Data Mining for Terahertz Generation Crystals

(Enoch) Sin-Hang Ho, Gabriel A. Valdivia-Berroeta, Zachary B. Zaccardi, Sydney K. F. Pettit, Bruce Wayne Palmer, Matthew J. Lutz, Claire Rader, Brittan P. Hunter, Natalie K. Green, Connor Barlow, Coriantumr Z. Wayment, Daisy J. Harmon, Paige Petersen, Stacey J. Smith, David J. Michaelis, Jeremy A. Johnson
Brigham Young University, Provo, UT, 84602 USA

Abstract—Non-centrosymmetric molecular crystals are effective in nonlinear optical applications, such as second-harmonic generation, and optical rectification, due to low dielectric constants and high molecular hyperpolarizabilities. We designed a combined method of data mining for non-centrosymmetric structure from the Cambridge Structural Database (CSD) and performing density functional theory calculations to discover new organic nonlinear optical crystals that generate intense terahertz (THz) radiation. To confirm our combination approach, we recrystallized and tested the newly discovered organic nonlinear materials. The results of THz experiments showed the THz generation capabilities are comparable to state-to-art THz generation crystals (DAST, OH-I and BNA).

I. INTRODUCTION

Designing solid organic materials with useful properties is always a continually developing field. Unique properties of organic crystals are closely related to the molecular building blocks and structure of the materials. However, it is time-consuming and costly to identify and develop possible candidates through trial and error. Fortunately, the advancement of computational methods and development of material databases greatly enhance the process of identifying possible candidates for specific uses. Computational methods, such as density functional theory (DFT) calculation, calculate the properties of molecules while material databases can be used to provide required input information for the calculation, such as atomic position, molecular bonding and charges of molecules. One potential use of such databases is to identify already existing materials that may be ideal for applications other than their original intended use. The idea that one can easily mine information about known materials for the development of new and extremely useful purposes will rapidly accelerate the discovery of new materials for many applications. This data mining approach also gives rise to new screening methodologies in the rapidly growing field of materials informatics.

Terahertz (THz) radiation is the span of frequencies between infrared and microwave in order of $10^{12} - 10^{15}$ Hz (wavelength range between 30 to 300 μm). THz light exhibits unique interactions with many materials because its frequency range is very similar to the vibrational frequencies of collective atom motions in a host of materials. Thus, THz light can efficiently interact with these collective motions and can be used to analyze and control material properties in ways that differ from other forms of radiation. Many emerging and potentially disruptive applications of THz spectroscopy are taking advantage of these unique interactions, including in bioimaging and security, chemical recognition, non-destructive chemical monitoring in industry and food processing, and wireless communication and high-speed computational devices.

Of the various methods available for generating THz light, optical rectification of infrared (IR) light with organic nonlinear optical (NLO) crystals is the most efficient method to produce high-intensity THz fields with extremely broad bandwidths. Organic crystals are efficient THz generators due to the fact that the organic molecules have high hyperpolarizability and the crystal packing can be improved by modifying the molecular building blocks.

In this presentation we combine data mining of known organic materials from the Cambridge Structural Database (CSD) with DFT calculation of key molecular properties to identify new candidate organic materials for intense terahertz (THz) generation. We then validate our combined data mining and computational approach to materials discovery by synthesizing and fully characterizing the crystallographic structures and the THz generation capability of four new THz generating organic materials via single-crystal X-rays diffractometry and electro-optical sampling.

![Image](315x263 to 565x391)

**Fig. 1.** A schematic diagram illustrating the combined data mining/computational approach, starting from data extraction, computation, to material identification.

II. RESULTS

Our data mining effort focused on compounds that contain conjugated π-systems, neutral molecules with light atoms (lighter than Br) and non-centrosymmetric crystal. A custom python program was developed to search the CSD database based on the above criteria. We also restricted to only search for structures with r refinement values less than 15% to filter out less reliable structures. From a starting pool of over 1 million materials, 15,782 compounds from the CSD were isolated based on our selected criteria. The custom python code also extracted the molecular files and the crystallographic information files of the resulting compounds. To evaluate the THz generation capability of the extracted structures, all the resulting compounds were submitted for DFT
promising THz generators, we pumped all four crystals with NMBA and TMOAT. The calculated hyperpolarizability (β) parameters to be our candidates for THz generation. Out of the packing order parameter (values between 0 to 1, a value of 1 represents ideal head-to-tail alignment) were also calculated through an automated python program. We selected 10 molecules with high hyperpolarizabilities and variety of order parameters to be our candidates for THz generation. Out of the candidates, four were synthesized on a large scale and were amenable to large crystal growth, namely PNPA, ZPAN, NMBA and TMOAT. The calculated hyperpolarizability (β) values are 3.16 x 10-29 esu, 9.52 x 10-29 esu, 8.19 x 10-29 esu and 5.08 x 10-29 esu; and the order parameter values (OP) are 1, 0.55, 0.84 and 0.39. The values are comparable to well-known THz generation crystals, namely DAST (β: 1.95 x 10-29 esu, OP: 0.83) and OH-1 (β: 8.96 x 10-29 esu, OP: 0.68). In order to verify that the structures of the grown crystals matched the reported CSD structures, the developed crystals were characterized through single crystal X-ray diffractometry. The structures of all four crystals are consistent with the previously reported structures.

To test the THz generation capability of the four new promising THz generators, we pumped all four crystals with either 1450-nm or 1250-nm pump wavelength based on the crystal colors and compared the spectrums to current state-of-the-art organic THz generators (DAST, OH-1 and BNA). The crystals were pumped with different wavelengths of light based on the colors of the crystals as yellow organic crystals often exhibit higher efficiency at shorter wavelengths comparing to the red organic crystals.

All four newly identified crystals exhibit efficient THz generation capability (Fig. 3). In the red crystal category, we compared THz generation from an 850 μm thick PNPA and a 600 μm thick ZPAN crystal with industry-standard THz generators DAST (470 μm) and OH-1 (300 μm) by irradiation at 1450 nm pump wavelength under the same conditions. The THz spectrum of PNPA exceeds that of both OH1 and DAST, with a spectral peak at 1 THz. (Fig. 3a). Even though the amplitude of the Fourier transform is less than other red crystals, ZPAN also generates THz with only small absorptions before 3THz. This result is consistent with the hyperpolarizability and order parameter calculations, which PNPA has the largest hyperpolarizability and order parameter values.

In the yellow crystal category, a 370 μm thick NMBA and a 540 μm thick TMOAT crystals were compared with a well-known yellow BNA crystal (400 μm thick) and red inorganic crystal GaP (300 μm) by irradiation at 1250 nm pump wavelength under the same conditions. The THz generation of NMBA is very broadband, and its spectroscopic shape and amplitude of the Fourier transform are comparable to that of BNA (Fig. 3b). TMOAT, on the other hand, is not as efficient as BNA and NMBA in THz generation. However, TMOAT still has a board spectrum with only one main absorption at 1 THz. The lower intensity of THz generation from TMOAT is consistent to our prediction based on the low calculated hyperpolarizability and order parameter values.

These results confirm that our combination approach of data mining and DFT calculation can indeed provide a productive and powerful method for identifying new organic materials for THz generation and even predicting relative THz generation capabilities. This combined approach is not limited to only the field of THz, but it is also highly applicable in any fields of materials discovery and development.

REFERENCES
1. Varghese, S.; Das, S., Role of Molecular Packing in Determining Solid-State Optical Properties of π-Conjugated Materials. The Journal of Physical Chemistry Letters 2011, 2 (8), 863-873.
2. Chan, W. L.; Deibel, J.; Mittelman, D. M., Imaging with terahertz radiation. Reports on Progress in Physics 2007, 70 (8), 1325-1379.
3. Baxter, J. B.; Guglietta, G. W., Terahertz Spectroscopy. Analytical Chemistry 2011, 83 (12), 4342-4368.
4. Gowen, A. A.; O’Sullivan, C.; O’Donnell, C. P., Terahertz time domain spectroscopy and imaging: Emerging techniques for food process monitoring and quality control. Trends in Food Science & Technology 2012, 25 (1), 40-46.
5. Akyildiz, I. F.; Jomet, J. M.; Han, C., Terahertz band: Next frontier for wireless communications. Physical Communication 2014, 12, 16-32.
6. Brunner, F. D. J.; Kwon, O. P.; Kwon, S.-J.; Jasbinšek, M.; Schneider, A.; Günter, P., A hydrogen-bonded organic nonlinear optical crystal for high-efficiency terahertz generation and detection. Opt. Express 2008, 16 (21), 16496-16508.