http://dx.doi.org/10.5012/bkcs.2012.33.3.855

MS-HEMs: An On-line Management System for High-Energy Molecules at ADD and BMDRC in Korea†

Sung Kwang Lee,[*] Soo Gyeong Cho,[‡,*] Jae Sung Park,[§] Kwang Yeon Kim,[§] and Kyoung Tae No[§,#]

Department of Chemistry, Hannam University, Daejeon 305-811, Korea. [*]E-mail: leesk@hnu.kr
[‡]Agency for Defense Development (ADD), P.O. Box 35-42, Daejeon 305-600, Korea. [*]E-mail: soocho@add.re.kr
[§]Bioinformatics & Molecular Design Research Center (BMDRC), Seoul 120-749, Korea
[#]Department of Biotechnology, Yonsei Univesity, Seoul 120-749, Korea
Received October 27, 2011, Accepted January 2, 2012

A pioneering version of an on-line management system for high-energy molecules (MS-HEMs) was developed by the ADD and BMDRC in Korea. The current system can manage the physicochemical and explosive properties of virtual and existing HEMs. The on-line MS-HEMs consist of three main routines: management, calculation, and search. The management routine contains a user-friendly interface to store and manage molecular structures and other properties of the new HEMs. The calculation routine automatically calculates a number of compositional and topological molecular descriptors when a new HEM is stored in the MS-HEMs. Physical properties, such as the heat of formation and density, can also be calculated using group additivity methods. In addition, the calculation routine for the impact sensitivity can be used to obtain the safety nature of new HEMs. The impact sensitivity was estimated in a knowledge-based manner using in-house neural network code. The search routine enables general users to find an exact HEM and its properties by sketching a 2D chemical structure, or to retrieve HEMs and their properties by giving a range of properties. These on-line MS-HEMs are expected be powerful tool for deriving novel promising HEMs.

Key Words : Management system, High-energy molecules (HEMs), Chemical search, Molecular descriptor, Impact sensitivity

Introduction

Recently there have been significant advancements in developing novel high-energy molecules (HEMs).[1-8] A range of new HEMs and relevant additives have been synthesized and applied to civil and military research. HNIW (2,4,6,8,10,12-hexanitrohexaazaisowurtzitane) is a good example of a highly powerful HEM.[9] The high power of HNIW can be attributed to the intriguing molecular structure and high nitrogen content. On the other hand, the highly sensitive nature of HNIW requires great caution when handling explosives. TATB (1,3,5-triamino-2,4,6-trinitrobenzene) and NTO (3-nitro-1,2,4-triazole-5-one), which were also introduced recently to the society of HEMs, show great insensitivity.[10-12] Nevertheless, formulation engineers dealing with these explosive molecules feel that TATB and NTO should have higher explosive performance.

Researchers who design and synthesize new HEMs have pushed themselves to derive more powerful, yet safe HEMs. However, it is difficult to find promising HEM candidates because more powerful HEMs are generally sensitive, whereas insensitive HEMs normally have less powerful. Owing to this inversely proportional tendency between performance and insensitivity, the search a promising novel HEM requires good strategies, not just a trial and error approach. It is important to understand the various molecular aspects and pinpoint the molecular aspects to enhance either the explosive performance or safety. The molecular design in drug discovery is a good example. Although the success ratio is extremely low in new drug design, the success guarantees high profit. For example, recent statistics shows that only one out of five thousand drugs eventually reach the market.[13] It is extremely important to select good candidate molecules in the early stages of development. One good way of identifying promising candidates is to have a good DB, which archives all previous information including the failures and possible virtual candidates. One of known HEMs DB is ICT thermochemical database[14] which contains more than 12,000 compounds. ICT database provides crystal density, melting point, boiling point, enthalpy of formation, and combustion energy. Although ICT database provides valuable information to HEM designers and synthetic chemists, it only archives experimental values of known molecules. Contrary to other HEM DBs including ICT thermochemical DB, our MS-HEMs were recently designed to collect and store virtual HEMs as well as known HEMs to derive a new successful HEM.

Architecture of the Management System

On-line MS-HEMs at ADD and BMDRC consists of three sets of architecture: web interface, calculation module and database. The web interface is written in PHP script langu-

†This paper is to commemorate Professor Kook Joe Shin's honourable retirement.

Figure 1. Main screen of the MS-HEMs.

age, and delivered using the Apache Web server, which is designed to handle the calculation module and database to provide HEMs information to the user or to store input data from the user. The calculation module is programmed using C language to calculate the various molecular descriptors. Quantitative structure-property relationships (QSPRs) need to be performed to derive the various explosive properties. In particular, the function to predict the impact sensitivity using previously stored in-house artificial neural network (ANN) weights, was added to the calculation module. The database implemented using Oracle 9i DBMS stores not only the 3D chemical structure using X-ray and computational methods but also the corresponding experimental and computational properties. Figure 1 shows the main screen of the MS-HEMs.

Service and Capabilities of MS-HEMs

Figure 2 shows a schematic diagram of the MS-HEMs. The characteristics of MS-HEMs assist the users in deriving novel HEMs by dealing with a great deal of virtual HEMs in the design stage.

Input Routine. All input data was stored in the database of the MS-HEMs through an input routine by the user. Table 1 summarizes the data fields in the database. All the HEMs records can be freely accessible to the public unless the user wishes to restrict the information of a specific HEM. In this case, the HEM and relevant records will be classified as confidential, and will be opened only to the restricted users and administrator. The user can add various names, such as the IUPAC name, common name and acronym, to identify the HEMs and upload the HEM in the form of a 3D MDL mol file[15] using the X3Dmole program.[16] The method type should be chosen according to the 3D structure generation including x-ray crystallography, two semi-empirical methods (AM1, PM3), two *ab initio* methods (HF/3-21G and HF/6-31G*) and two density functional methods (B3LYP/6-31G* and BP/6-31G**). If other generation methods have been performed at other levels of theories, one should choose 'etc' to write the method type.

Molecular orbital methods (semi-empirical and *ab initio*) have been used to examine the performance and sensitivity of the HEMs from the electronic structure, electronic descriptors, such as molecular electronegativity, ionization potential, HOMO-LUMO energy gap, excess charges of specific atoms and electrostatic potentials of specific bonds. These descriptors are often used as input variables to correlate the chemical structures with explosive and chemical properties of HEMs.

The physical properties, such as heat of formation, heat of combustion, heat of sublimation and crystal density can be obtained from both experiments and computation. They can provide an excellent estimation of the explosive performance. The boiling and melting points are obtained mainly

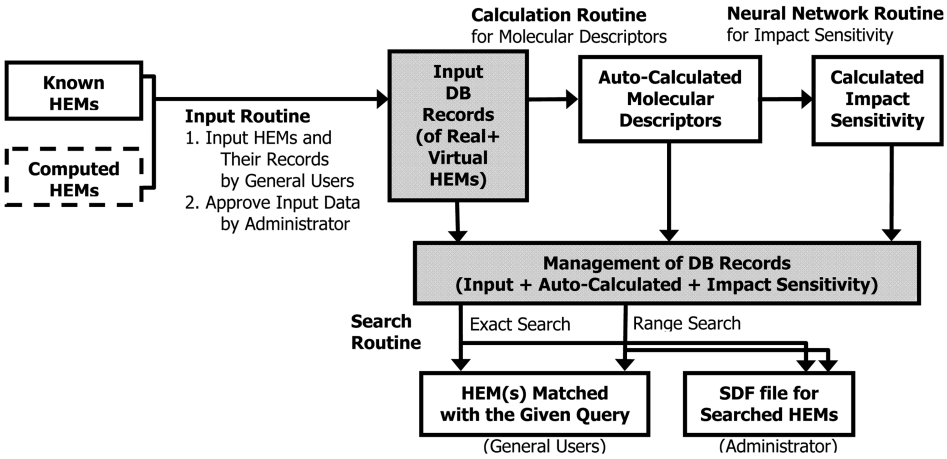

Figure 2. Structure of the MS-HEMs.

from experiments, and are important factors in processing the HEM formulations to determine which process can be used to develop new explosive formulations with new HEMs.

The explosive properties, such as C-J pressure and detonation velocity, are listed in the input screen, in addition to the impact, friction and electrostatic sensitivities. The DB can manage both the experimentally measured and computational estimated values of the explosive performance and sensitivity.

In addition, the assignment of molecule types should be inputted in a database system according to the presence of a ring, heteroatom and aromaticity. This classification will be helpful as a good categorizer when one wishes to design a new HEM of a specific molecular type and search query molecules within that region. Since some high molecular weight HEMs can have several molecular types at a same time, the users may check more than one molecular type within a category of cyclic or acyclic. This part plans to automatically assign molecular types in the calculation routine.

Table 1. Input Fields of On-line MS-HEMs

• Security Level: confidential/normal	• Phase (at 25 °C) · Solid · Solid-Liquid · Liquid · Liquid-Gas · Gas
• Record Type:experimental/ computational	
• Common name:	
• Acronym:	
• IUPAC name:	
• Molecule type: · Cyclic - benzenes - aromatic heterocycles - aliphatic heterocycles - aliphatic hydrocarbons - salt with cycles · Acyclic - acyclic hydrocarbons - acyclic chain with heteroatoms - salt without cycles	• Aromaticity: Y/N • Synthesized?: Y/N • Molecular electronegativity • Ionization potential (eV) • HOMO-LUMO energy gap (eV) • Heat of formation (kcal/mol) • Heat of combustion (kcal/mol) • Heat of sublimation (kcal/mol) • Density(g/cc)
• Structure creation method · X-ray · AM1 · PM3 · HF/6-31G* · HF/3-21G · B3LYP/6-31G* · BP86/6-31G**-1 · BP86/6-31G**-2 · etc	• Boiling point (°C) • Melting point (°C) • C-J Pressure (kbar) • Detonation velocity (m/sec) • Partition coefficient (logP) • Molecular reflectivity index • Impact sensitivity (cm) • Friction sensitivity • Electrostatic sensitivity

Calculation Routine. Once the 3D chemical structure of a new HEM is uploaded into the DB, MS-HEMs run a calculation routine to automatically calculate the various compositional and topological molecular descriptors. Table 2 lists all the molecular descriptors calculated in this routine including the molecular weight, oxygen balance, and number of specific atoms as compositional descriptors, as well as the number of specific bonds, groups, rings, and rotatable bonds as topological descriptors. In addition, the impact sensitivity, which is one of the most important descriptors in explaining the safety features of new HEMs, is computed from the ANN based on the QSPR model. The QSPR model of the impact sensitivity will be explained in Section 3.3. These descriptors are expected to be utilized to future studies predicting the explosive performance and safety nature of new HEMs. To predict the explosive property and safety nature more accurately, it is important to expand the automatic calculation routine to calculate more topological and 3-D geometric descriptors.

Besides the compositional and topological descriptors, the heat of formation and crystalline density were also calculated in this routine. These descriptors are very important for estimating the detonation velocity and pressure. The Benson[17-19] and Salmon method[20,21] were implemented to calculate the heat of formation. Benson method is based on the 2nd order group additivity method, which considers the

Table 2. Molecular Descriptors Calculated Automatically by the MS-HEMs

Molecular weight
Oxygen valence OB100 = (2*C - H - 2*N - 2*(CO_2))/MW
Number of specific groups NO_2-C(sp^2), NO_2-C(sp^3), NO_2-N, NO_2-O, -C(NO_2)$_3$, -CO_2, -NH_2, -OH, CH in aromatic ring of nitroaromatics,
Number of aromatic rings Number of aliphatic rings Number of rotatable bonds
Impact sensitivity (logH_{50}, cm) Heat of formation (kcal/mol, gas & solid states) Crystal density (g/cc)

atomic bond and group contribution to the heat of formation in the gas phase. The Salmon method is similar to the Benson method but can also estimate the heat of formation in the solid state according to the number of nitrogen-containing molecules. The Ammon method[22-24] was implemented to calculate the crystal density. This method employs group additivity methods where the molar volume is obtained by summing up the volumes of the appropriate atoms or functional groups.

ANN Routine for Impact Sensitivity. One of the important features in the calculation routine is to estimate the impact sensitivity of a new HEM from the chemical structure. It is difficult to obtain the reliable impact sensitivity value because the values can vary considerably according to the conditions and equipment used. Therefore, although the user adds the experimental value to the MS-HEMs, the impact sensitivity value estimated by this code may provide a second chance to contemplate the quality of the experimental value. In most cases, a reasonable estimation of the impact sensitivities of HEMs is useful for screening the potential HEM because all the virtual and some of the existing HEMs are not clarified by their safety nature, which is measured by the impact sensitivity values. Moreover, it is useful to develop a good QSPR model that employs calculated molecular descriptors from the 2D chemical structure to facilitate a quick and easy assessment of the impact sensitivity. During the uploading stage of the MS-HEMs DB, the ANN model was performed to obtain a reasonable estimation of the impact sensitivity using a range of HEMs in this module.

In a previous study,[25] ANN architecture was optimized using compositional and topological descriptors. The database published by Storm, Stine, and Kramer,[26,27] who archived impact sensitivities of 234 HEMs, was used to train the ANN model. The 17 (combining compositional and topological parameters)-2-1 architecture provided the best result among those tested in the present study by enhancing the correlation coefficient (r^2) and reducing the standard error of the prediction (SEP) in the test set.

This study checked whether this architecture depended upon the selection of different molecules in the validation and test sets. These new predictions by switching molecules

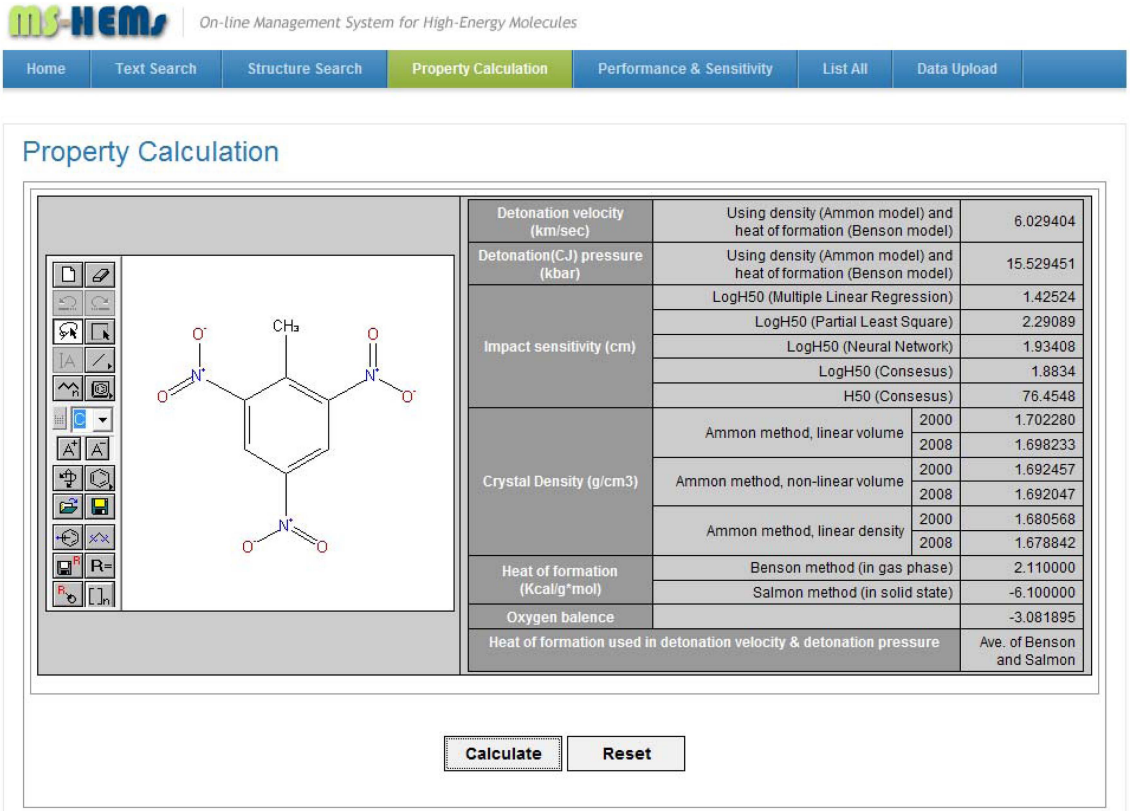

Figure 3. The screen layout for predicting the properties of TNT.

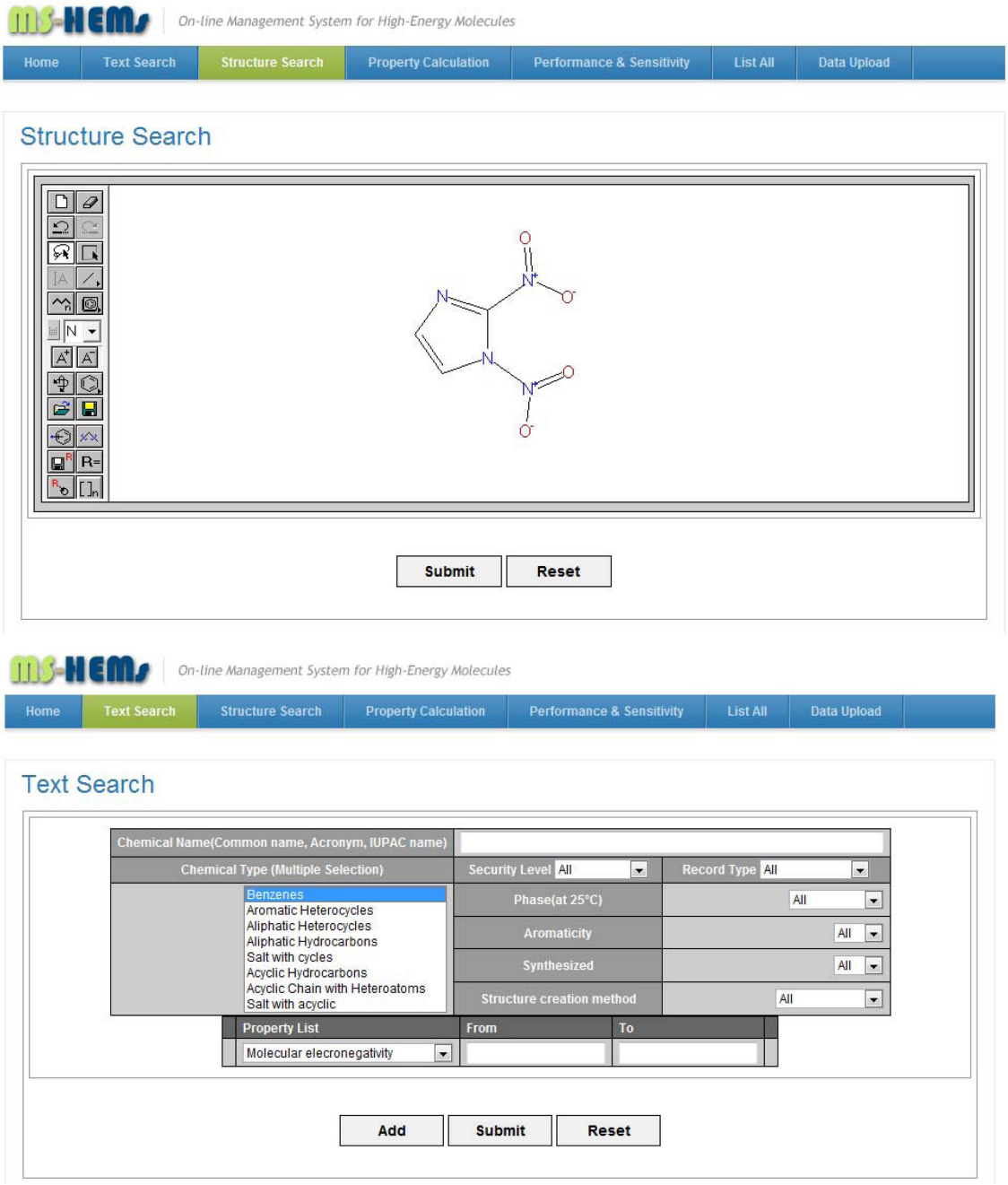

Figure 4. The screen layouts for the structure search (top) and text search (below).

in the validation and test sets provided SEP and r^2 values, which are similar to the original prediction. This ANN architecture has been implemented in the in-house ANN structure inside the MS-HEMs. Simple compositional and topological molecular descriptors, which were calculated in the calculation routine implemented in the MS-HEMs, were employed as input neurons. Figure 3 gives a screenshot of the predicting properties of TNT from the calculation and ANN routine.

Search Routine. Two user interfaces, e.g. a structure search and text search, have been implemented in MS-HEMs. Both interfaces provide access to certain HEMs along with their explosive and chemical properties. Figure 4 shows the screen layout of the search. The structure search enables the users to find the HEM and its properties by inputting the 2D molecular structure. This exact search mechanism from 2D molecular structure has been implemented by following Ullman's isomorphism method.[28] The DBdrawX program,[29] which was developed by BMDRI and interfaced to the ActiveX program, has been used to sketch the input 2D chemical structure. A MDL mol file, which is generated using other commercial chemical drawing software, such as ChemDraw and Accelrys Draw, can be used to input the 2D chemical structure. We are currently implementing a partial structure search method into the on-line DB to retrieve the HEMs, which have an input molecular

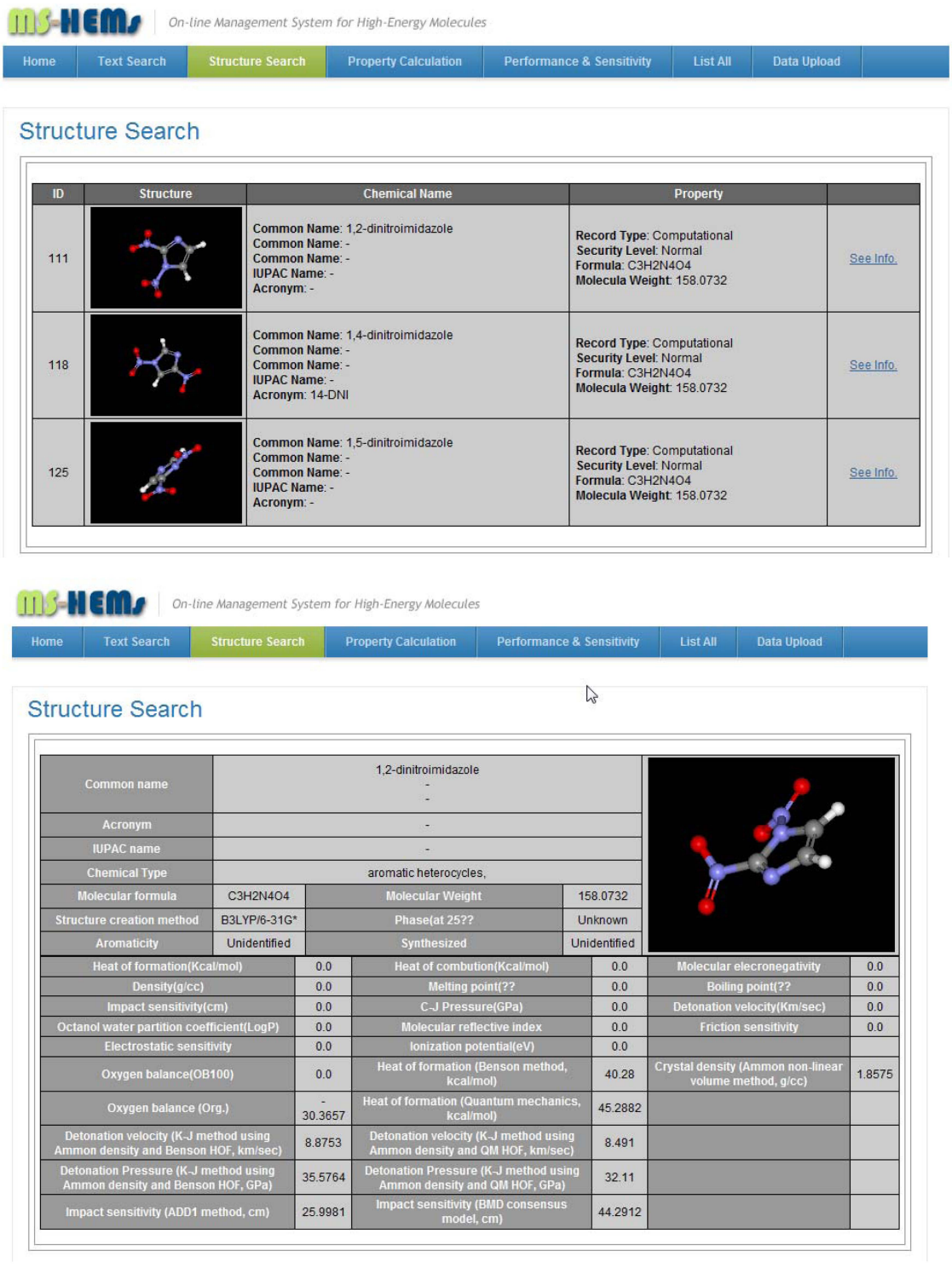

Figure 5. The list of HEMs retrieved (top) and chemical information of the specific HEM retrieved (below) after performing a search.

fragment.

Text search mode is also shown in Figure 4. In this mode, the users can retrieve the target HEMs by inputting the IUPAC name, common name, structure creation method, molecular type, and phase. In addition, the users can identify HEMs with eight different properties. These results can be presented by a SD file of the selected HEMs and their properties, when approved by the administrator.

Concluding Remarks

Chemical database (DB) is an extremely important resource in developing novel materials and provides good reservoir of candidate molecules. Our continuing interest in developing novel HEMs prompted us to develop an on-line MS-HEMs. The program and DB have been installed linux-based PCs at the BMDRC in Korea, and is currently avai-

lable on the web (http://add.bmdrc.org). The on-line MS-HEMs can manage virtual HEMs designed by theoretical scientists as well as existing HEMs synthesized previously. A combined set of virtual and real HEMs should work as an important knowledge-based resource in deriving new HEMs that have excellent explosive performance and enhanced safety nature. Owing to the virtual molecules dealt with this DB, a large proportion of input records consist of molecular descriptors that can be obtained from quantum mechanical calculations. The input screen has a user-friendly menu, allowing the users to input records easily. The DB has a calculation routine, where a range of compositional and topological descriptors, heat of formation, and crystal density can be computed automatically when a new HEM is inputted into the MS-HEMs. In addition, the impact sensitivity can be calculated using in-house ANN codes. The search routine allows the general users to find the exact HEM and its properties by inputting a 2D chemical structure sketch, or to retrieve the HEMs and their properties by giving a range of properties. The on-line DB for HEMs is a potential molecule pool that combines real and virtual HEMs, and provides important knowledge in locating good HEM candidates with a sufficiently safe nature.

Acknowledgments. This study was financially supported by ADD and the Defence Acquisition Program Administration (DAPA).

References

1. Cooper, P. W. *Explosives Engineering*; VCH: New York, 1996.
2. Miller, R. S. Research on New Energetic Materials. In *Decomposition, Combustion, and Detonation Chemistry of Energetic Materials*; Brill, T. B., Russell, T. P., Tao, W. C., Wardle, R. B., Eds.; Materials Research Society: Pittsburgh, 1996; p 3.
3. Agrawal, J. P.; Field, J. E. *Prog. Energy Combust. Sci.* **1998**, *24*(1), 1.
4. Rice, B. M. Overiew of Research in Energetic Materials. In *Energetic Materials. Part 1. Decomposition, Crystal and Molecular Properties*; Politzer, P., Murray, J. S., Eds.; Elsevier: Amsterdam, 2003; p 1.
5. Akhavan, J. *The Chemistry of Explosives*, 2nd ed.; The Royal Society of Chemistry: Cambridge, 2004.
6. Sikder, A. K.; Sikder, N. *J. Hazard. Mater.* **2004**, *112*(1-2), 1.
7. Badgujar, D. M.; Talawar, M. B.; Asthana, S. N.; Mahulikar, P. P. *J. Hazard. Mater.* **2008**, *151*(2-3), 289.
8. Borman, S. *Chem. Eng. News* **1994**, *72*(3), 18.
9. Simpson, R. L.; Urtiew, P. A.; Ornellas, D. L.; Moody, G. L.; Scribner, K. J.; Hoffman, D. M. *Propellants, Explosives, Pyrotechnics* **1997**, *22*(5), 249.
10. Lee, K.-Y.; Chapman, L. B.; Cobura, M. D. *J. Energ. Mater.* **1987**, *5*, 27.
11. Tarver, C. M. *Propellants, Explosives, Pyrotechnics* **1990**, *15*(4), 132.
12. Singh, G.; Kapoor, I. P. S.; Tiwari, S. K.; Felix, P. S. *J. Hazard. Mater.* **2001**, *81*(1-2), 67.
13. Brennan, M. B. *Chem. Eng. News* **2000**, *78*(23).
14. *ICT Thermochemical Database*, ver. 7.0; Institute of Chemical Technology: Pfinztal, Germany, 2004.
15. CTfile. http://accelrys.com/products/informatics/cheminformatics/ctfile-formats/no-fee.php (accessed August, 25, 2011).
16. *X3Dmole*, 1.0; BMDRC: 2007.
17. Benson, S. W. *Thermochemical Kinetics*, 2nd ed.; John Wiley & Sons Inc.: New York, 1976; p 22.
18. Cohen, N. *J. Phys. Chem. Ref. Data* **1996**, *25*(6), 1411.
19. Cohen, N.; Benson, S. W. *Chem. Rev.* **1993**, *93*(7), 2419.
20. Salmon, A.; Dalmazzone, D. *J. Phys. Chem. Ref. Data* **2006**, *35*(3), 1443.
21. Salmon, A.; Dalmazzone, D. *J. Phys. Chem. Ref. Data* **2007**, *36*(1), 19.
22. Ammon, H. L.; Mitchell, S. *Propellants, Explosives, Pyrotechnics* **1998**, *23*(5), 260.
23. Ammon, H. L. *Propellants, Explosives, Pyrotechnics* **2008**, *33*(2), 92.
24. Ammon, H. L. *Struct. Chem.* **2001**, *12*(3-4), 205.
25. Cho, S. G.; No, K. T.; Goh, E. M.; Kim, J. K.; Shin, J. H.; Joo, Y. D.; Seong, S. *Bull. Korean Chem. Soc.* **2005**, *26*(3), 399.
26. Storm, C. B.; Stine, J. R.; Kramer, J. F. *Sensitivity Relationships in Energetic Materials*; LA-UR-89-2936, Los Alamos Nat. Lab., NM, 1989.
27. Storm, C. B.; Stine, J. R.; Kramer, J. F. Sensitivity Relationships in Energetic Materials. In *Chemistry and Physics of Energetic Materials*, 2nd ed.; Bulusu, S. N., Ed.; Kluwer Academic Publishers: Dordrecht, Netherlands, 1990; p 605.
28. Ullmann, J. R. *J. ACM* **1976**, *23*(1), 31.
29. *DBDrawX*, 1.0; http://www.bmdrc.org/04_product/07_dbdrawx.asp, BMDRC: Seoul, 2007.