# Correlation of Boiling Points with Molecular Structure for Chlorofluoroethanes

Terry S. Carlton

Department of Chemistry, Oberlin College, Oberlin, Ohio 44074-1083

Two six-descriptor models, each descriptor derived directly from molecular structure, were fitted to the normal boiling points ($T_b$) of ethane plus the 52 known chlorofluoroethanes by multiple linear regression ($s$ = 2.6 K, $R^2$ = 0.998). Both models predict $T_b$ values of 351 ± 2 and 366 ± 2 K for unknown 1,1-dichloro-2-fluoroethane and 1,1,1-trichloro-2-fluoroethane, respectively. Several descriptors treat specially the atoms that remain exterior after X atoms in $CH_2X$ groups (X = Cl, F) and H atoms in $CHX_2$ groups attach electrostatically to identical groups of adjacent molecules. Electrostatic repulsion between opposing C−X bonds reduces $T_b$ by 20−30 K for half of the compounds.

## INTRODUCTION

Many mathematical models have been developed for fitting and predicting boiling points ($T_b$). Comprehensive models that treat organic compounds of considerable functional diversity have achieved standard deviations ($s$) of 8−12 K in their fits of boiling points.[1−4]

Models for restricted classes of compounds can be more accurate. For example, Hansen and Jurs[5] fitted $T_b$s of 123 $C_2$−$C_{10}$ alkenes and cycloalkenes with $s$ of 1.78 K. Until now, however, the environmentally and industrially important class chlorofluoroalkanes has been harder to fit. The most successful published models for this class are those of Balaban et al.[6] They achieved $s$ = 8.3 K for $C_1$−$C_2$ chlorofluoroalkanes and $s$ = 8.3 K for $C_1$−$C_4$. The present investigation seeks to achieve for the $C_2$ chlorofluoroalkanes an accuracy closer to that of Hansen and Jurs.[5]

Normal $T_b$s are known for all but two of the 54 possible chlorofluroethanes (including ethanes with only chlorine or only fluorine). This wealth of data leaves little to predict, but it permits demanding tests of models. The descriptors in our models will be based solely on molecular structure. Successful descriptors for chlorofluoroethanes may subsequently prove useful in modeling and predicting the $T_b$s of haloalkanes with more carbons or additional halogens, many of which have never been synthesized.

To predict unknown values is one goal of quantitative structure-property relationship (QSPR) modeling. Another is to understand better the structural basis for a property. In the case of chlorofluoroethane $T_b$s, some of these structure−property correlations are fairly obvious. Chlorine atoms contribute markedly to these $T_b$s, presumably because they make a much larger contribution to dispersion forces than do fluorine or hydrogen atoms. The groups $CHX_2$ and $CH_2X$ (X = Cl, F) can engage in dipole−dipole attraction, and these groups likewise markedly increase $T_b$s.

More subtle electrostatic interactions may also be important for these $T_b$s. Woolf[7] pointed out that some of the patterns in chlorofluoroethane $T_b$s could be attributed to intermolecular repulsions between C−F bond dipoles, weaker repulsions between C−H dipoles, and attractions between C−F and C−H. He postulated that repulsions would be

greatest between $CF_3$ groups. Carlton[8] proposed that because highly polar C−X bonds point outward from the entire surface of perhaloalkane molecules, "hedgehog" repulsion makes negative contributions to the $T_b$s of perhaloalkanes. The spines of a rolled-up hedgehog repel every intruder, but molecular hedgehogs repel only molecules of the same surface polarity, typically negative.

Hedgehog repulsion might also be important for certain molecules whose surfaces do not consist entirely of outward-directed, highly polar bonds. For example, molecular complexes could be completely covered by negative spines even if the surfaces of their constituent molecules had both positive and negative sites. Attraction between positive and negative sites on adjacent molecules might both hold the complex together and keep the positive sites from appearing on the surface of the complex. Another situation could involve an uncomplexed molecule with positive and neutral surface sites in addition to negative spines. Those sites could accommodate spines from adjacent molecules, thereby providing relief from hedgehog repulsion. But if the spines per molecule outnumbered the relieving sites, then the excess spines would have to oppose spines on other molecules. The present investigation will test the importance of hedgehog repulsion in both these situations as well as in perhaloethanes.

**The Physical Model.** Underlying our QSPR mathematical models is a physical model that highlights certain relationships between intermolecular forces and structural features. One relationship is strong dispersion attraction by chlorine atoms. Another relationship is hedgehog repulsion between C−X bonds. The model also builds on three dichotomies: bipolar versus monopolar structural groups; primary versus secondary intermolecular interactions; and exterior versus interior X and H atoms. These dichotomies will now be defined and explained.

The model distinguishes among $CH_2X$, $CHX_2$, $CX_3$, and $CH_3$ groups. It assumes that each $CH_2X$ group is associated electrostatically with identical groups on two adjacent molecules. One association involves the hydrogen atoms of the group; the other involves its halogen atom. The model assumes the same for $CHX_2$ groups. We call $CH_2X$ and $CHX_2$ *bipolar* groups because their positive and negative regions are both accessible to other molecules.

CHLOROFLUOROETHANE BOILING POINT−STRUCTURE CORRELATION

*J. Chem. Inf. Comput. Sci., Vol. 38, No. 2, 1998* **159**

We call a $CX_3$ group *monopolar* because only its negative X atoms are accessible to other molecules. The other half of the molecule blocks access to the positive end of the $CX_3$ group. Thus, the electrostatic interaction between two $CX_3$ groups will be repulsive, whereas monopolar $CH_3$ and $CX_3$ will attract each other because of their opposite polarities. The model does not assume that a monopolar group is associated with a specific number of groups from other molecules.

We will find it useful to classify halogen and hydrogen atoms as either *exterior* or *interior*. This classification refers not to the isolated molecule, for which all atoms except carbon are exterior, but to the molecule when every bipolar group is associated with two other molecules. These associations are called *primary* interactions. Once primary interactions are established, the hydrogen atom in any $CHX_2$ group is shielded by a pair of X atoms on another molecule, but all other hydrogen atoms (in $CH_2X$ or $CH_3$ groups) remain exposed. The following considerations show why. The most favorable electrostatic attractions between two $CHX_2$ groups place the H atom of one group in contact with the X atoms of the other. Space-filling Corey−Pauling−Koltun (CPK) models show that when a $CHX_2$ group is associated in this way with two other $CHX_2$ groups, its X atoms still point outward and are almost as exposed as in the isolated molecule. By contrast, the H atom is in the groove between two X atoms on an adjacent molecule; this configuration considerably reduces the accessibility of the H atom to a third molecule. We designate the H atom of a $CHX_2$ group as interior and the X atoms as exterior.

With a $CH_2X$ group, the situation is reversed: the H atoms point outward, and X is interior (though more accessible, because of its larger size, than was H in the $CHX_2$ case). We classify halogen atoms as interior in $CH_2X$ groups but exterior in all other groups, whereas hydrogen atoms are interior in $CHX_2$ groups but exterior in all others.

The following two-stage model of molecular interactions suggests why the exterior/interior classification of peripheral atoms may be important. First, a molecule with one or more bipolar groups undergoes primary associations as previously described. Second, the residual molecular surface undergoes further interactions, called *secondary* interactions. According to the idealized model, this surface consists of exterior atoms only, and interior atoms should be less important for the secondary interactions. Electrostatically, exterior X atoms should repel exterior X atoms of another molecule but attract exterior H atoms. These secondary electrostatic interactions, being primarily dipole−dipole, should be much less important for an interior X atom because its primary association prevents it from being approached along its C−X axis. (An interior X atom is also less exposed for secondary dispersion attractions than is an exterior X atom of the same element, but dispersion attractions, unlike electrostatic interactions of polar bonds, are not very sensitive to the direction of approach.)

Our mathematical models will include several novel descriptors that are based on these dichotomies. If we find that these descriptors contribute significantly to successful modeling of chlorofluoroethane $T_b$s, then the primary/secondary model may also suggest useful descriptors not only for larger haloalkanes but for molecules with other bipolar groups.

## COMPUTATIONAL DETAILS

**Data Base.** Table 1 lists experimental values of the normal $T_b$s as obtained from the literature. Values for 1,1-dichloro-2-fluoroethane and 1,1,1-trichloro-2-fluoroethane are unavailable because these compounds have not been synthesized.

**Mathematical Models.** Two mathematical models, A and B, were used. Each model was a linear combination of six structure-based descriptors plus a constant $y$ intercept. Model A used $tCl$, $eClsq$, bip, $xsX$, D5, and D6 as descriptors. Model B used $tCl$, $eClsq$, $nCHX_2$, $nCH_2X$, D9, and D10. Definitions of the descriptors follow, and their values are listed in Table 1. Descriptor $tCl$ is the total number of chlorine atoms. Descriptor $eClsq$ is the square of $eCl$, the number of exterior chlorine atoms; $eCl$ equals $tCl$ minus the number of $CH_2Cl$ groups. Descriptor $xsX$ is the number of excess exterior X atoms (X = Cl, F); it is defined as the number of exterior X atoms minus the number of exterior H atoms unless this difference is negative, in which case $xsX$ equals zero. Descriptor bip, the number of bipolar groups, equals the sum of descriptors $nCHX_2$ and $nCH_2X$, which are the numbers of $CHX_2$ and $CH_2X$ groups, respectively. Descriptors D5, D6, and D10 are nonzero only for molecules in which all exterior atoms are hydrogens. For these molecules, D5 equals the number of $CH_2X$ groups, D6 equals the number of $CH_3$ groups, and D10 equals 1. Descriptor D9 equals 1 for molecules with a $CH_3$ group plus one or more exterior X atoms. It equals zero for all other molecules.

**Calculation Methods.** Values of descriptor regression coefficients and the $y$ intercept were optimized by multiple linear regression by the analysis tool Regression of the spreadsheet application Microsoft Excel 5.0 on a Power Macintosh. The Excel analysis tool Correlation calculated pairwise correlations among descriptors. Regressions were performed separately for $T_b$s of the following sets of compounds from Table 1: the first 47 (each of which has exterior X atoms), the entire set of 53 (which adds the six compounds that have no exterior X atoms), and the first 50 (with model A only).

## RESULTS

Normal $T_b$s of the four sets of chlorofluoroethanes were fitted by models with $m$ descriptors appropriate to the structural features represented in the sets. For fitting the first 47 compounds, model A did not require descriptors D5 or D6 and model B did not require descriptor D10. Statistical results with model A were

$$n = 47, m = 4, s = 1.82, R^2 = 0.9991, F = 12172$$

$$n = 50, m = 5, s = 1.92, R^2 = 0.9990, F = 8906$$

$$n = 53, m = 6, s = 2.57, R^2 = 0.9984, F = 4777$$

Statistical results with model B were

$$n = 47, m = 5, s = 1.89, R^2 = 0.9991, F = 9034$$

$$n = 53, m = 6, s = 2.61, R^2 = 0.9983, F = 4612$$

**Table 1.** Experimental Normal Boiling Points $T_b$ and Descriptor Values for All 52 Chlorofluoroethanes Plus Ethane

| split formula | | $T_b$ (K) | ref[a] | descriptors | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $t$Cl | $e$Clsq | bip | $xs$X | D5 | D6 | $n$CHX$_2$ | $n$CH$_2$X | D9 | D10 |
| CCl$_3$ | CCl$_3$ | 457.6 | D | 6 | 36 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CCl$_3$ | CCl$_2$F | 411.1 | T | 5 | 25 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CCl$_2$F | CCl$_2$F | 365.88 | T | 4 | 16 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CCl$_3$ | CClF$_2$ | 365.91 | T | 4 | 16 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CCl$_2$F | CClF$_2$ | 320.74 | T | 3 | 9 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CCl$_3$ | CF$_3$ | 318.61 | T | 3 | 9 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CClF$_2$ | CClF$_2$ | 276.58 | T | 2 | 4 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CCl$_2$F | CF$_3$ | 276.59 | T | 2 | 4 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CClF$_2$ | CF$_3$ | 234.08 | T | 1 | 1 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CF$_3$ | CF$_3$ | 195.22 | T | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| CCl$_3$ | CHCl$_2$ | 433.03 | T | 5 | 25 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CCl$_3$ | CHClF | 390.2 | T | 4 | 16 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CCl$_2$F | CHCl$_2$ | 389.8 | T | 4 | 16 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CCl$_3$ | CHF$_2$ | 346 | T | 3 | 9 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CCl$_2$F | CHClF | 345.7 | T | 3 | 9 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CClF$_2$ | CHCl$_2$ | 345.1 | T | 3 | 9 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CCl$_2$F | CHF$_2$ | 303.4 | E | 2 | 4 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CClF$_2$ | CHClF | 302.70 | T | 2 | 4 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CF$_3$ | CHCl$_2$ | 300.81 | T | 2 | 4 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CF$_3$ | CHClF | 261.19 | T | 1 | 1 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CClF$_2$ | CHF$_2$ | 261.41 | T | 1 | 1 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CF$_3$ | CHF$_2$ | 225.06 | T | 0 | 0 | 1 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| CHCl$_2$ | CHCl$_2$ | 418.3 | T | 4 | 16 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 0 |
| CHCl$_2$ | CHClF | 375.6 | F | 3 | 9 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 0 |
| CHCl$_2$ | CHF$_2$ | 333 | T | 2 | 4 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 0 |
| CHClF | CHClF | 332.8 | G | 2 | 4 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 0 |
| CHClF | CHF$_2$ | 290.4 | E | 1 | 1 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 0 |
| CHF$_2$ | CHF$_2$ | 253.10 | T | 0 | 0 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 0 |
| CCl$_3$ | CH$_2$Cl | 403.35 | T | 4 | 9 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| CCl$_2$F | CH$_2$Cl | 361.2 | T | 3 | 4 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| CClF$_2$ | CH$_2$Cl | 319.30 | T | 2 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| CF$_3$ | CH$_2$Cl | 279.25 | T | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| CCl$_2$F | CH$_2$F | 321.6 | H | 2 | 4 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| CClF$_2$ | CH$_2$F | 285 | J | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| CF$_3$ | CH$_2$F | 247.03 | T | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| CHCl$_2$ | CH$_2$Cl | 387.00 | T | 3 | 4 | 2 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| CHClF | CH$_2$Cl | 347.0 | T | 2 | 1 | 2 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| CHF$_2$ | CH$_2$Cl | 308.3 | T | 1 | 0 | 2 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| CHClF | CH$_2$F | 308.2 | K | 1 | 1 | 2 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| CHF$_2$ | CH$_2$F | 276.85 | T | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| CCl$_3$ | CH$_3$ | 347.23 | T | 3 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| CCl$_2$F | CH$_3$ | 305.20 | T | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| CClF$_2$ | CH$_3$ | 264.05 | T | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| CF$_3$ | CH$_3$ | 225.86 | T | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| CHCl$_2$ | CH$_3$ | 330.45 | T | 2 | 4 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| CHClF | CH$_3$ | 289.30 | T | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| CHF$_2$ | CH$_3$ | 249.10 | T | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| CH$_2$Cl | CH$_2$Cl | 356.66 | T | 2 | 0 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 1 |
| CH$_2$Cl | CH$_2$F | 325.91 | T | 1 | 0 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 1 |
| CH$_2$F | CH$_2$F | 283.7 | T | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 1 |
| CH$_2$Cl | CH$_3$ | 285.42 | T | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| CH$_2$F | CH$_3$ | 235.43 | T | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| CH$_3$ | CH$_3$ | 184.57 | L | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 |

[a] D: Antoine equation (from ref T) for vapor pressure of liquid was extrapolated to 1 atm; E: Haszeldine, R. N.; Young, J. C.; Polyfluoroalkyl compounds of silicon. V. Reaction of trichlorosilane with chlorotrifluoroethylene, and halogen abstraction by silyl radicals. *J. Chem. Soc.* **1960**, 4503−4508; F: Hauptschein, M.; Bigelow, L. A. The directing influence of substituents on the chlorination of halogenated ethanes and propanes. *J. Am. Chem. Soc.* **1951**, *73*, 5591−5593; G: Average of $T_b$ values from ref H for meso (59.9 °C) and (±) (59.4 °C) forms; H: Bissell, E. R.; Fields, D. B. Addition of fluorine to halogenated olefins. *J. Org. Chem.* **1964**, *29*, 1591−1593; J: Haszeldine, R. N.; Steele, B. R. Addition of free radicals to unsaturated systems. XIV. The direction of radical addition to trifluoroethylene. *J. Chem. Soc.* **1957**, 2800−2806; K: Rausch, D. A.; Davis, R. A.; Osborne, D. W. The addition of fluorine to halogenated olefins by means of metal fluorides. *J. Org. Chem.* **1963**, *28*, 494−497; L: *TRC Thermodynamic Tables—Hydrocarbons;* Thermodynamics Research Center, The Texas A&M University System: College Station, TX, **1976**, a-1010; T: *TRC Thermodynamic Tables—Non-Hydrocarbons;* Thermodynamics Research Center, The Texas A&M University System: College Station, TX, **1973**, k-7240, k-7241; **1981**, k-7040; **1996**, k-6880, k-6881; **1997**, m-7350−m-7354.

Figure 1 compares with experimental $T_b$s the values calculated by model A for the 53-compound set. Table 2 reports residuals (negative errors) of individual $T_b$s calculated by both models for all sets. Table 3 reports values of regression coefficients. Table 4 reports for the 53-compound set the pairwise correlations among the six descriptors for each model.

The models of Balaban et al.[6] used as descriptors the numbers of C, Cl, and F atoms and the topological indices $W$ (Weiner index) and $J$ index. Using all these except $J$
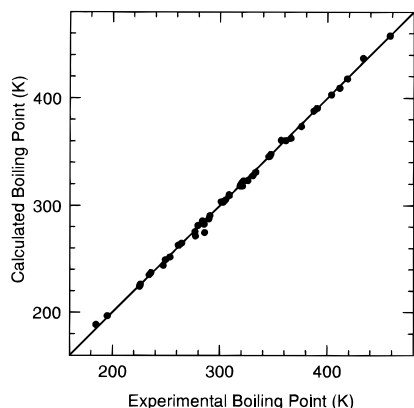
**Figure 1.** Calculated versus experimental $T_b$s for all 52 chlorofluoroethanes plus ethane using the six-descriptor regression model A.

(four descriptors), they fitted $T_b$s for 62 $C_1$−$C_2$ chlorofluoroalkanes with the following results: $n = 62$, $s = 8.3$, $R^2 = 0.984$.

To compare their models with ours, we fitted the 52 chlorofluoroethane $T_b$s ($CH_3CH_3$ omitted) to two four-descriptor models. Results with their descriptors $tCl$, $tF$, $W$, and $J$ were: $n = 52$, $s = 6.10$, $R^2 = 0.9898$, $F = 1142$, whereas a four-descriptor ($tCl$, $exClsq$, $bip$, $xsX$) adaptation of model A gave $n = 52$, $s = 4.06$, $R^2 = 0.9955$, $F = 2598$.

Devotta and Pendyala[9] supplemented the Joback[10,11] group contribution method by adding group contributions for −$CF_3$, >$CF_2$, and >$CF^−$ groups that are not attached to hydrogen, as well as corrections for the following four conditions: perfluorinated; partially fluorinated with or without other halogens; partially halogenated without fluorine; and perhalogenated with or without fluorine. In addition, six of the original Joback descriptors were needed to treat the chlorofluoroethanes. Devotta and Pendyala[9] used Joback's values of the coefficients for five of these six (all except the number of F atoms) as well as Joback's intercept. We calculated the predictions of the $T_b$s by this model for ethane plus all 52 chlorofluoroethanes and found $s$ to be 29.3 K. To give the modified Joback model a better chance, we reoptimized to the 53 $T_b$s all 13 descriptor coefficients and the intercept. The statistical results then were: $n = 53$, $s = 6.15$, $R^2 = 0.9896$, $F = 381$.

Both present models are markedly superior to these previous ones in accuracy and statistical significance for correlations of chlorofluoroethane $T_b$s.

Woolf[7] analyzed patterns of chlorofluoroethane $T_b$s and their increments, particularly as functions of molar mass and among isomers. Although he did not present a mathematical model, he predicted unknown $T_b$s by those patterns.

The four-descriptor version of model A in this investigation predicts the $T_b$s of unknown 1,1-dichloro-2-fluoroethane and 1,1,1-trichloro-2-fluoroethane to be 350.9 and 365.7 K, respectively. The five-descriptor version of model B predicts 350.7 and 365.9 K, respectively, for these compounds. Because both compounds have exterior X atoms, the estimated uncertainty of 2 K is based on the scatter of calculated $T_b$s for the 47-compound set. Table 5 shows that the present predictions for both compounds are at the high end of predictions from the literature.

## DISCUSSION

Models A and B have several similarities but important differences. The statistical results of both models are very similar, and both models correlate chlorofluoroethane $T_b$s better than any previous model. Both models are based on the distinction between exterior and interior atoms. Bipolar groups are important in both models, but only model B distinguishes between $CHX_2$ and $CH_2X$ groups. Only model A has a descriptor that models directly the importance of hedgehog repulsion and a descriptor that models the non-hedgehog attractions of bipolar groups.

**Descriptors.** Descriptors $tCl$ and $eClsq$ are the only ones that depend on whether a halogen atom X is Cl or F. Descriptor $tCl$ is the total number of Cl atoms. It is also, however, the number of F atoms replaced by Cl. [Table 1 makes this evident. For any fixed quartet of values for $bip$, $xsX$, D5, and D6 (or for $nCHX_2$, $nCH_2X$, D9, and D10), each increase of $tCl$ by 1 corresponds to replacement of one F by Cl.] This replacement increases dispersion attraction considerably.

The uncertainty of the coefficient of $tCl$ is ≤2% even though a Cl atom occupies an interior site in 10 of the 53 compounds. This result is remarkable because an interior Cl is less exposed than an exterior Cl. One might have expected the dispersion force of an interior Cl to be less effective in raising the $T_b$ than that of an exterior Cl. A possible explanation is based on the fact that this coefficient is not the total contribution per Cl atom to the $T_b$, merely the increase over the contribution per F atom. Because CPK models show that an interior F atom is less exposed than an interior Cl atom, perhaps a change from an exterior to an interior site reduces the dispersion attraction by both atoms considerably, but by almost the same amount. Another possible explanation is that even an exterior Cl atom uses only a fraction of its exposed surface at any one time due to restrictions on molecular packing, a fraction that might not be much larger than the total exposed surface of an interior Cl.

Descriptor $eClsq$, the square of the number of exterior Cl atoms, approximates the nonlinear increase of $T_b$ with number of Cl atoms. The small coefficient and $t$ statistic of $eClsq$ (Table 3) indicate that it is far less significant than any other descriptor except D5 or D10. Descriptor $eClsq$ is nevertheless important because its values for some compounds are large. When $eClsq$ was omitted from the 47-compound regression calculation with model A, $s$ changed from 1.82 to 3.82 K. When $tClsq$, the square of the *total* number of Cl atoms, was used in place of $eClsq$, $s$ became 2.39 K. Possibly, $eClsq$ reflects a mass-action effect for Cl atoms from adjacent molecules.

Descriptor $xsX$ models hedgehog repulsion. It equals the number of exterior X atoms in excess of exterior H atoms. Its coefficient approximates the negative contribution to the $T_b$ due to hedgehog repulsion by exterior C−X bonds that cannot be accommodated by H atoms. Remarkably, the hedgehog contribution to $T_b$s is the same for a C−Cl bond as a C−F bond. This result is shown by the small uncertainty of the $xsX$ coefficient even though it applies to both Cl and F.

The effect of $xsX$ on 28 of the 53 $T_b$s is distinctly larger than on the others. For these 28, $xsX$ has values of 4, 5, or

**Table 2.** Boiling Point Residuals for Fits of Two Models to 47, 50, and 53 Compounds

| split formula | | $T_b$ (K) | residuals[a] (K) model A 47 | model A 50 | model A 53 | model B 47 | model B 53 |
|---|---|---|---|---|---|---|---|
| $CCl_3$ | $CCl_3$ | 457.6 | −0.90 | −1.00 | −0.57 | −1.03 | −0.78 |
| $CCl_3$ | $CCl_2F$ | 411.1 | 1.45 | 1.42 | 1.54 | 1.19 | 1.31 |
| $CCl_2F$ | $CCl_2F$ | 365.88 | 2.96 | 2.98 | 2.90 | 2.62 | 2.68 |
| $CCl_3$ | $CClF_2$ | 365.91 | 2.99 | 3.01 | 2.93 | 2.65 | 2.71 |
| $CCl_2F$ | $CClF_2$ | 320.74 | 2.43 | 2.47 | 2.31 | 2.05 | 2.13 |
| $CCl_3$ | $CF_3$ | 318.61 | 0.30 | 0.34 | 0.18 | −0.08 | 0.00 |
| $CClF_2$ | $CClF_2$ | 276.58 | 0.78 | 0.81 | 0.68 | 0.39 | 0.56 |
| $CCl_2F$ | $CF_3$ | 276.59 | 0.79 | 0.82 | 0.69 | 0.40 | 0.57 |
| $CClF_2$ | $CF_3$ | 234.08 | −1.34 | −1.34 | −1.33 | −1.68 | −1.36 |
| $CF_3$ | $CF_3$ | 195.22 | −1.92 | −1.99 | −1.72 | −2.19 | −1.63 |
| $CCl_3$ | $CHCl_2$ | 433.03 | −4.31 | −4.34 | −4.19 | −4.15 | −4.12 |
| $CCl_3$ | $CHClF$ | 390.2 | −0.41 | −0.40 | −0.44 | −0.33 | −0.36 |
| $CCl_2F$ | $CHCl_2$ | 389.8 | −0.81 | −0.80 | −0.84 | −0.73 | −0.76 |
| $CCl_3$ | $CHF_2$ | 346 | 0.01 | 0.04 | −0.09 | 0.04 | 0.03 |
| $CCl_2F$ | $CHClF$ | 345.7 | −0.29 | −0.26 | −0.39 | −0.26 | −0.27 |
| $CClF_2$ | $CHCl_2$ | 345.1 | −0.89 | −0.86 | −0.99 | −0.86 | −0.87 |
| $CCl_2F$ | $CHF_2$ | 303.4 | −0.09 | −0.07 | −0.16 | −0.06 | 0.01 |
| $CClF_2$ | $CHClF$ | 302.70 | −0.79 | −0.77 | −0.86 | −0.76 | −0.69 |
| $CF_3$ | $CHCl_2$ | 300.81 | −2.68 | −2.66 | −2.75 | −2.65 | −2.58 |
| $CF_3$ | $CHClF$ | 261.19 | −1.91 | −1.93 | −1.87 | −1.84 | −1.61 |
| $CClF_2$ | $CHF_2$ | 261.41 | −1.69 | −1.71 | −1.65 | −1.62 | −1.39 |
| $CF_3$ | $CHF_2$ | 225.06 | 0.23 | 0.15 | 0.47 | 0.38 | 0.84 |
| $CHCl_2$ | $CHCl_2$ | 418.3 | 0.00 | 0.01 | 0.00 | 0.50 | 0.37 |
| $CHCl_2$ | $CHClF$ | 375.6 | 1.92 | 1.94 | 1.85 | 2.37 | 2.26 |
| $CHCl_2$ | $CHF_2$ | 333 | 1.82 | 1.83 | 1.78 | 2.27 | 2.25 |
| $CHClF$ | $CHClF$ | 332.8 | 1.62 | 1.63 | 1.58 | 2.07 | 2.05 |
| $CHClF$ | $CHF_2$ | 290.4 | −0.39 | −0.41 | −0.32 | 0.10 | 0.23 |
| $CHF_2$ | $CHF_2$ | 253.10 | 0.58 | 0.50 | 0.85 | 1.16 | 1.51 |
| $CCl_3$ | $CH_2Cl$ | 403.35 | 0.40 | 0.49 | 0.08 | 0.13 | −0.46 |
| $CCl_2F$ | $CH_2Cl$ | 361.2 | 0.75 | 0.84 | 0.45 | 0.48 | −0.02 |
| $CClF_2$ | $CH_2Cl$ | 319.30 | −0.76 | −0.71 | −0.95 | −0.99 | −1.34 |
| $CF_3$ | $CH_2Cl$ | 279.25 | −2.54 | −2.55 | −2.53 | −2.69 | −2.81 |
| $CCl_2F$ | $CH_2F$ | 321.6 | −1.64 | −1.62 | −1.69 | −1.80 | −2.04 |
| $CClF_2$ | $CH_2F$ | 285 | 2.15 | 2.13 | 2.21 | 2.02 | 1.94 |
| $CF_3$ | $CH_2F$ | 247.03 | 2.45 | 2.37 | 2.71 | 2.41 | 2.55 |
| $CHCl_2$ | $CH_2Cl$ | 387.00 | −1.14 | −1.06 | −1.40 | −0.99 | −1.59 |
| $CHClF$ | $CH_2Cl$ | 347.0 | −0.75 | −0.71 | −0.91 | −0.56 | −1.00 |
| $CHF_2$ | $CH_2Cl$ | 308.3 | −1.18 | −1.19 | −1.14 | −0.91 | −1.12 |
| $CHClF$ | $CH_2F$ | 308.2 | −2.34 | −2.36 | −2.25 | −2.05 | −2.22 |
| $CHF_2$ | $CH_2F$ | 276.85 | 4.59 | 4.50 | 4.87 | 4.96 | 5.01 |
| $CCl_3$ | $CH_3$ | 347.23 | −0.69 | −0.66 | −0.79 | 1.12 | 0.93 |
| $CCl_2F$ | $CH_3$ | 305.20 | −0.22 | −0.20 | −0.29 | 1.59 | 1.48 |
| $CClF_2$ | $CH_3$ | 264.05 | −0.99 | −1.00 | −0.95 | 0.86 | 0.92 |
| $CF_3$ | $CH_3$ | 225.86 | −0.90 | −0.98 | −0.67 | 1.03 | 1.31 |
| $CHCl_2$ | $CH_3$ | 330.45 | 2.28 | 2.29 | 2.23 | −0.43 | −0.63 |
| $CHClF$ | $CH_3$ | 289.30 | 1.51 | 1.49 | 1.58 | −1.16 | −1.20 |
| $CHF_2$ | $CH_3$ | 249.10 | −0.41 | −0.49 | −0.15 | −3.00 | −2.81 |
| $CH_2Cl$ | $CH_2Cl$ | 356.66 | 9.97[b] | −2.57 | −4.31 | −9.81[b] | − 3.81 |
| $CH_2Cl$ | $CH_2F$ | 325.91 | 16.43[b] | 3.82 | 2.40 | −3.25[b] | 3.02 |
| $CH_2F$ | $CH_2F$ | 283.7 | 11.44[b] | −1.25 | −2.35 | −8.14[b] | −1.61 |
| $CH_2Cl$ | $CH_3$ | 285.42 | − 1.31[b] | −7.61[b] | 10.52 | 3.48[b] | 10.15 |
| $CH_2F$ | $CH_3$ | 235.43 | 14.08[b] | −20.46[b] | −2.01 | −9.19[b] | −2.26 |
| $CH_3$ | $CH_3$ | 184.57 | 42.19[b] | −42.27[b] | −4.26 | − 12.84[b] | −5.50 |

[a] Residual = $T_b$(exp) − $T_b$(calc).  [b] Predicted value.

6, corresponding to contributions of −20 K to −30 K to the $T_b$. For the other 25, $xsX$ is 0 or 1, corresponding to 0 K or −5 K.

The single descriptor bip suffices to model the effect of electrostatic interactions by bipolar groups $CHX_2$ and $CH_2X$. When model A was adapted by using separate descriptors $nCHX_2$ and $nCH_2X$ in place of bip, no $T_b$ residual for the 47-compound correlation changed by >0.01 K, and neither coefficient for the new descriptors differed by >0.015% from that for bip. This insensitivity to the type of bipolar group is remarkable. It is attributable in part to the small variation of dipole moments for these structural groups, as indicated by the following molecular dipole moments[12]: $CH_2ClCH_3$, 2.05 D; $CH_2FCH_3$, 1.94 D; $CHCl_2CH_3$, 2.06 D; $CHClFCH_3$, 2.07 D; and $CHF_2CH_3$, 2.27 D, (1 D = 3.33564 × $10^{-30}$ C/m.) The exterior X atoms of $CHX_2$ cause hedgehog repulsion, whereas the exterior H atoms of $CH_2X$ relieve hedgehog repulsion. Apparently (unless fortuitous cancellations are involved) descriptor $xsX$ models this hedgehog difference so successfully that bip needs to model only electrostatic attractions between bipolar groups.

Hydrogen bonding between C—H and C—F bonds has been proposed as the sole explanation[13] (or as a possible alternative to dipole−dipole attractions[14]) for anomalously

CHLOROFLUOROETHANE BOILING POINT−STRUCTURE CORRELATION

*J. Chem. Inf. Comput. Sci., Vol. 38, No. 2, 1998* **163**

**Table 3.** Coefficients of Descriptors for Boiling Point Correlations

| | model A | | | | | |
|---|---|---|---|---|---|---|
| | $n = 47$ | | | $n = 53$ | | |
| descriptor | coef | SE | $t$ stat[a] | coef | SE | $t$ stat[a] |
| intercept | $226.76 \pm 0.79$ | | 288.0 | $226.53 \pm 1.09$ | | 207.7 |
| $t$Cl | $37.22 \pm 0.43$ | | 86.5 | $37.46 \pm 0.57$ | | 65.9 |
| $e$Clsq | $1.06 \pm 0.09$ | | 12.2 | $1.01 \pm 0.12$ | | 8.7 |
| bip | $22.75 \pm 0.39$ | | 58.0 | $22.73 \pm 0.55$ | | 41.1 |
| $xs$X | $-4.94 \pm 0.13$ | | $-38.9$ | $-4.93 \pm 0.18$ | | $-27.6$ |
| D5 | | | | $7.04 \pm 0.80$ | | 8.8 |
| D6 | | | | $-18.85 \pm 1.20$ | | $-15.7$ |

| | model B | | | | | |
|---|---|---|---|---|---|---|
| | $n = 47$ | | | $n = 53$ | | |
| descriptor | coef | SE | $t$ stat[a] | coef | SE | $t$ stat[a] |
| intercept | $197.41 \pm 0.81$ | | 242.9 | $196.85 \pm 1.03$ | | 191.9 |
| $t$Cl | $37.32 \pm 0.46$ | | 81.4 | $37.58 \pm 0.59$ | | 63.6 |
| $e$Clsq | $1.04 \pm 0.09$ | | 11.3 | $1.00 \pm 0.12$ | | 8.3 |
| $n$CHX$_2$ | $27.27 \pm 0.44$ | | 62.3 | $27.37 \pm 0.59$ | | 46.2 |
| $n$CH$_2$X | $47.22 \pm 0.77$ | | 61.2 | $47.62 \pm 0.88$ | | 54.4 |
| D9 | $27.43 \pm 0.88$ | | 31.3 | $27.70 \pm 1.18$ | | 23.5 |
| D10 | | | | $-6.79 \pm 1.49$ | | $-4.6$ |

[a] The $t$ statistic equals (regression coefficient)/(standard error).

**Table 4.** Correlation Matrices of the Descriptors ($R$ Values)

| | model A | | | | | |
|---|---|---|---|---|---|---|
| | $t$Cl | $ex$Clsq | bip | $xs$X | D5 | D6 |
| $t$Cl | 1.000 | | | | | |
| $ex$Clsq | 0.901 | 1.000 | | | | |
| bip | $-0.205$ | $-0.299$ | 1.000 | | | |
| $xs$X | 0.462 | 0.520 | $-0.338$ | 1.000 | | |
| D5 | $-0.227$ | $-0.227$ | 0.319 | $-0.349$ | 1.000 | |
| D6 | $-0.266$ | $-0.170$ | $-0.150$ | $-0.262$ | 0.164 | 1.000 |

| | model B | | | | | |
|---|---|---|---|---|---|---|
| | $t$Cl | $ex$Clsq | $n$CHX$_2$ | $n$CH$_2$X | D9 | D10 |
| $t$Cl | 1.000 | | | | | |
| $ex$Clsq | 0.901 | 1.000 | | | | |
| $n$CHX$_2$ | 0.011 | $-0.006$ | 1.000 | | | |
| $n$CH$_2$X | $-0.269$ | $-0.367$ | $-0.332$ | 1.000 | | |
| D9 | $-0.181$ | $-0.148$ | $-0.100$ | $-0.249$ | 1.000 | |
| D10 | $-0.313$ | $-0.264$ | $-0.316$ | 0.579 | $-0.139$ | 1.000 |

**Table 5.** Predicted Boiling Points of Unknown Chlorofluoroethanes

| | | $T_b$(K) | |
|---|---|---|---|
| source | descriptors | CHCl$_2$CH$_2$F | CCl$_3$CH$_2$F |
| present models A and B | 4 | $351 \pm 2$ | $366 \pm 2$ |
| Balaban et al.[a] | 4 | 350 | 358, 360 |
| Balaban et al.[b] | 4 | 346 | 364 |
| Modified Joback[c] | 13 | 348 | 367 |
| Woolf[d] | N/A | 348 | 362 |

[a] Reference 6, their models. [b] Coefficents of $t$Cl, $t$F, $W$, $J$ reoptimized to 52 chlorofluoroethane boiling points. [c] Reference 9. Coefficients reoptimized to boiling points of ethane plus 52 chlorofluoroethanes. [d] Reference 7.

high $T_b$s of fluoroethanes that contain CH$_2$F or CHF$_2$ groups. However, hydrogen bonding is unlikely to account for this behavior, which is explained more plausibly by electrostatic attraction of bipolar groups, as modeled quantitatively by bip, $n$CHX$_2$, and $n$CH$_2$X. Chlorofluoroethanes with CH$_2$X or CHX$_2$ groups are very weak hydrogen-bond donors,[15] and fluoroalkanes are very weak hydrogen-bond acceptors.[16]

Model B does not have a descriptor like $xs$X to model hedgehog repulsion, which causes the coefficient of descrip-

tor $n$CHX$_2$ to be smaller than that of $n$CH$_2$X. A CHX$_2$ group has two more exterior X atoms than CH$_2$X and two fewer exterior H atoms to relieve hedgehog repulsion. This interpretation agrees with the fact (cf., Table 3) that:

$$\text{coef}(n\text{CHX}_2) = \text{coef}(n\text{CH}_2\text{X}) + 4 \times \text{coef}(xs\text{X})$$

to well within the uncertainties of these coefficients.

Descriptor D9 is similarly related to relief of hedgehog repulsion. The reference compound for model B is CF$_3$CF$_3$ in the sense that the value of every descriptor is zero. (Model A has CF$_3$CH$_3$ as its reference compound.) All C−F bonds in CF$_3$CF$_3$ contribute to hedgehog repulsion, but all C−F bonds in CF$_3$CH$_3$ are compensated by C−H bonds. This compensation stabilizes liquid CF$_3$CH$_3$ relative to CF$_3$CF$_3$. Descriptor D9 models this stabilization's effect on the $T_b$. Its coefficient is ∼2.0 K smaller, however, than six times $-$coef($xs$X), possibly because D9 is used for compounds CHX$_2$CH$_3$ in addition to CX$_3$CH$_3$.

The negative coefficient of D10 corrects for the fact that $n$CH$_2$X overcompensates for hedgehog relief in compounds that have no exterior X atoms. The fact that D10 is also successful for CH$_3$CH$_3$ may be fortuitous.

Descriptors D5 and D6 are essentially indicator variables. Only four of the compounds have nonzero values for D5, and only three of them for D6. Descriptors D5 and D6 specify the numbers of CH$_2$X and CH$_3$ groups, respectively, but only in molecules in which all exterior atoms are H. In all other molecules that contain CH$_2$X or CH$_3$, these groups provide relief from hedgehog repulsion, but there cannot be hedgehog relief in the absence of exterior C−X bonds. Hedgehog relief makes a positive but unknown contribution to the $T_b$ of CF$_3$CH$_3$. This contribution is incorporated into the value of the $y$ intercept, which is also the estimate by the model of the $T_b$ of CF$_3$CH$_3$ because the value of every descriptor is zero for this molecule. The $y$ intercept is the same for every compound, so a negative adjustment must be added to the $T_b$ model for CH$_3$CH$_3$. This adjustment for CH$_3$CH$_3$ is approximated by twice the D6 coefficient, which is observed to be negative.

An analysis of the changes in going from CF$_3$CH$_2$F to CH$_2$FCH$_2$F is slightly more complicated but also involves the absence of a need for hedgehog relief. We therefore would expect the coefficient of D5 to be negative like that of D6. Surprisingly, it is positive. In addition, the $T_b$s of CH$_2$ClCH$_3$ and CH$_2$FCH$_3$ present a special challenge to any model. A five-descriptor model A (D6 not needed) that correlates $T_b$s of the first 50 compounds predicts $T_b$s that are high by 7.6 K for CH$_2$ClCH$_3$ and by 20.5 K for CH$_2$-FCH$_3$ (Table 2). No single additional descriptor could correlate both $T_b$s satisfactorily. With D6 as the additional descriptor, the correlation of all 53 $T_b$s has residuals of 10.5, $-2.0$, and $-4.3$ K for CH$_2$ClCH$_3$, CH$_2$FCH$_3$, and CH$_3$CH$_3$, respectively. The 10.5 K residual is largely responsible for increasing $s$ to 2.57 K (from 1.92 K for the five-descriptor, 50-compound correlation). Although D6 correlates $T_b$s acceptably for both CH$_2$FCH$_3$ and CH$_3$CH$_3$, this result may be merely fortuitous in view of the unexplained sign of the D5 coefficient.

Descriptors D5 and D6 are useful but not essential for the chlorofluoroethanes. This fact was demonstrated by the four-descriptor model A results for $n = 52$ (cf., the *Results*

section). The residual for ethane (which was omitted from the set) was −41 K. Residuals for other compounds with nonzero values of D5 or D6 ranged from −14 to +15 K. Magnitudes of all other residuals were 4.6 K or less, and *s* for the 52 compounds was 4.06 K.

The $T_b$s of haloalkanes without exterior X atoms may prove easier to relate to those of alkanes than to those of haloalkanes that have exterior X atoms. Testing of this speculation would require going beyond the haloethanes, which include few members without exterior X atoms.

Models A and B are both remarkably economical in their choices of descriptors. They do not use descriptors for the number of H atoms, of F atoms, or of X atoms. They ignore the difference between the dispersion energies of F and H atoms. The distinction between exterior and interior atoms is crucially important, yet no descriptor specifies the number of exterior X atoms or exterior H atoms. However, the total number of exterior atoms does equal six minus bip (and hence six minus $n$CHX$_2$ minus $n$CH$_2$X).

Model A does not discriminate between CHX$_2$ and CH$_2$X groups, as has already been discussed. It ignores CX$_3$ groups. It acknowledges CH$_3$ groups only in the three CH$_3$-containing compounds in which all exterior atoms are H.

Models A and B both ignore rearrangements of exterior halogen atoms. Exchange of an exterior Cl on one carbon with an exterior F on the other carbon changes the experimental $T_b$ by at most 2 K, and by <0.6 K for eight of the 10 isomeric pairs corresponding to this exhange.

Because of descriptor $e$Clsq, neither model ignores exchange of an interior Cl with an exterior F. The experimental $T_b$ of CClF$_2$CH$_2$F is 6 K higher than that of CF$_3$CH$_2$Cl. The $T_b$ changes by −0.1 K and 2.3 K for the other two exchanges of this sort for which data are available. Perhaps the $T_b$ of CClF$_2$CH$_2$F should be remeasured.

Coefficients of $t$Cl, bip, $xs$X, $n$CHX$_2$, $n$CH$_2$X, and D9 were particularly precise, as shown by their large $t$ statistics (Table 3). In addition, those coefficients remained virtually unchanged when the 47-compound correlations were augmented with the six compounds that lack exterior X atoms and hence require additional descriptors. These facts suggest that those six descriptors are individually significant, not just parameters in successful models.

Which of these six descriptors could contribute usefully to models for related molecules? Descriptors $t$Cl, $n$CHX$_2$, and $n$CH$_2$X appear promising because they refer to structural groups that remain well defined in larger haloalkanes. The other three descriptors may also prove useful, but only experimentation with new models will tell.

**Summary Conclusions.** Two new models, each with six descriptors, reproduce chlorofluoroethane $T_b$s with unprecedented accuracy. Discriminations between primary and secondary interactions and between exterior and interior atoms (in the special senses of this paper) are important to the definitions and success of both models. Two descriptors of model A and one descriptor of model B apply only to the six compounds for which every exterior atom is H.

Together, the descriptors $t$Cl and $e$Clsq model the slightly nonlinear increase of $T_b$ with number of Cl atoms. The contribution per Cl to $T_b$ ranges from 38 K for molecules with one Cl to 43 K for C$_2$Cl$_6$. The descriptor bip models primary electrostatic attractions of bipolar groups. These electrostatic attractions contribute 23 K per group (either CH$_2$X or CHX$_2$) to $T_b$.

Electrostatic repulsion between exterior C—X bonds makes a negative contribution to $T_b$ to the extent that these bonds are not compensated by exterior C—H bonds. This hedgehog contribution is the same for a C—Cl bond as for a C—F bond and equals the $xs$X coefficient. The hedgehog contribution reduces $T_b$ for CX$_3$CX$_3$ by 30 K.

### REFERENCES AND NOTES

(1) Egolf, L. M.; Wessel, M. D.; Jurs, P. C. Prediction of boiling points and critical temperatures of industrially important organic compounds from molecular structure. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 947−956.
(2) Katritzky, A. R.; Mu, L.; Lobanov, V. S. Correlation of boiling points with molecular structure. 1. A training set of 298 diverse organics and a test set of 9 simple inorganics. *J. Phys. Chem.* **1996**, *100*, 10400−10407.
(3) Wang, S.; Milne, G. W. A.; Klopman, G. Graph theory and group contributions in the estimation of boiling points *J. Chem. Inf Comput. Sci.* **1994**, *34*, 1242−1250.
(4) Wessel, M. D.; Jurs, P. C. Prediction of normal boiling points for a diverse set of industrially important organic compounds from molecular structure. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 841−850.
(5) Hansen, P. J.; Jurs, P. C. Prediction of olefin boiling points from molecular structure. *Anal. Chem.* **1987**, *59*, 2322−2327.
(6) Balaban, A. T.; Basak, S. C.; Colburn, T.; Grunwald, G. D. Correlation between structure and normal boiling points of haloalkanes C$_1$−C$_4$ using neural networks. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1118−1121.
(7) Woolf, A. A. Boiling point relations in the halogenated ethane series. *J. Fluorine Chem.* **1990**, *50,* 89−99.
(8) Carlton, T. S.; Winkle, R. J. Calculations of electrostatic intermolecular repulsion in tetrafluoromethane: evidence for a 'hedgehog' contribution to perfluoroalkane volatility. *J. Fluorine Chem.* **1993**, *65*, 1−6.
(9) Devotta, S.; Pendyala, V. R. Modified Joback group contribution method for normal boiling point of aliphatic halogenated compounds. *Ind. Eng. Chem. Res.* **1992,** *31*, 2042−2046.
(10) Joback, K. G.; Reid, R. C. Estimation of pure-component properties from group contributions. *Chem. Eng. Comm.* **1987,** *57*, 233−243.
(11) Reid, R. C.; Prausnitz, J. M.; Poling, B. E. *The Properties of Gases and Liquids,* 4th ed.; McGraw-Hill: New York, 1987.
(12) Lide, D. R. Dipole moments of molecules in the gas phase. In *CRC Handbook of Chemistry and Physics,* 76th ed.; Lide, D. R., Ed.; CRC: Boca Raton, FL, 1995; pp 9.42−9.50.
(13) Hudlicky, M. *Chemistry of Organic Fluorine Compounds,* 2nd ed.; Halsted: New York, 1976; pp 533−538, 546.
(14) Smart, B. E. Physical and physicochemical properties. In *Chemistry of Organic Fluorine Compounds II: A Critical Review;* Hudlicky, M.; Pavlath, A. E., Eds.; American Chemical Society: Washington, D.C., 1995; pp 979−1010.
(15) Abraham, M. H.; Grellier, P. L.; Prior, D. V.; Duce, P. P.; Morris, J. J.; Taylor, P. J. Hydrogen bonding. Part 7. A scale of solute hydrogen-bond acidity based on log K values for complexation in tetrachloromethane. *J. Chem. Soc., Perkin Trans. 2,* **1989,** 699−711.
(16) Abraham, M. H.; Duce, P. P.; Prior, D. V.; Barratt, D. G.; Morris, J. J.; Taylor, P. J. Hydrogen bonding. Part 9. Solute proton donor and proton acceptor scales for use in drug design. *J. Chem. Soc., Perkin Trans. 2* **1989**, 1355−1375, Table 4.