic groups must be greater than or equal to the number of acyclic bonds in the ring groups. This requirement can be enforced through the following constraint

$$Y_{\text{clique}}^R \sum_{i \in \mathfrak{A}} n_i S_{ai} \geq Y_{\text{clique}}^R \sum_{i \in \mathfrak{R}} n_i S_{ai} \tag{33}$$

This last constraint will eliminate the group combination: 3 ($_\text{r}$ > CH—$^\text{r}$) and 1 (—F), which is feasible to all previous structural constraints.

The concept used to derive Eq. 33 can be extended to the case of double acyclic bonds

$$Y_{\text{clique}}^R \sum_{i \in \mathfrak{A}} n_i D_{ai} \geq Y_{\text{clique}}^R \sum_{i \in R} n_i D_{ai}$$

Property targets and other bounds

This case study requires that the new refrigerant's enthalpy of vaporization at T_{evp} be greater than or equal to that of $R12$ (18.4 KJ/g − mol)

$$\Delta H_{ve} \geq 18.4 \tag{34}$$

The motivation behind the above constraint is that a higher ΔH_{ve} reduces the amount of refrigerant required. Also, a smaller liquid heat capacity C_{pla} reduces the amount of refrigerant vapor generated in the expansion valve. For this

reason, the case study requires that the new refrigerant's liquid heat capacity at T_{avg} be less than or equal to that of $R12$ (32.2 cal/g − mol·K)

$$C_{pla} \leq 32.2 \tag{35}$$

We require that the new refrigerant's vapor pressure at T_{evp} be greater than or equal to 1.4 bar

$$P_{vpe} \geq 1.4 \tag{36}$$

The above constraint is required as the lowest pressure in the refrigeration cycle must be greater than atmospheric pressure (1.1 bar) to reduce the possibility of air and moisture leaking into the system.

Equation 37 requires that the new refrigerant's vapor pressure at T_{cnd} be less than 14 bar. This is because a high system pressure increases the size, weight, and cost of the equipment. A pressure ratio of 10 is considered the maximum allowable for a refrigeration cycle.

$$P_{vpc} \leq 14 \tag{37}$$

Finally, we have nonnegativity and integrality requirements

$$C_{p0a}, C_{pla}, \Delta H_{vb}, \Delta H_{ve} \geq 0$$

$$P_c, P_{vpc}, P_{vpcr}, P_{vpe}, P_{vper} \geq 0$$

$$T_{avgr}, T_b, T_{br}, T_c, T_{cndr}, T_{evpr} \geq 0$$

$$h, G \geq 0$$

$$\beta \leq O$$

$$Y_i = 0 \text{ or } 1 \quad i \in \{M, A, C, R, H\}$$

$$Y_i = 0 \text{ or } 1 \text{ for } i \in \{SDx, SDy, SDz\}$$

$$Y_i = 0 \text{ or } 1 \text{ for } i \in \{STx, STy, STz\}$$

$$Y_{\text{clique}}^R = 0 \text{ or } 1$$

$$Z_i: \text{nonnegative integer, for } i \in \{B, S, SR, D, DR, T\}$$

$$n_R: \text{nonnegative integer}$$

$$0 \leq n_i \leq N_{\max}, \quad n_i: \text{integer for } i = 1, \ldots, N$$

$$\alpha, \omega, k: \text{unrestricted in sign}$$

Note that bounds on most continuous and integer variables of the problem can be inferred through interval arithmetic operations using the bounds in Eqs. 34−37, the non-negativity requirements, and the problem constraints. Table 3 shows bounds on all continuous and integer variables that were used in all calculations. As seen, fairly generous bounds on all problem variables were used. For instance, up to 15 occurrences of each structural group are allowed in the final design.

Objective function

As explained above, high enthalpy of vaporization and low liquid heat capacity are desirable as far as the efficiency of

Table 3. Variable Bounds

	Lower Bound	Upper Bound		Lower Bound	Upper Bound
α	-5.4	6100	P_{vpe}	1.4	14
β	-16000	0	P_{vper}	0.001	1
ΔH_{vb}	10	110	T_{avgr}	0.001	1
ΔH_{ve}	18.4	100	T_b	50	1000
ω	-1	1.3	T_{br}	0.001	1
C_{p0a}	1	135	T_c	100	2000
C_{pla}	1	32.2	T_{cndr}	0.001	1
G	0.48	2500	T_{evpr}	0.001	1
h	0	5300	Z_B	0	172
k	-1	10	Z_D	0	60
n_i	0	15	Z_{DR}	0	30
P_c	2	200	Z_S	0	330
P_{vpc}	1.4	14	Z_{SR}	0	232
P_{vpcr}	0.001	1	Z_T	0	15

the refrigeration cycle is concerned. The design objective for this case study is to find a refrigerant at least as efficient as $R12$, that is, a refrigerant with a larger heat of vaporization, ΔH_{ve} and a smaller liquid heat capacity C_{pla}, than that of $R12$. These two objectives are combined into that of maximizing their ratio

$$\max \frac{\Delta H_{ve}}{C_{pla}} \qquad (38)$$

Remarks on the Formulation

(1) The objective function in Eq. 38 was chosen for consistency with the earlier articles that studied the same problem. It should be obvious that many other objective functions are possible and might be more meaningful. For example, one could use an expression for the coefficient of performance of the refrigeration cycle, or a measure of the ozone depletion potential of the refrigerant, or a combination of economic and environmental objectives.

(2) The resultant formulation is a mixed-integer optimization model with multiple local optima due to the presence of the highly nonlinear property estimation formulae.

(3) By considering 44 candidate groups and allowing up to 15 members of each group to be present in the final solution, the total number of combinatorial designs exceeds 10^{85}. Most of these will be rendered infeasible by the model constraints.

(4) Constraints 34 and 35 were introduced based on optimality arguments. Alternatively, one may relax these constraints.

(5) For this case study, the model involves 52 integer, 14 binary (0–1) variables, 21 continuous variables, and 92 constraints (this includes the J–S and a subset of the O–M structural feasibility constraints).

(6) The righthand side of constraint (9) can also be recast as

$$n_i \le N_{\max}Y_A \qquad i \in \mathcal{C}$$

From an algorithmic point of view, disaggregating this and similar constraints is beneficial since it usually leads to tighter relaxations. This disaggregation was used in our computations and increased the number of problem constraints to 236.

(7) The solution to the above model provides a set of atoms and submolecular groups to be combined to form a molecule. However, the model provides no information on *how* the selected groups should exactly be bonded to form a connected molecule. The bonding of the selected groups can be determined by solving another optimization problem. In particular, consider a graph whose nodes correspond to the groups selected by the above model. Using the valences of the selected groups, assign arcs for all possible group interconnections over this graph. Then, define binary variables to select arcs subject to constraints satisfying valence requirements, the absence of submolecules, and the absence of cycles made up by acyclic bonds. All feasible solutions to this model would need to be found; this is within the capabilities of the algorithm developed in the next section. However, as the computational results section shows, this model does not need to be solved for the specific refrigerant design problem addressed in this paper. In particular, we show that, for this application, the solution set consists of small molecules whose bonding can be easily obtained by inspection.

Global Optimization Algorithm

In the description of the algorithm, we use P to denote the MINLP model to be solved over the set X of variable bounds shown in Table 3. Let P_k be problem P defined over $X_k \subseteq X$. The following algorithm integrates constraint programming with mathematical programming tools to solve P.

Algorithm branch-and-reduce

Initialization Step. Set the upper and lower bounds of the search tree $UB \leftarrow +\infty$ and $LB \leftarrow -\infty$. Put problem P into $\mathcal{L}$, the list of "active" problems to be solved. Set $k = 1$ and begin iteration.

Iteration k:

Step 1. *Selection:* If $\mathcal{L}$ is empty, stop. Otherwise, select and remove P_k from $\mathcal{L}$. Go to Step 2.

Step 2. *Constraint Propagation:* Preprocess P_k using feasibility-based range reduction techniques.

Step 3. *Bounding:* Lower and upper bound P_k. Update the lower and upper bounds of the search tree and delete inferior nodes. If the current node is deleted, go to Step 1. Otherwise, go to Step 4.

Step 4. *Postprocessing:* Postprocess P_k using both feasibility-based and optimality-based range reduction techniques. If the bounds on variables are substantially improved, go to Step 3. Otherwise, go to Step 5.

Step 5. *Branching:* Choose a variable on which to branch. Divide the range of the variable at some point to create subdomains X_{k_1} and X_{k_2}, and add two subproblems P_{k_1} and P_{k_2} to $\mathcal{L}$. Set $k \leftarrow k + 1$ and go to Step 1.

We note that the above algorithm applies to the solution of a general class of MINLP models. Thus, it could be used to solve a variety of molecular design problems that are modeled as MINLPs, including those that utilize group contribution methods that require information about nearest neighbors and connectivity of the entire molecule (cf. Achenie et al. (2002)).

Refer to Ryoo and Sahinidis (1995, 1996), Ghildyal and Sahinidis (2001), Shectman and Sahinidis (1998), and Tawar-

malani and Sahinidis (1999, 2002b), for algorithmic details; and Dorneich and Sahinidis (1995), Liu et al. (1996), Sahinidis and Tawarmalani (2000), and Tawarmalani and Sahinidis (2002b) for other applications of this algorithm. In these works, we provide detailed discussions on the construction of suitable relaxations, convergent branching rules, and feasibility- and optimality-based range reduction rules for the global optimization of mixed-integer nonlinear programs. We detail here only two problem-specific steps.

Lower bounding and feasibility-based range reduction

Figure 2 makes it clear that, for given values of the integer variables of the formulation, a forward calculation through appropriate problem constraints makes it possible to compute all physical properties involved and, thus, test feasibility of the given integer point. Therefore, instead of relying on time-consuming local searches to identify feasible solutions, at every node of the branch-and-reduce tree, we round the relaxation solution and use a single pass through the appropriately ordered set of constraints in order to evaluate physical properties and test for feasibility of the rounded solution. Improved solutions are used to update the incumbent.

We note that the same ordering of constraints and interval arithmetic operations can be used to infer bounds on the continuous problem variables once the branch-and-reduce algorithm has reduced the ranges of some of the integer variables. This leads to a very efficient mechanism for collapsing variable bounds as the search progresses deeper in the tree.

Finding all feasible solutions

It is common practice to find alternative feasible solutions to an integer programming formulation by introducing the binary expansions

$$n_i = \sum_{j=1}^{\lfloor \log_2(N_{max}) \rfloor} 2^{j-1} y_{ij}$$

Once the problem is solved in terms of binary variables, it is straightforward to exclude a binary combination y^k by introducing the well-known *linear* integer cut

$$\sum_{(i,j) \in \mathfrak{B}^k} y_{ij} - \sum_{(i,j) \in \mathfrak{N}^k} y_{ij} \leq \| \mathfrak{B}^k \| - 1$$

where $\mathfrak{B}^k = \{(i,j) | y_{ij}^k = 1\}$, $\mathfrak{N}^k = \{(i,j) | y_{ij}^k = 0\}$, and $\| \mathfrak{B}^k \|$ is the cardinality of $\mathfrak{B}^k$. Alternatively, one can work directly in the space of the integer variables by introducing the following more intuitive, yet *nonlinear*, cut

$$\sum_i \left(n_i^k - n_i \right)^2 \geq 1$$

Several solutions of the optimization formulation can then be obtained by solving a series of models in which integer cuts are successively introduced to exclude the previous models' optimal solutions from further consideration. This allows the identification of as many feasible solutions as desired. How-

ever, such an approach requires the search of a number of branch-and-bound trees.

Instead of the successive introduction of integer cuts, we prefer to modify Step 3 of the branch-and-reduce algorithm as follows:

Step 3. *Bounding:* Lower and upper bound P_k. Update the lower and upper bounds of the search tree, *record all feasible solutions found,* and delete *the current node if it is infeasible or all integer variables are fixed*. If the current node is deleted, go to Step 1. Otherwise, go to Step 4.

In other words, instead of deleting all inferior nodes, we delete only the current node when it becomes infeasible or a point. Nodes where feasible solutions are identified are branched further until they become points in the search space or infeasible. The advantage of this scheme vs. using the integer cuts is that all feasible solutions are identified through a single application of branch-and-reduce, thus avoiding a great deal of duplicate work in repetitive search trees.

It is important to note that, once Step 3 of the algorithm has been modified as above, care must be exercised in the application of the optimality-based range reduction and probing techniques of Ryoo and Sahinidis (1995). These techniques require an appropriate lower bound for optimality-based fathoming. Instead of using the lower bounds provided by feasible solutions identified during the search, we make use of a bound obtained by solving a linear relaxation of the problem.

Finally, we note that it is straightforward to modify this algorithm to provide only the K best solutions of the model for any prespecified integer value of K. In every iteration, one would need to maintain the K best solutions found thus far and delete nodes probably worse than the Kth best known solution. In this case, the Kth best known solution provides an appropriate lower bound to be used for optimality-based range reduction and probing in the formulae of Ryoo and Sahinidis (1995).

Computational Results

The proposed algorithm was implemented using BARON (Sahinidis, 1996), a general purpose global optimization package. Two modifications to the general purpose solver were made for the specific problem: the problem-specific lower bounding step and the fathoming rule for the identification of all feasible solutions. The capability to identify all, or the K best, solutions of an integer program was added to the general purpose BARON solver after successful implementation in the context of this molecular design problem.

All runs were performed on a 332 MHz IBM RS/6000 workstation model 43P with 128 MB memory and a LINPACK score of 59.9. Solution of the linear programming relaxation subproblems was done through CPLEX 6.0 (CPLEX, 1997).

We began computational experimentations with this problem and the branch-and-reduce algorithm in the mid-1990s. The initial attempts did not include the specialized local search algorithm and were based on an early implementation of BARON. At that time, over 150 h of CPU time were required for the algorithm to identify a few feasible solutions, but global optimality was not proven and the solutions identi-

Table 4. Complete Set of Solutions to the Refrigerant Design Problem

	Molecular Structure	$\dfrac{\Delta H_{ve}}{C_{pla}}$		Molecular Structure	$\dfrac{\Delta H_{ve}}{C_{pla}}$
FNO	$F-N=0$	1.2880	C_2HO_2F	$F-O-CH=C=0$	0.8705
CHClO	$Cl-CH=O$	1.1804	C_3H_4	$CH_2=C=CH_2$	0.8656
FSH	$F-SH$	1.1697	C_2H_6	CH_3-CH_3	0.8632
CH_3Cl	CH_3-Cl	1.1219	C_3H_3FO	$(F-)(CH3-)>C=C=O$	0.8531
C_2HClO_2	$O=C<(-CH=O)(-Cl)$	1.1207	NHF_2	$F-NH-F$	0.8468
CClFO	$(F-)(Cl-)>C=O$	1.0179	C_3H_3FO	$(F-)(O=CH-)>C=CH_2$	0.8291
C_3H_2O	$CH\equiv C-CH=0$	1.002	CHO_3F	$O=CH-O-O-F$	0.8288
$CFONH_2$	$O=CH-NH-F$	0.9893	C_2HOF	$CH\equiv C-O-F$	0.8263
ClFO	$Cl-O-F$	0.9822	C_3H_3F	$CH\equiv C-CH_2-F$	0.7802
C_3H_2O	$CH_2=C=C=O$	0.9672	CHF_2Cl	$(F-)(F-)>CH-Cl$	0.7770
C_3H_4O	$CH_3-CH=C=O$	0.9619	C_2H_3OF	$CH_2=CH-O-F$	0.7685
C_3H_3O	$CH_2=CH-CH=O$	0.9408	NF_2Cl	$(F-)(F-)>N-Cl$	0.7658
C_3H_4	$CH_3-C\equiv CH$	0.9278	C_2H_6NF	$(CH3-)(CH3-)>N-F$	0.6817
C_2F_2	$F-C\equiv C-F$	0.9299	N_2HF_3	$(F-)(F-)>N-NH-F$	0.6711
CH_2ClF	$F-CH2-Cl$	0.9202	$C_2H_2OF_2$	$CH2=C<(-O-F)(-F)$	0.6705
$C_2H_4O_2$	$CH3-O-CH=O$	0.9178	$C_3H_2F_2$	$(F-)(F-)>CH-C\equiv CH$	0.6686
CHNOFCl	$O=CH-N<(-F)(-Cl)$	0.9167	C_2HNF_2	$CH\equiv C-N<(-F)(-F)$	0.6587
C_3H_6O	$(CH_3-)(CH_3-)>C=O$	0.8978	$C_3H_4F_2$	$(F-)(F-CH_2-)>C=CH_2$	0.6377
C_3H_3FO	$F-CH=CH-CH=O$	0.8868	$C_3H_4F_2$	$(F-)(F-)>CH-CH=CH_2$	0.6263
C_3HFO	$(F-)(CH\equiv C-)>C=O$	0.8785	$C_2H_3NF_2$	$CH2=CH-N<(-F)(-F)$	0.6176
C_3H_6O	$CH_3-CH_2-CH=O$	0.8773	CH_3NOF_2	$(F-)(CH_3-)>N-O-F$	0.6139
$CNHF_2O$	$O=C<(-F)(-NH-F)$	0.8763	$C_3H_3F_3$	$(_r>CH-^r)_3(-F)_3$	0.5977

fied were not globally optimal. Currently, BARON takes 45 min. of CPU time to find all feasible solutions of the above model. This amounts to producing candidate compounds at the rate of one solution per minute. What made this possible is a number of equally important developments:

Table 5. Admissible Solutions

	Molecular Structure	$\dfrac{\Delta H_{ve}}{C_{pla}}$
FNO	$F-N=O$	1.2880
FSH	$F-SH$	1.1697
CH_3Cl	CH_3-Cl	1.1219
C_2HClO_2	$O=C<(-CH=O)(-Cl)$	1.1207
ClFO	$(Cl-)(-O-)(-F)$	0.9822
C_3H_4O	$CH_3-CH=C=O$	0.9619
C_3H_4	$CH_3-C\equiv CH$	0.9278
C_2F_2	$F-C\equiv C-F$	0.9229
CH_2ClF	$F-CH2-Cl$	0.9202
C_2HO_2F	$F-O-CH=C=O$	0.8705
C_3H_4	$CH_2=C=CH_2$	0.8656
C_2H_6	CH_3-CH_3	0.8632
C_3H_3FO	$(F-)(CH3-)>C=C=O$	0.8531
NHF_2	$F-NH-F$	0.8468
C_2HOF	$CH\equiv C-O-F$	0.8263
C_3H_3F	$CH\equiv C-CH_2-F$	0.7802
CHF_2Cl	$(F-)(F-)>CH-Cl$	0.7770
C_2H_3OF	$CH_2=CH-O-F$	0.7685
NF_2Cl	$(F-)(F-)>N-Cl$	0.7658
C_2H_6NF	$(CH3-)(CH3-)>N-F$	0.6817
N_2HF_3	$(F-)(F-)>N-NH-F$	0.6711
$C_2H_2OF_2$	$CH2=C<(-O-F)(-F)$	0.6705
$C_3H_2F_2$	$(F-)(F-)>CH-C\equiv CH$	0.6686
C_2HNF_2	$CH\equiv C-N<(-F)(-F)$	0.6587
$C_3H_4F_2$	$(F-)(F-CH_2-)>C=CH_2$	0.6377
$C_3H_4F_2$	$(F-)(-F)>CH-CH=CH_2$	0.6263
$C_2H_3NF_2$	$CH2=CH-N<(-F)(-F)$	0.6176
CH_3NOF_2	$(F-)(CH_3-)>N-O-F$	0.6139
$C_3H_3F_3$	$(_r>CH-^r)_3(-F)_3$	0.5977

• A feasibility-based presolver (Ghildyal and Sahinidis, 2001) that reduces the upper bounds of group occurrences from 15 to five at the root node of the search tree;

• Aggressive probing to reduce ranges of five variables by solving minimization/maximization LPs on each node of the branch-and-reduce tree;

• The entirely linear programming based relaxation of Tawarmalani and Sahinidis (1999), which eliminated the need for the solution of convex nonlinear programming relaxations that are often hard to solve with current nonlinear programming technology;

• The branching technique of Tawarmalani and Sahinidis (1999), which is based on an elaborate mechanism that selects the branching and probing variables from among the original problem variables, as well as their functional expressions that appear in the mathematical program;

• The development of tighter relaxations based on the product disaggregation concept of Tawarmalani et al. (2002) that, in turn, is based on the theory of convex extensions (Tawarmalani & Sahinidis 2001, 2002a);

• Implementation of the specialized lower bounding scheme that made possible the identification of good feasible solutions pretty early during the search without the need to resort to a nonlinear programming solver;

• Development of the algorithm to find all feasible solutions through a single branch-and-reduce tree.

The complete set of solutions identified for the above model is shown in Table 4. The results demonstrate some of the weaknesses, as well as strengths of the group contribution approach to molecular design.

In Table 4, there are 15 compounds involving groups from column one of Table 2 and $=O$ or $-O-$. While mathematically acceptable, such combinations lead to inaccurate prop-

erty estimation as compared to selecting the mixed groups directly from the fourth column of Table 2. Clearly, the latter choice leads to compounds that do not satisfy the imposed property requirements. Nanda (2001) develops a set of constraints that render such combinations infeasible in the MINLP formulation. Inclusion of these constraints to our formulation further reduces the CPU time to about 30 min.

Once the 15 inadmissible compounds are eliminated, the solutions of Table 5 remain. Among the remaining 29 compounds, one notices CH_3Cl, a well-known former refrigerant. One also observes that, among these 29 compounds, there are 41, 38, 5, 11, and 8 groups containing C, F, Cl, O, and N, respectively. It is rather remarkable that the model clearly points towards CFCs as the ideal refrigerants.

Among the solutions of Table 5, all but the following nine compounds were identified in earlier works on this problem

(1) $F—N=O$
(2) $Cl—O—F$
(3) $CH_3—CH=C=O$
(4) $F—O—CH=C=O$
(5) $CH_2=C=CH_2$
(6) $(F—)(CH3—)>C=C=0$
(7) $CH\equiv C—O—F$
(8) $CH_2=CH—O—F$
(9) $CH2=C<(—O—F)(F—)$

Except for the first compound on this list (nitrosyl fluoride), which is available commercially, minimal information currently exists in the literature about the other compounds. While all of these compounds seem to be unstable, toxic, or otherwise not fit to be used as refrigerants, their identification demonstrates that the proposed methodology has the potential of discovering novel compounds.

Conclusions

This article has demonstrated that the state-of-the-art in global optimization of mixed-integer nonlinear programs coupled with a suitable formulation has now reached the level of maturity required to solve realistic models for molecular design. In addition to rediscovering CFCs, this methodology has yielded a number of potential replacement refrigerants.

Several possible future directions are suggested by the results of this article. First, it would be interesting to identify classes of compounds for which the search algorithm and structural feasibility constraints developed in this article are sufficient to resolve the combinatorial difficulty of the molecular design search problem. As our algorithm and structural feasibility constraints are independent of the particular molecule being designed, they could be used for the design of industrial solvents, polymers, and pharmaceuticals. In the context of these compounds, it may be necessary to extend the model to use more complex group contribution techniques, such as those that account for interactions between groups and group connectivity. Finally, it would be interesting to systematically experiment with the applicability of the proposed MINLP algorithm to solve molecular design problems with larger search spaces. Promising results in the solution of problems with thousands of feasible molecular structures were recently reported (Sahinidis, 2002). In the context of the *Freon 12* replacement problem itself, relaxing some of the property target constraints and expanding the search region seems a promising future research direction.

Acknowledgments

Notation

Indices

i, j = Index of group i or j, where $i, j = 1, \ldots, N$

Index sets

$\mathscr{A}$ = union of groups having only acyclic bonds
$\mathscr{A}_k$ = union of acyclic groups having exactly k bonds ($k = 1, 3,$ or 4)
$\mathscr{B}$ = union of groups having odd number of free bonds
$\mathscr{C}$ = union of groups having only cyclic bonds
$\mathscr{D}$ = union of groups having acyclic double bonds
$\mathscr{D}/\mathscr{S}$ = union of groups having acyclic double bonds and no acyclic single bonds
$\mathscr{D}\mathscr{R}$ = union of groups having cyclic double bonds
$\mathscr{H}$ = union of groups having more than one acyclic bond
$\mathscr{M}$ = union of groups having both acyclic and cyclic bonds
$\mathscr{O}$ = union of groups having only one acyclic bond
$\mathscr{R}$ = union of groups having cyclic bonds
$\mathscr{S}$ = union of groups having acyclic single bonds
$\mathscr{S}/\mathscr{D}$ = union of groups having acyclic single bonds and no double bonds
$\mathscr{S}/\mathscr{J}$ = union of groups having acyclic single bonds and no triple bonds
$\mathscr{S}\mathscr{D}$ = union of groups having acyclic single bonds and acyclic double bonds
$\mathscr{S}\mathscr{R}$ = union of groups having cyclic single bonds
$\mathscr{S}\mathscr{J}$ = union of groups having acyclic single bonds and acyclic triple bonds
$\mathscr{J}$ = union of groups having acyclic triple bonds
$\mathscr{J}/\mathscr{S}$ = union of groups having acyclic triple bonds and no single bonds

Parameters

$\Delta H_{vb,i}$ = contribution of group i to enthalpy of vaporization at boiling temperature
a_i = number of atoms of group i
b_i = available free bonds of group i
C_{p0ai} = constant contribution of group i to ideal gas heat capacity
C_{p0bi} = first-order contribution of group i to ideal gas heat capacity
C_{p0ci} = second-order contribution of group i to ideal gas heat capacity
C_{p0di} = third-order contribution of group i to ideal gas heat capacity
D_i = number of available acyclic double bonds of group i
D_i^H = number of available acyclic double bonds of higher-bonded group i
D_i^R = number of available cyclic double bonds of group i
N = total number of group types in the model, $N = 44$
N_{max} = maximum number of each group allowed, $N_{max} = 15$
P_{ci} = contribution of group i to critical pressure
S_i = number of available acyclic single bonds of group i
S_i^H = number of available acyclic single bonds of higher-bonded group i
S_i^R = number of available cyclic single bonds of group i
T_i = number of available acyclic triple bonds of group i
T_i^H = number of available acyclic triple bonds of higher-bonded group i
T_{avg} = average temperature, $T_{avg} = 294.26$ K $= (70°F)$

T_{bi} = contribution of group i to boiling temperature
T_{ci} = contribution of group i to critical temperature
T_{cnd} = condensing temperature, T_{cnd} = 316.48 K = (110°F)
T_{evp} = evaporating temperature, T_{evp} = 272.04 K = (30°F)

Variables

α = intermediate variable used to calculate acentric factor ω
β = intermediate variable used to calculate acentric factor ω
ΔH_{vb} = enthalpy of vaporization at boiling temperature, KJ/g−mol
ΔH_{ve} = enthalpy of vaporization at evaporating temperature, KJ/g−mol
ω = acentric factor
C_{p0a} = ideal gas heat capacity at average temperature, cal/g−mol·K
C_{pla} = liquid heat capacity at average temperature, cal/g−mol·K
G = intermediate variable used in the Riedel-Plank-Miller equation
h = intermediate variable used in the Riedel-Plank-Miller equation
k = intermediate variable used in the Riedel-Plank-Miller equation
n_i = number of groups of type i selected
P_c = critical pressure, bar
P_{vpc} = vapor pressure at condensing temperature, bar
P_{vpcr} = reduced vapor pressure at condensing temperature
P_{vpe} = vapor pressure at evaporating temperature, bar
P_{vper} = reduced vapor pressure at evaporating temperature
T_{avgr} = reduced average temperature
T_b = boiling temperature, K
T_{br} = reduced boiling temperature;
T_c = critical temperature, K
T_{cndr} = reduced condensing temperature
T_{evpr} = reduced evaporating temperature
Y_i = 0−1 variables modeling structural disjunctions, $i \in \{ \mathfrak{M}, \mathfrak{A}, \mathfrak{C}, \mathfrak{R} \}$
Z_i = integer variables modeling structural requirements, $i \in \{ \mathfrak{B}, \mathfrak{S}, \mathfrak{SR}, \mathfrak{D}, \mathfrak{DR}, \mathfrak{I} \}$

Literature Cited

Achenie, L. E. K., R. Gani, and V. Venkatasubramanian, eds., *Computer Aided Molecular Design: Theory and Practice*, Elsevier, New York (2002).

Brignole, E. A., S. B. Bottini, and R. Gani, "A Strategy for the Design and Selection of Solvents for Separation Processes," *Fluid Phase Equilibria*, **29**, 125 (1986).

Calm, J. M., "Refrigerant Safety," *ASHRAE J.*, 17 (1994).

Calm, J. M., and D. A. Didion, "Trade-offs in Refrigerant Selections: Past, Present, and Future," *Int. J. Refrig.*, **21**, 308 (1998).

Camarda, K. V., and C. D. Maranas, "Optimization in Polymer Design Using Connectivity Indices," *Ind. Eng. Chemistry Res.*, **38**: 1884−1892 (1999).

Churi, N., and L. E. K. Achenie, "Novel Mathematical Programming Model for Computer Aided Molecular Design," *Ind. Eng. Chemistry Res.*, **35**, 3788 (1996).

Chrui, N., and L. E. K. Achenie, "On the Use of a Mixed Integer Nonlinear Programming Model for Refrigerant Design," *Ind. Eng. Chemistry Res.*, **35**, 3788 (1996).

Chrui, N., and L. E. K. Achenie, "On the Use of a Mixed Integer Nonlinear Programming Model for Refrigerant Design," *Int. Trans. of Operational Res.*, **4**, 45 (1997a).

Churi, N., and L. E. K. Achenie, "The Optimal Design of Refrigerant Mixtures for a Two-Evaporator Refrigeration System," *Comput. Chem. Eng.*, **21**, S349 (1997b).

Constantinou, L., K. Bagherpour, R. Gani, J. A. Klein, and D. T. Wu, "Computer Aided Product Design: Problem Formulations, Methodology and Applications," *Comput. Chem. Eng.*, **20**, 685 (1996).

Constantinou, L., and R. Gani, "New Group Contribution Method for Estimating Properties of Pure Compounds," *AIChE J.*, **40**, 1697 (1994).

CPLEX, *CPLEX 6.0 User's Manual*, ILOG CPLEX Division, Incline Village, NV (1997).

Devillers, J., and C. Putavy, "Designing Biodegradable Molecules from the Combined Use of a Backpropagation Neural Network and a Genetic Algorithm," *Genetic Algorithms in Molecular Design*, J. Devillers, ed., Academic Press, London, pp. 303−314 (1996).

Dorneich, M. C., and N. V. Sahinidis, "Global Optimization Algorithms for Chip Layout and Compaction," *Eng. Optimization*, **25**, 131 (1995).

Duvedi, A. P., and L. E. K. Achenie, "Designing Environmentally Safe Refrigerants Using Mathematical Programming," *Chem. Eng. Sci.*, **51**, 3727 (1996).

Friedler, F., L. T. Fan, L. Kalotai, and A. Dallos, "A Combinatorial Approach for Generating Candidate Molecules with Desired Properties based on Group Contribution," *Comput. Chem. Eng.*, **22**, 809 (1998).

Gani, R., and A. Fredenslund, "Computer-Aided Molecular and Mixture Design with Specified Constraints," *Fluid Phase Equilibria*, **82**, 39 (1993).

Gani, R., B. Nielsen, and A. Fredenslund, "A Group Contribution Approach to Computer-Aided Molecular Design," *AIChE J.*, **37**, 1318 (1991).

Ghildyal, V., and N. V. Sahinidis, "Solving Global Optimization Problems with BARON," *From Local to Global Optimization. A Workshop on the Occasion of the 70th Birthday of Professor Hoang Tuy*, A. Migdalas, P. Pardalos, and P. Varbrand, eds., Kluwer Academic Publishers, Boston, MA, pp. 205−230 (2001).

Horvath, A. L., *Molecular Design: Chemical Structure Generation from the Properties of Pure Organic Compounds*, Elsevier, New York (1992).

Joback, K. G., and R. C. Reid, "Estimation of Pure-Component Properties from Group-Contributions," *Chem. Eng. Comm.*, **57**, 233 (1987).

Joback, K. G., and G. Stephanopoulos, "Designing Molecules Possessing Desired Physical Property Values, *Proc. of 1989 Foundations of Computer-Aided Process Design Conf.*, Snowmass, CO, Elsevier, Amsterdam, 195 (1990).

Joback, K. G., and G. Stephanopoulos, "Searching Spaces of Discrete Solutions: The Design of Molecules Possessing Desired Physical Properties," *Advances in Chem. Eng.*, **21**, 257 (1995).

Kuijpers, L., "Refrigerants: The Present Situation and Future Prospects," *ASHRAE J.*, 22 (1994).

Liu, M. L., N. V. Sahinidis, and J. P. Shectman, "Planning of Chemical Process Networks via Global Concave Minimization," *Global Optimization in Engineering Design*, I. E. Grossmann, ed., Kluwer Academic Publishers, Boston, MA pp. 195−230 (1996).

Macchietto, S., O. Odele, and O. Omatsone, "Design of Optimal Solvents for Liquid-Liquid Extraction and Gas Absorption Processes," *Trans. of the Inst. of Chem. Eng.*, **68**, 429 (1990).

Maranas, C. D., "Optimal Computer-Aided Molecular Design: A Polymer Design Case Study," *Ind. & Eng. Chemistry Res.*, **35**, 3403 (1996).

Maranas, C. D., "Optimal Molecular Design Under Property Prediction Uncertainty," *AIChE J.*, **43**, 1250 (1997a).

Maranas, C. D., "Optimization Accounting for Property Prediction Uncertainty in Polymer Design," *Comput. Chem. Eng.*, **21**, S1019 (1997b).

Marcoulaki, E. C., and A. C. Kokossis, "Molecular Design Synthesis Using Stochastic Optimization as a Tool for Scoping and Screening," *Comput. Chem. Eng.*, **22**, S11 (1998).

Marcoulaki, E. C., and A. C. Kokossis, "On the Development of Novel Chemicals Using a Systematic Optimisation Approach. Part II. Optimisation Framework," *Chem. Eng. Sci.*, **55**, 2529 (2000a).

Marcoulaki, E. C., and A. C. Kokossis, "On the Development of Novel Chemicals using a Systematic Optimisation Approach. Part II. Solvent Design," *Chem. Eng. Sci.*, **55**, 2547 (2000b).

Nagengast, B. A., "Refrigerants: 160 Years of Change," *ASHRAE J.*, 54 (1995).

Nanda, G., "Design of Efficient Secondary Refrigerants," Master's Thesis, Dept. of Chemical Engineering, University of Illinois, Urbana, IL (2001).

Naser, S. F., and R. L. Fournier, "A System for the Design of an Optimum Liquid-Liquid Extractant Molecule," *Comput. & Chem. Eng.*, **15**, 397 (1991).

Odele, O., and S. Macchietto, "Computer Aided Molecular Design: A Novel Method for Optimal Solvent Selection," *Fluid Phase Equilibria*, **82**, 47 (1993).

Ourique, J. E., and A. S. Telles, "Computer-Aided Molecular Design with Simulated Annealing and Molecular Graphs," *Comput. & Chem. Eng.*, **22**, S615 (1998).

Pistikopoulos, E. N., and S. K. Stefanis, "Optimal Solvent Design for Environmental Impact Minimization," *Comput. Chem. Eng.*, **22**, 717 (1998).

Pretel, E. J., P. A. Lopez, S. B. Bottini, and E. A. Brignole, "Computer-Aided Molecular Design of Solvents for Separation Processes," *AIChE J.*, **40**, 1349 (1994).

Raman, V. S., and C. D. Maranas, "Optimization in Product Design with Properties Correlated with Topological Indices," *Comput. Chem. Eng.*, **22**, 747 (1998).

Reid, R. C., J. M. Prausnitz, and B. E. Poling, *The Properties of Gases and Liquids*, 4th ed., McGraw-Hill, New York (1987).

Riedel, L., "Kritischer koeffizient, dichte des gesattigten dampfes und verdampfungswarme," *Chemie Ingenieur Technik*, **26**, 679 (1954).

Rowlinson, J. S., *Liquids and Liquid Mixtures*, Butterworth, London (1969).

Ryoo, H. S., and N. V. Sahinidis, "Global Optimization of Nonconvex NLPs and MINLPs with Applications in Process Design," *Comput. Chem. Eng.*, **19**, 551 (1995).

Ryoo, H. S., and N. V. Sahinidis, "A Branch-and-Reduce Approach to Global Optimization," *J. of Global Optimization*, **8**, 107 (1996).

Sahinidis, N. V., "BARON: A General Purpose Global Optimization Software Package," *J. of Global Optimization*, **8**, 201 (1996).

Sahinidis, N. V., "Challenges in Informatics: Optimization in Computational Biology, Chemistry, and Medicine," Paper 240b, AIChE Meeting, Indianapolis, IN (2002).

Sahinidis, N. V., and M. Tawarmalani, "Applications of Global Optimization to Process and Molecular Design," *Comput. Chem. Eng.*, **24**, 2157 (2000).

Shectman, J. P., and N. V. Sahinidis, "A Finite Algorithm for Global Minimization of Separable Concave Programs," *J. of Global Optimization*, **12**, 1, (1998).

Siddhaye, S., K. V. Camarda, E. Topp, and M. Southard, "Design of Novel Pharmaceutical Products via Combinatorial Optimization, *Comput. Chem. Eng.*, **24**, 701 (2000).

Sinha, M., L. E. K. Achenie, and G. M. Ostrovsky, "Environmentally Benign Solvent Design by Global Optimization," *Comput. Chem. Eng.*, **23**, 1381 (1999).

Smithart, E., "Change and Progress in the HVAC Industry," *Energy Eng.*, **92**, 64 (1995).

Tawarmalani, M., S. Ahmed, and N. V. Sahinidis, "Product Disaggregation and Relaxations of Mixed-Integer Rational Programs," *Optimization and Eng.*, **3**, 281 (2002).

Tawarmalani, M., and N. V. Sahinidis, "Global Optimization of Mixed Integer Nonlinear Programs: A theoretical and Computational Study," *Mathematical Programming*, Submitted (1999).

Tawarmalani, M., and N. V. Sahinidis, "Semidefnite Relaxations of Fractional Programs via Novel Techniques for Constructing Convex Envelopes of Nonlinear Functions," *J. of Global Optimization*, **20**, 137 (2001).

Tawarmalani, M., and N. V. Sahinidis, "Convex Extensions and Convex Envelopes of l.s.c. Functions," *Mathematical Programming*, **93**, 247 (2002a).

Tawarmalani, M., and N. V. Sahinidis, *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications, Vol. 65, Nonconvex Optimization and Its Applications*, Vol. 65, Kluwer Academic Publishers, Dordrecht (2002b).

Thornton, R. C., "The History and Growth of HVAC&R since the 1930s," *ASHRAE J.*, S68 (1995).

Vaidyanathan, R. and M. El-Halwagi, "Computer-Aided Synthesis of Polymers and Blends with Target Properties," *Ind. Eng. Chemistry Res.*, **35**, 627 (1996).

Vaidyanathan, R., Y. Gowayed, and M. El-Halwagi, "Computer-Aided Design of Fiber Reinforced Polymer Composite Products," *Comput. Chem. Eng.*, **22**, 801 (1998).

Venkatasubramanian, V., K. Chan and J. M. Caruthers, "Computer-Aided Molecular Design Using Genetic Algorithms," *Comput. Chem. Eng.*, **18**, 833 (1994).

Venkatasubramanian, V., K. Chan, and J. M. Caruthers, "Genetic Algorithm Approach for Computer-Aided Molecular Design," *ACS Symposium Series Volume on Applications of Computer-Aided Molecular Design*, C. Reynolds, K. Holloway, and H. Cox, eds., ACS Washington, 396 (1995).

Venkatasubramanian, V., A. Sundaram, K. Chan, and J. M. Caruthers, "Computer-Aided Molecular Design Using Neural Networks and Genetic Algorithms," *Genetic Algorithms in Molecular Design*, J. Devillers, ed., Academic Press, pp. 271−302 (1996).

Watson, K. M., "Thermodynamics of the Liquid State," *Ind. Eng. Chemistry Res.*, **35**, 398 (1943).

Wuebbles, D. J., "The Relative Efficiency of a Number of Halocarbons for Destroying Stratospheric Ozone," Report UCID-18924, Lawrence Livermore National Laboratory, Livermore, CA (1981).

Wuebbles, D. J. and J. Edmonds, *A Primer on Green Gases*, Lewis Publishers, Chelsea, MI (1991).