Published in final edited form as:

J Am Soc Mass Spectrom. 2013 July ; 24(7): 1026–1036. doi:10.1007/s13361-013-0633-x.

Decoding Split and Pool Combinatorial Libraries with Electron Transfer Dissociation Tandem Mass Spectrometry

Mohosin Sarkar^{1,2}, Bruce D. Pascal^{3,4}, Caitlin Steckler⁴, Claudio Aquino¹, Glenn C. Micalizio¹, Thomas Kodadek^{1,2}, and Michael J. Chalmers^{4,5}

¹Department of Chemistry, The Scripps Research Institute, Scripps Florida, Jupiter, FL 33458, USA

²Department of Cancer Biology, The Scripps Research Institute, Scripps Florida, Jupiter, FL 33458, USA

³TRI Informatics, The Scripps Research Institute, Scripps Florida, Jupiter, FL 33458, USA

⁴Mass Spectrometry and Proteomics, The Scripps Research Institute, Scripps Florida, Jupiter, FL 33458, USA

⁵Molecular Therapeutics, The Scripps Research Institute, Scripps Florida, Jupiter, FL 33458, USA

Abstract

Screening of bead-based split and pool combinatorial chemistry libraries is a powerful approach to aid the discovery of new chemical compounds able to interact with, and modulate the activities of, protein targets of interest. Split and pool synthesis provides for large and well diversified chemical libraries, in this case comprised of oligomers generated from a well-defined starting set. At the end of the synthesis, each bead in the library displays many copies of a unique oligomer sequence.

Because the sequence of the oligomer is not known at the time of screening, methods for decoding of the sequence of each screening “hit” are essential. Here we describe an electron transfer dissociation (ETD) based tandem mass spectrometry approach for the decoding of mass-encoded split and pool libraries. We demonstrate that the newly described “chiral oligomers of pentenoic amides (COPAs)” yield non-sequence-specific product ions upon collisional activated dissociation; however, complete sequence information can be obtained with ETD. To aid in the decoding of libraries from MS and MS/MS data, we have incorporated ⁷⁹Br/⁸¹Br isotope “tags” to differentiate N- and C-terminal product ions. In addition, we have created “Hit-Find,” a software program that allows users to generate libraries in silico. The user can then search all possible members of the chemical library for those that fall within a user-defined mass error.

Keywords

MS; MS/MS; ETD; ECD; COPA; Combinatorial; Split-pool

Introduction

Since its recognition by Lam et al in 1991, the split and pool technique has become a common method for the synthesis of “One-Bead-One-Compound” combinatorial libraries [1]. These compound collections, in which each bead contains many copies of a single small molecule, have proven to be powerful tools in biomedical research for the identification of lead compounds for drug discovery and medicinal chemistry [2].

Split and pool methods have been successfully applied to synthesize peptides, peptoids, chemical oligomers, and other peptidomimetic libraries (azapeptide, azatide, vinylogous peptide, oligoureia, oligocarbamate, etc.). Libraries have been created that access over one million compounds. Because of the spatially separable nature of this type of library, they are suitable for the screening of the biological targets. Libraries synthesized by the split and pool method have been successfully applied to many targets, including antibodies, proteases, protein kinases (and other enzymes), adhesion molecules, G-protein coupled receptors, artificial receptors, bacteria, and even whole cells, for the discovery of new compounds with novel biological properties [2, 3].

Chemical and mass spectrometry-based methods have been widely applied for the structural elucidation of oligomers cleaved from single beads. In a recent study Pei and coworkers described high-throughput sequencing of hybrid oligomers based on partial Edman degradation and MALDI MS [4]. In other work, the structures of peptoid, peptide, azapeptide, and other peptidomimetic oligomers have been elucidated from single beads by MALDI MS/MS [5, 6] or ESI MS/MS [7]. Collision-activated dissociation (CAD) methods are based on the masses of product ions formed by breaking the amide ($O=CR-NH-$) bonds. However, because of the limited methods and technologies available for decoding the sequence of unknown compounds from a single bead, the synthesis of large libraries has been limited to the peptides, peptoids, and peptidomimetics only. It should be noted that mass-based decoding can only be successful with mass-encoded libraries (isomeric monomeric units cannot be differentiated by mass alone). However, because of the sensitivity and speed of modern instruments, the use of mass spectrometry remains the method of choice for the decoding of mass-encoded split and pool libraries.

Electron transfer dissociation (ETD) [8] is a gas phase ion fragmentation technique that exhibits similar dissociation characteristics to electron capture dissociation (ECD) [9]. In general, both ETD [10] and ECD [11] provide a mechanism for peptide ion fragmentation that yields complementary information to that obtained with “slow heating” dissociation mechanisms, including CAD. During ETD and ECD, an electron is transferred to, or captured by, the positively charged ion, leading to bond cleavage. Whereas CAD is charge site directed and preferentially cleaves peptide ions at the CO–N amide bond yielding *b* and *y* type ions, ECD and ETD dissociate the peptide at the *N*-alkyl bond to form *c* and *z*[•] (or *c*[•] and *z* ions) yielding a more random series of ions. This random fragmentation allows extensive sequence information to be obtained from protein and peptide ions [11]. Unfortunately, propensity for CAD to cleave the weakest bond in a given peptide ion can lead to the preferred loss of labile post-translational modifications and neutral losses of NH₃ and H₂O. These neutral losses may result in a lack of sequence-specific product ion information [10, 11]. It has been demonstrated the ECD/ETD can fragment peptide ions without losses of labile PTMs, such as phosphorylation and glycosylation [12, 13]. The application of activated ion (AI) ECD and ETD combined with supplemental activation have improved the efficiency of the ECD/ETD event [14, 15]. In summary, ECD and ETD often yield structural information that cannot be obtained with CAD.

We have recently described a new class of chiral and conformationally constrained oligomers, termed “chiral oligomers of pentenoic amides” (COPAs) that can be used to construct mass encoded split and pool synthetic libraries for screening applications. From a 160,000 member COPA library, we were able to identify a noncovalent synthetic ligand to the DNA binding domain of the transcription factor p53 (for more details please refer to [16]). During this work, we discovered that these COPA compounds (unlike peptides and peptoids) could not be sequenced by CAD MS/MS (MALDI ToF-ToF and ESI ion trap). Therefore, we hypothesized that ETD may yield sequence-specific information through cleavage of the alpha-carbon-nitrogen bond. This was indeed the case and we were able to successfully sequence a number of hits from the p53 screen. Due to space limitations, we were unable to describe the ETD sequencing in any detail within the original COPA manuscript [16].

In this work, we describe how COPA compounds dissociate upon CAD and ETD. We apply ETD to sequence COPAs and other “CAD-intractable” compounds from related libraries (synthesized using 5-chloromethyl isoxazole-4-carboxylic acid and 4-bromocrotonic acid as halo acid substituents). All of these new submonomers made it possible to make large libraries with excellent functional diversity, novel main chain backbone diversity, and good conformational rigidity, which hold great promise to be a valuable source of protein ligands. Finally, we describe a software application titled “Hit-Find” that can be used together with accurate mass information to help identify compounds from “One-Bead-One-Compound” libraries. Together, the combination of accurate precursor ion mass measurement, ETD fragmentation, and the use of Hit-Find provide a robust method for sequencing hits from mass-encoded combinatorial screening libraries. To date, we have successfully sequenced hundreds of COPA-derived compounds with the use of ETD.

Methods

COPA Library Design and Synthesis

The COPA library was synthesized on TentaGel resin (1.5 g, ~520,000 beads/g, 0.55 mmol/g loading capacity) from Rapp Polymere GmbH, Germany. The approximate total number of beads was 780000, which was enough to get 4 to 5 duplicate copies of each library member (total number of beads/library diversity, $780000/160000=4.8$). The 4-mer library was synthesized on the TentaGel resin following one-bead-one-compound split and pool techniques [1, 16, 17]. Ten amine building blocks (as shown in Figure 3) were used at four positions. With COPA (*S*) and (*R*) isomers at four different positions, the total diversity of the library was 160,000. A polyamide linker (Figure 3) was synthesized first on TentaGel MB NH₂ resin (160 μ m) following standard solid-phase peptide synthesis and microwave assisted sub-monomer protocols [18, 19]. To avoid the artifacts of the MALDI-TOF MS, MALDI matrix, and tandem MS/MS with low molar mass compounds a longer linker was designed. A Met residue was used to allow the removal of the compound from TentaGel resin by CNBr treatment. A positively charged amino side chain was used to increase the ionization efficiency in mass spectrometry. The amine, 4-bromophenylethylamine, was used to help differentiate the sets of signals derived from N-terminal product ions (marked as c') and the C-terminal product ions (marked as z') that appear as a 2 Da isotopic doublet due the ⁷⁹Br/⁸¹Br isotope ratio within the z' fragments.

TentaGel resin (1.50 g, 0.72 mmol) was swelled in anhydrous DMF for 1 h. The beads were treated with 5 equiv. of Fmoc-Met-(OH) (1.35 g, 0.24 mmol), 5 equiv. of HOBt (487 mg, 3.60 mmol), 5 equiv. of HBTU (1.37 g, 3.60 mmol), and 5 equiv. of *N*-methylmorpholine (396 μ L, 3.60 mmol) with gentle shaking at room temperature for 3 h in a 50 mL glass reaction vessel from ChemGlass. The beads were washed thoroughly with DMF. The Fmoc group was removed by treating resin with 20 % piperidine in DMF for 20 min (2 $\times$) and the beads were thoroughly washed again with DMF. The rest of the linker peptoid was synthesized using microwave-assisted solid-phase submonomer methods for peptoid synthesis using boc-diaminobutane, 4-Br-pehenthylamine, methoxypropylamine, and isopropoxy propylamine.

For library synthesis TentaGel beads were split equally to distribute into 2 reaction vessels and subjected to the coupling of (*R*)- or (*S*)-D₃ pentenoic acid isomers. The beads (750 mg, 0.36 mmol in each vessel) were treated with 5 equiv. of (*R*)-isomer (291 mg, 1.8 mmol) or (*S*)-D₃ isomer (298 mg, 1.8 mmol), 7 equiv. of diisopropylcarbodiimide (DIC, 379 μ L, 2.52 mmol), and 5 equiv. of HOAt (245 mg, 1.8 mmol) in 15.0 mL of anhydrous DMF in a 25 mL disposable reaction column for solid-phase synthesis from Intavis AG. The coupling was carried out at 37 °C for 3 h with gentle shaking. The beads were thoroughly washed with DMF and the beads were pooled together and mixed thoroughly before subjecting them to amine displacement. The beads were split equally to distribute into 10 reaction vessels (150 mg, 0.072 mmol in each vessel) and each reaction vessel was treated with 3.0 mL of 1 M solution of one of 10 amines (10 amines in 10 different reaction vessels). The reaction was carried out at 37 °C for 3 h with gentle shaking. The beads were thoroughly washed with DMF, pooled together and mixed thoroughly. The acylation reaction with COPA (*R*) and (*S*)

isomers and the displacement reaction with 10 amines were repeated three more times to synthesize 4-mer COPA library. The side chain protecting groups were removed by treating pooled resins with 20 mL of 94 % trifluoroacetic acid (TFA), 2 % triisopropylsilane (TIS), 2 % thioanisole, and 2 % water with gentle shaking for 2 h at room temperature. The resins were washed with DCM thoroughly, dried under vacuum and stored at -20°C (see [16] for detailed experimental procedure).

Synthesis of Oligomers Using Isoxazole and Crotonic Acids

A polyamide linker (as shown in Figures 5 and 6) was synthesized following microwave assisted solid-phase submonomer protocols as described above. Following the linker a 5-mer oligomer was synthesized in which a novel isoxazole unit was sandwiched between normal benzyl amine and 2-methoxyethyl amine derived peptoid units. Peptoid units were synthesized following the protocols described above. The coupling of isoxazole unit was carried out by treating resin beads (100 mg, 0.055 mmol) with 10 equiv. of DIC (diisopropyl carboxiimide, 0.55 mmol) and 10 equiv. of 5-(chloromethyl)isoxazole-4-carboxylic acid in 2.0 mL DMF. The acylation reaction was carried out by applying microwave irradiation for 15 s (2 $\times$) set to deliver 10 % power with 30 s interval in between. The displacement reaction was carried out by treating the beads with 2 mL of 1 M benzylamine and by applying microwave irradiation as described above. The coupling of 4-bromocrotonic acids was carried out by at 37°C for 1 h by treating resin beads (100 mg, 0.055 mmol) with 5 equiv. of DIC and 5 equiv. of 4-bromocrotonic acid in 2.0 mL DMF. The displacement reaction was carried out using 1 M benzylamine in 2.0 mL DMF for 1 h at 37°C .

Decoding of COPA Sequence from Single Beads with Tandem Mass Spectrometry

COPA oligomers from the library and the small molecule oligomers (synthesized using isoxazole and bromocrotonic acid) were cleaved from random individual beads by treating with 20 μL of CNBr solution (50 mg CNBr in 1 mL 5:4:1 acetonitrile: acetic acid: H_2O) in a 96-well plate overnight at room temperature. The CNBr cocktail was evaporated to dryness in a high capacity Speedvac system (Explorer 2000). The COPA compound from a single bead was dissolved in 25 μL of 1:1 $\text{CH}_3\text{CN}:\text{H}_2\text{O}$ containing 0.1 % formic acid, and 1 μL of which was spotted on MALDI plate with α -cyano-4-hydroxycinnamic acid for MALDI-TOF MS. The rest of the compound was infused into the mass spectrometer with either a (1) 25 syringe coupled to a 30 μM stainless steel emitter (Proxeon) coupled with a nano ESI source, or (2) with an automated chip based ESI source (Triversa Nanomate; Advion Biosciences). High resolution product ion data were acquired with an LTQ-Orbitrap XL (Thermo Scientific) equipped with ETD at a resolving power of 60,000 at m/z 400. Low resolution data were acquired with a linear ion trap (LTQ-XL ETD, Thermo Scientific). For all ETD experiments the precursor ion isolation width was set to 5 Th with supplemental activation of the charge reduced precursor ion species. ETD reaction time was set to 100 ms for $[\text{M}+3\text{H}]^{3+}$ ions and 150 ms for $[\text{M}+2\text{H}]^{2+}$ ions. Reduced ETD reaction times may be applied; however, some loss in efficiency of fragmentation was observed in some cases (data not shown). The ETD reaction times provided here yielded close to optimal fragmentation efficiency for all compounds characterized to date. Spectra are the average of ~ 40 scans.

Hit-Find

Hit-Find was developed using Java/Swing technology for the core and graphical user interface (GUI) components, with the backend data stored in a SQLite database format. The program initially stores the individual elemental compositions for each residue in a “look up” table. This table is then used to assemble the cumulative elemental composition of all possible oligomer combinations. For each elemental composition, a theoretical isotopic distribution is calculated using the software qmass [20, 21], and the corresponding monoisotopic mass is stored in the database. These masses are then indexed, and can be searched against measured masses of “hit” compounds. Those members of the library, within a user-specified error tolerance, are reported back in the GUI.

Results and Discussion

Proof-of-Concept Compounds

The workflow of the split and pool synthesis protocol is shown in Figure 1a. Following this protocol it is possible to generate highly diverse chemical libraries for screening campaigns containing in excess of 100,000 compounds. However, prior to construction of a library synthesized with a novel class of chemical monomers (such as COPA), it is essential to demonstrate that the compounds can be reliably synthesized, and that the sequence of each member of the library can be accurately determined. Therefore, prior to generating a COPA library, we synthesized a series of proof-of-concept COPA compounds for evaluation.

The sequence of one proof-of-concept COPA compound is shown in Figure 1b. The compound contains constant N- and C-terminal regions connected by monomeric peptoid and COPA units that are generated from the coupling of an amine and chloropentenoic acids. Single COPA monomeric units are colored red and blue within Figure 1b. The peptoid monomeric unit is colored green.

MS/MS Characterization of COPA Compounds

In order to be suitable for use in a mass encoded screening library, the chemical identity of each COPA molecule must be able to be determined after the compound is cleaved from the bead. Mass spectrometry is an attractive technique for sequencing library hits and CAD has been widely applied to peptide and peptoid libraries. Within our laboratories, MALDI ToF-ToF has been the technique of choice for the sequencing of library hits [18]. It was quickly demonstrated that our model COPA compounds, including the compound shown in Figure 1b, did not fragment to yield sequence specific product ions from the CAD of the $[M+H]^+$ precursor ion (data not shown). In order to determine if the precursor ion charge state would impact the CAD of the COPA compound we performed ESI MS/MS within a linear ion trap instrument. The product ion spectrum obtained from the CAD of the $[M+2H]^{2+}$ precursor ion is shown in Figure 2a. A number of dominant neutral losses were observed, including losses of NH_3 from the terminal primary amines located within side chains of the C-terminal linker region. Neutral losses of 104 Da were also observed and were attributed to side chain loss. A number of ions were generated that provided some limited structural information. The ions at m/z 359 and 487 provide some information relating to the location of one of the two butylamine side chains located within the C-terminal linker region; however, no COPA

unit sequence information could be determined. Based upon these results, it was concluded that any newly synthesized COPA library could not be decoded with CAD based tandem mass spectrometry.

Electron transfer dissociation has been shown to provide access to fragmentation pathways that are often complementary to CAD. Therefore we performed ETD of the $[M+2H]^{2+}$ precursor ion generated from the ESI of the COPA compound shown in Figure 1b. The ETD product ion spectrum is shown in Figure 2b and in contrast to CAD, the ETD spectrum yielded extensive sequence-specific product ions. Abundant product ions were observed, along with the charge reduced radical $[M+2H]^{1+\bullet}$ ion. In a behavior similar to that observed for peptide ions [14, 15], it was noted that the intensity of the COPA product ions could be increased with supplemental activation of the charge reduced precursor ion.

After manual inspection of the data, the location of each bond cleavage along with the corresponding product ions are shown in Figure 2b. It was apparent that the complete sequence of this COPA molecule could be readily determined based upon the delta mass shift observed between cleavages at sequential *N*-alkyl bonds. For example, the COPA monomer highlighted in Figure 1b can be identified within the compound from the delta mass between ions at m/z 689.41 and 918.56. The adjacent COPA monomer can be identified from the delta mass measured between the ions at m/z 918.56 and 1133.69. It should be noted that certain product ions exhibited the loss of $H^{\bullet}$, as was observed with the product ion labeled “2” shown within the inset to Figure 2b. This behavior is consistent with peptide fragmentation during ECD and ETD [9, 15]. A proposed scheme for the generation of product ions 3, 6, and 7 are shown in Supplemental Figure 1.

In order to increase our confidence that we correctly interpreted the spectrum shown in Figure 2b, we calculated the mass error between the theoretical m/z , and the measured m/z for a number of product ions. The assigned elemental composition and associated mass error (ppm) are detailed within Table 1. It should be noted that this spectrum was acquired with a resolution of $\sim 60,000$ at m/z 400 within the Orbitrap mass analyzer. All product ions were measured with a few ppm of the theoretical value, thereby increasing our confidence in the assignment of the fragmentation behavior of the COPA compound.

Together, these data show that we have a robust, sequence-specific fragmentation technique that can be used to sequence unknown members of any COPA library.

COPA Ion ETD Product Ion Nomenclature

Based loosely upon the Roepstorff rules of peptide ion fragmentation [22], we propose the use of c' and z' to assign the identity of ions generated from the ETD of COPA (and other monomeric) compounds originating from split and pool combinatorial libraries.

IMPORTANT NOTE: It should be emphasized that the COPA synthesis procedure originates from the first monomeric unit and terminates with the addition of the polyamide linker region (Figure 3). Therefore, as shown in Figures 3 and 4, the “N-terminus” of the COPA compound is on the right hand side of the molecule as it is typically drawn. In order to be consistent with peptide ion fragmentation and yet differentiate between cleavage at the amide (peptide) bond versus cleavage of the *N*-alkyl (COPA) bond, we have chosen to

annotate these N-terminal ions as c' ions. The complimentary C-terminal ions are described as z' ions (Figure 4).

Adding a Bromine Isotope “Tag” to Aid De-Novo Sequencing

The ability to rationally design the compounds prior to generation of the library allows for the inclusion of isotope “tags” to differentiate between the N-terminal and C-terminal product ions. Having determined that we could sequence the proof-of-concept COPA compounds with ETD, we decided to modify the constant C-terminal region of the COPA compound to include a bromine atom. Because each bromine atom contains a ^{81}Br isotope of (almost) equal abundance to the ^{79}Br isotope, the z' ions could be clearly distinguished from the c' ions (see z'_6 within the inset of Figure 4). Therefore, in the final design of the COPA library [16], we included a bromine atom to help identify N-terminal z' fragment ions during the de-novo sequencing of the library compounds (see inset to Figure 4).

The 160,000 Member COPA Tetramer Library

The design of the 160,000 member COPA library used in our previous work is shown in Figure 3. Each COPA monomer is formed through the reaction between the amine building blocks and the chloropentenoic acids (Figure 3, left). For structure elucidation by MS, the (S) isomer of the chloropentenoic acid was labeled with deuterium (CD_3 at C2) to correlate differences in mass of fragment ions with absolute stereochemistry of the COPA isomer.

In order to assess the quality and fidelity of the library synthesis, a number of beads were picked and the COPA compounds were cleaved to assess purity and also to further demonstrate the ability of ETD to decode the library. Figure 4 shows the ETD product ion spectrum generated from the $[\text{M}+3\text{H}]^{3+}$ precursor ion. The sequence was readily determined with complementary information being obtained from complete c' and z' series. In the case of the z' ions, each isotope cluster was readily identified due to the $^{79}\text{Br}/^{81}\text{Br}$ isotope.

Hit-Find

In order to facilitate the MS based sequencing of unknown compounds originating from mass encoded split and pool chemical libraries we developed a Java based software application titled “Hit-Find.” A screen shot from the program is shown in Supplemental Figure 2. Within the application the user can generate the elemental composition, and theoretical m/z and z values for every possible compound. The user can then search the measured m/z value and z against this list to identify all possible elemental compositions within a user defined ppm range. For example, the COPA compound shown within Figure 4 has a COPA monomeric sequence of “ippa-napiD3-iaa-iaa-ndap”. The elemental composition of the compound is $\text{C}_{81}\text{H}_{131}\text{D}_3\text{BrN}_{13}\text{O}_{12}$ which has a calculated $[\text{M}+\text{H}]^+$ value of 1563.97651 Th. A Hit-Find search of the 160,000 member COPA library yielded only one elemental composition (48 possible compounds) within a 5 ppm mass error. With an increased mass error of 50 ppm, the number of possible elemental compositions increases to seven and the number of possible compounds expands to 419. In future editions of the software we will include product ion m/z data. It is appealing to envisage a software platform that operates in a similar fashion to peptide database search engines.

ETD Sequencing of Additional COPA-Derived Libraries

In order to make structurally diverse and conformationally rigid oligomers we investigated other alternatives to the pentenoic acid (in COPAs) or bromoacetic acid (in peptoids) subunits. Therefore, we synthesized heterogeneous libraries based upon 5-(chloromethyl)isoxazole-4-carboxylic acid and also from 4-bromocrotonic acid derived oligomers. Like COPAs, the lack of diagnostic sequence specific ions from collisional activated dissociation (CAD) of $[M+H]^+$ precursor ions precluded the use of MALDI MS/MS for decoding their structure. But, in both cases the test compounds from the new libraries exhibited sequence specific fragmentation patterns with ETD and were able to be sequenced correctly. Figure 5 shows the structure (**a**) and high resolution ETD fragmentation spectrum (**b**) for the pentamer in which isoxazole unit was sandwiched between benzylamine derived peptoid units. Figure 6 shows the structure (**a**) and ETD fragmentation spectrum (**b**) for the pentamer in which two 4-bromocrotonic units were sandwiched between benzylamine derived peptoid units.

Mechanism of Fragmentation

Although the mechanism of fragmentation is still unknown, we hypothesized that the presence of a quaternized amide nitrogen within the protonated parent ion captures an electron and leads to the cleavage of one of two *N,N*-alkyl bonds. As shown in Figure 7a (COPA, peptoid, isoxazol, and crotonic acid oligomers) the backbone fragmentation is favored because the product ions generated by ETD are stabilized by the presence of resonance structures. The cleavage of side chains was favored only when stable side chain radical ions were generated by ETD (Figure 7b). Among dozens of side chains we have used for oligomer synthesis, side chain fragmentations were observed only in place of benzylic substituents.

Conclusions

It is clear that the compounds within our split and pool COPA libraries are intractable to decoding with CAD based MS/MS techniques. Neither ESI CAD (linear ion trap) of multiply protonated COPA ions nor MALDI CAD (ToF/ToF) analysis of singly protonated ions yielded sequence-specific information. In contrast, we demonstrate that ETD yielded extensive sequence information from which we could determine the monomeric sequence of compounds cleaved from the library beads. The ability to sequence the COPA compounds with ETD allowed us to move forward and generate a 160,000 member library that was used to discover the first noncovalent synthetic ligand to the DNA binding domain of p53 [2]. Without ETD, the results from the p53 screen could not have been decoded, and the screen would not have been successful. We also demonstrate how ETD can be applied to other “CAD intractable” compounds, including oligomers synthesized from isoxazole and bromocrotonic acids. In order to simplify the de-novo sequencing of hits from these split and pool libraries, we designed $^{79}\text{Br}/^{81}\text{Br}$ isotope “tags” within the linker region to differentiate between *c'* and *z'* ions (our proposed nomenclature for these ions is described above). Finally, we describe a software application, titled Hit-Find, to aid in the sequencing of compounds from libraries (<http://proteomics.florida.scripps.edu/Informatics.html>). The

ETD method has proven to be robust, and over the last year we have successfully decoded hundreds of hits from screening of various libraries (predominantly COPA).

In summary, the application of ETD to the de-novo sequencing of the oligomeric split and pool libraries has significantly expanded the number of chemical libraries available for use by medicinal chemistry and drug discovery efforts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

G.C.M acknowledges financial support from the Fidelity Biosciences Research Initiative. T.K. acknowledges support from the NHLBI (N01-HV-00242). The authors thank Patrick Griffin for access to the LTQ-Orbitrap with ETD.

References

1. Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature*. 1991; 354:82–84. [PubMed: 1944576]
2. Lam KS, Lebl M, Krchnak V. The “one-bead-one-compound” combinatorial library method. *Chem. Rev.* 1997; 97:411–448. [PubMed: 11848877]
3. Ramnarayanan R, Chan BC, Salvitti MA, Mallouk TE, Falihi FM, Davis J, Galloway DB, Bare SR, Willis RR. Directed-sorting method for synthesis of bead-based combinatorial libraries of heterogeneous catalysts. *J. Combinat. Chem.* 2006; 8:199–212.
4. Thakkar A, Cohen AS, Connolly MD, Zuckermann RN, Pei D. High-throughput sequencing of peptoids and peptide-peptoid hybrids by partial edman degradation and mass spectrometry. *J. Combinat. Chem.* 2009; 11:294–302.
5. Simpson LS, Kodadek T. A cleavable scaffold strategy for the synthesis of one-bead one-compound cyclic peptoid libraries that can be sequenced by tandem mass spectrometry. *Tetrahedron Lett.* 2012; 53:2341–2344. [PubMed: 22736872]
6. Aditya A, Kodadek T. Incorporation of heterocycles into the backbone of peptoids to generate diverse peptoid-inspired one bead one compound libraries. *ACS Combinat. Sci.* 2012; 14:164–169.
7. Ruijtenbeek R, Versluis C, Heck AJ, Redegeld FA, Nijkamp FP, Liskamp RM. Characterization of a phosphorylated peptide and peptoid and peptoid-peptide hybrids by mass spectrometry. *J. Mass Spectrom.* 2002; 37:47–55. [PubMed: 11813310]
8. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. USA.* 2004; 101:9528–9533. [PubMed: 15210983]
9. Zubarev RA, Kelleher NL, McLafferty FW. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* 1998; 120:3265–3266.
10. Udeshi ND, Shabanowitz J, Hunt DF, Rose KL. Analysis of proteins and peptides on a chromatographic timescale by electron-transfer dissociation MS. *FEBS J.* 2007; 274:6269–6276. [PubMed: 18021239]
11. Cooper HJ, Hakansson K, Marshall AG. The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom. Rev.* 2005; 24:201–222. [PubMed: 15389856]
12. Hakansson K, Chalmers MJ, Quinn JP, McFarland MA, Hendrickson CL, Marshall AG. Combined electron capture and infrared multiphoton dissociation for multistage MS/MS in a Fourier transform ion cyclotron resonance mass spectrometer. *Anal. Chem.* 2003; 75:3256–3262. [PubMed: 12964777]

13. Wiesner J, Premisler T, Sickmann A. Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. *Proteomics*. 2008; 8:4466–4483. [PubMed: 18972526]
14. Chalmers MJ, Hakansson K, Johnson R, Smith R, Shen JW, Emmett MR, Marshall AG. Protein kinase A phosphorylation characterized by tandem Fourier transform ion cyclotron resonance mass spectrometry. *Proteomics*. 2004; 4:970–981. [PubMed: 15048979]
15. Swaney DL, McAlister GC, Wirtala M, Schwartz JC, Syka JE, Coon JJ. Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. *Anal. Chem.* 2007; 79:477–485. [PubMed: 17222010]
16. Aquino C, Sarkar M, Chalmers MJ, Mendes K, Kodadek T, Micalizio G. A biomimetic polyketide-inspired approach to small-molecule ligand discovery. *Nat. Chem.* 2012; 4:99–104. [PubMed: 22270625]
17. Burkoth TS, Fafarman AT, Charych DH, Connolly MD, Zuckermann RN. Incorporation of unprotected heterocyclic side chains into peptoid oligomers via solid-phase submonomer synthesis. *J. Am. Chem. Soc.* 2003; 125:8841–8845. [PubMed: 12862480]
18. Olivos HJ, Alluri PG, Reddy MM, Salony D, Kodadek T. Microwave-assisted solid-phase synthesis of peptoids. *Org. Lett.* 2002; 4:4057–4059. [PubMed: 12423085]
19. Zuckermann RN, Kerr JM, Kent SBH, Moos WH. Efficient method for the preparation of peptoids [oligo(*N*-substituted glycines)] by submonomer solid-phase synthesis. *J. Am. Chem. Soc.* 1992; 114:10646–10647.
20. Rockwood AL, Haimi P. Efficient calculation of accurate masses of isotopic peaks. *J. Am. Soc. Mass Spectrom.* 2006; 17:415–419. [PubMed: 16458531]
21. Rockwood AL, Van Orman JR, Dearden DV. Isotopic compositions and accurate masses of single isotopic peaks. *J. Am. Soc. Mass Spectrom.* 2004; 15:12–21. [PubMed: 14698550]
22. Roepstorff P, Fohlman J. Letter to the editors. *Biol. Mass Spectrom.* 1984; 11:601–601.

Figure 1.

(a) Split and pool synthesis. A general scheme illustrating the process followed in split and pool library synthesis. **(b)** Chemical composition of a proof-of-concept COPA compound showing linker (black), peptoid (green), and COPA monomers (red and blue)

Figure 2.

CAD and ETD of COPA compound (1). **(a)** Product ion spectrum obtained from the linear ion trap CAD of the $[M+2H]^{2+}$ precursor ion. Dominant neutral losses of NH_3 and COPA monomer side chains are observed. Minimal sequence specific information is obtained from CAD of either the $[M+H]^+$ or $[M+2]^{2+}$ precursor ions. **(b)** High resolution (Orbitrap) ETD product ion spectrum of the $[M+2H]^{2+}$ precursor ion. Abundant product ions from cleavage of the C- α -N bond are observed and the complete sequence of COPA (1) can be determined.

It should be noted that loss and gain of $H^\bullet$ are observed from some product ions, including #5 [see inset to (b)]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. The 160,000 member COPA library. Left, amine and acid building blocks used to generate the library. Upper right, linker regions. Lower right, final COPA tetramer library with a chemical diversity of 160,000 compounds

Figure 4.

ETD sequencing of unknown COPAs aided by a bromine isotope tag. ETD spectrum showing the sequence obtained from COPA compound from a single random bead picked during the library quality evaluation process. Multiple sequence-specific product ions are observed and the sequence of the compound is readily decoded as "ippa-napiD3-iaa-iaa-ndap." Inclusion of the Br atom within the linker region allows all the z' product ions to be readily identified by the observed isotopic signature (inset)

Figure 5.

ETD sequencing of isoxazole-COPA compound. (a) Structure of oligomer synthesized with 5-(chloromethyl)isoxazole-4-carboxylic acid, (b) decoding of the oligomer sequence with ETD of the $[M+2H]^{2+}$ ion. ETD data were acquired with an Orbitrap mass analyzer operating at a resolving power of 60,000 at m/z 400

Figure 6. ETD sequencing of crotononic-COAP compound. **(a)** Structure of the oligomer synthesized with 4-bromorotononic acid, **(b)** decoding of the oligomer sequence with ETD of the $[M + 2H]^{2+}$ ion. ETD data were acquired with a linear ion trap

(a) Backbone fragmentation **Figure 7.**

Plausible fragmentation mechanisms for oligomers containing halo acids (pentenoic acid, isoxazol, crotonic acid, bromoacetic acid). **(a)** Capture of an electron by the precursor ions ($2H^+$) leading to the backbone fragmentation of the molecule and subsequent product ions detected by ETD. **(b)** Capture of an electron by the precursor ions ($2H^+$) leading to the cleavage of the side chains and subsequent product ions detected by ETD. Side chain fragmentation observed only for benzylic substituents

Table 1

Sequence specific product ions arising from the ETD of the COPA shown in Figure 2b. Note that the calculated theoretical m/z values do not include the mass of the odd electron (0.00054 Da) proposed for ions 4–7. Calculated theoretical m/z values correspond to the protonated $[M+nH]^{n+}$ ion. For product ions 1–7 the ions are singly protonated $[M+H]^+$. Theoretical values were calculated with IsoPro 3.1

Ion #	Proposed	$[M+nH]^{n+}$	$[M+nH]^{n+}$	m/z	PPM
	Elemental composition	Theoretical	Measured		
1	C64 H91 N9 O7	1098.71138	1098.709	−0.0024	−2.2
2	C58 H79 N7 O6	970.61644	970.6141	−0.0023	−2.4
3	C52 H68 N6 O4	841.53746	841.5342	−0.0033	−3.9
4	C24 H43 N6 O7 •	528.32657	528.3249	−0.0017	−3.2
5	C34 H54 N7 O8•	689.41064	689.4087	−0.0019	−2.8
6	C49 H73 N8 O9•	918.5573	918.5552	−0.0021	−2.3
7	C63 H90 N9 O10•	1133.68831	1133.6857	−0.0026	−2.3
$[M+2H]^{2+}$	C70 H98 N10 O10	620.38063	620.3788	−0.0018	−2.9