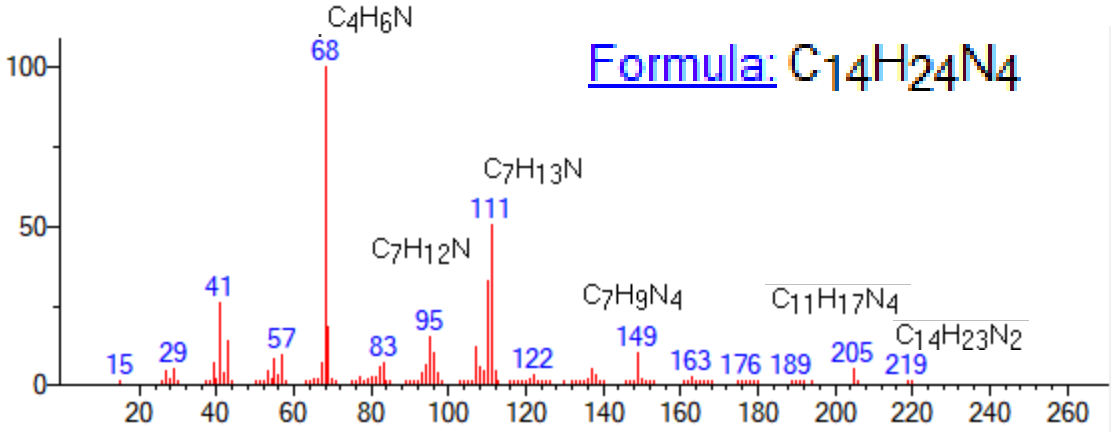
Comparison of Google Image Search to CAS and ChemSpider for Unknown Identifications Using Molecular Formula from Accurate Mass EI or MSMS

-For many years, we have used accurate mass data and either ChemSpider or SciFinder (CAS) to identify unknowns in commercial samples, extractables, etc.
-We refer[1,2,3] to these as *"known unknowns"*
-They are ***generally known*** in a variety such of databases such as ChemSpider or the CAS Registry
-***However***, their EI or MSMS spectra are not found in commercially available libraries
-This presentation compares using Google Image as another approach instead of ChemSpider or SciFinder for these type of identifications
-Possibly a viable approach which could compliment ChemSpider and SciFinder

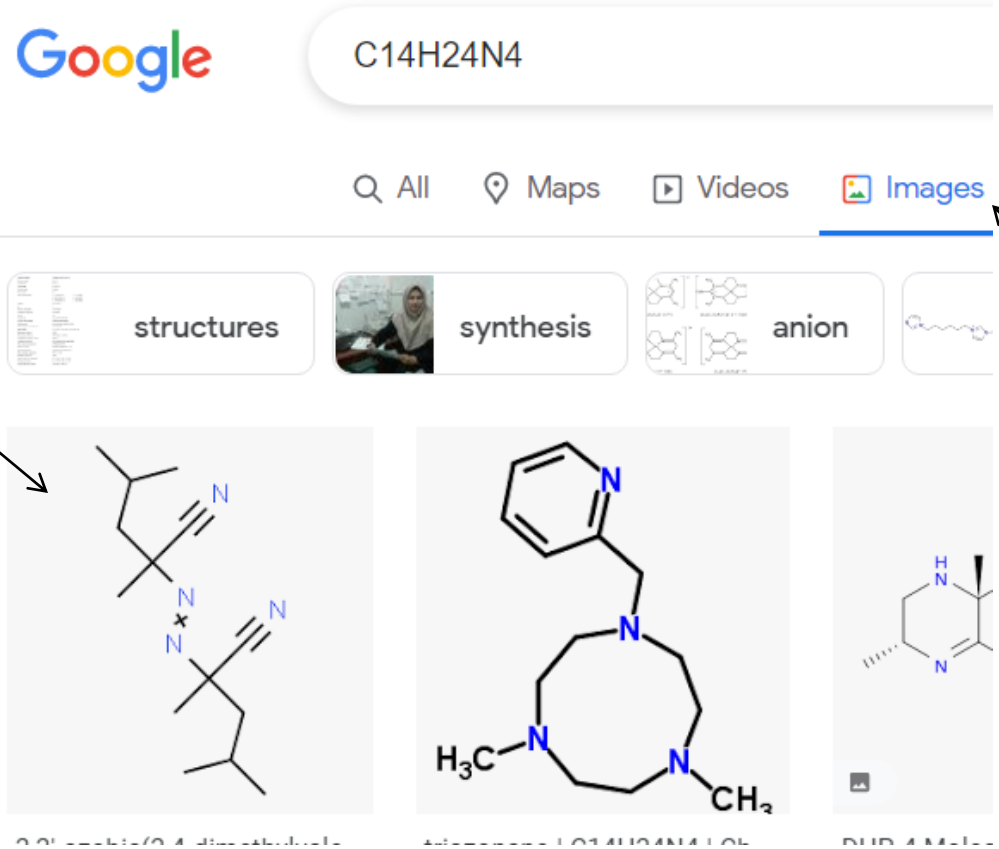
Example:

Accurate Mass EI and CI Data Obtainable from Orbitrap
Identity ***Not Found*** In NIST or Wiley EI Libraries Searches



Google Search Results for $C_{14}H_{24}N_4$

-View results of Google search using ***image*** option
-Nitrile-azo compound shows up 7 times out of the 23 structures displayed
-the proposed structure fits the EI fragmentation pattern

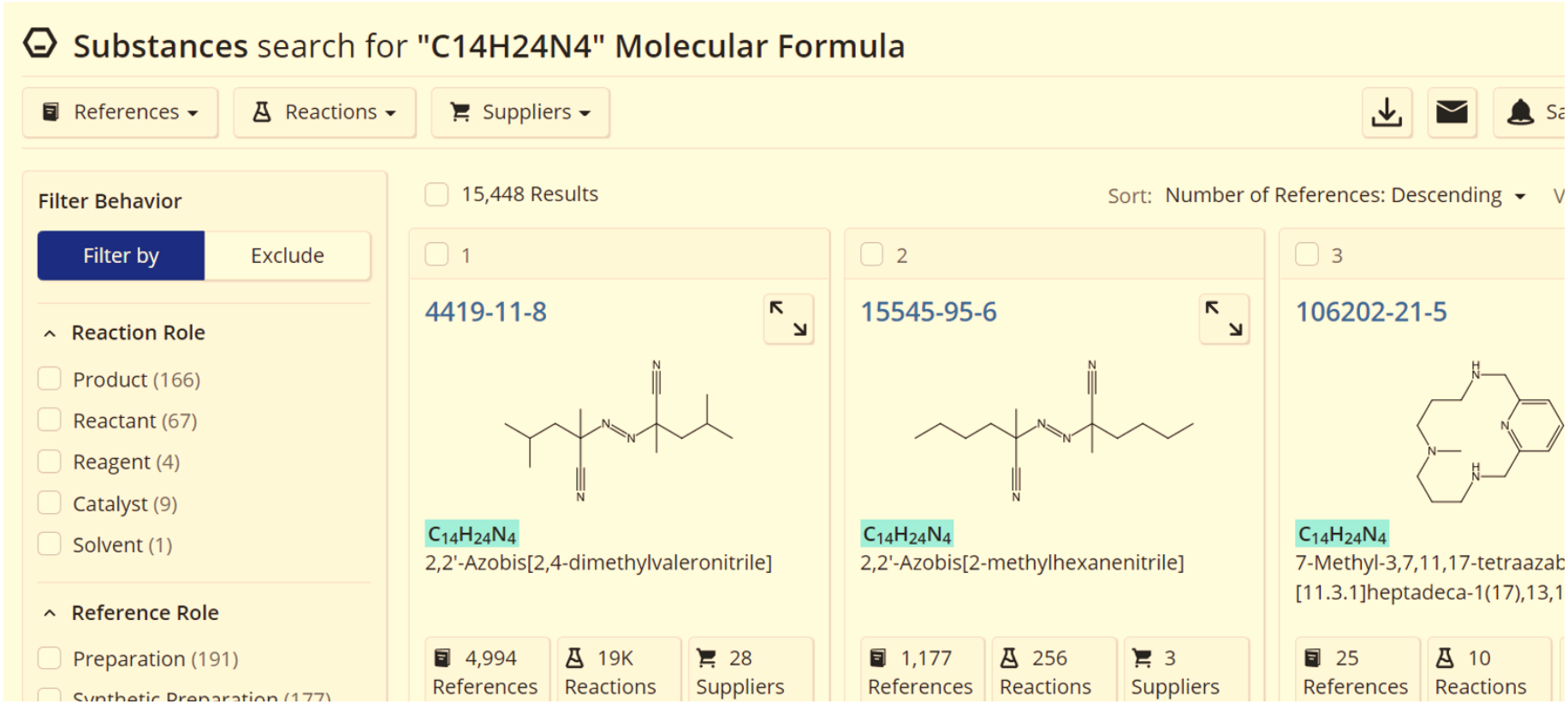


Correct Structure

Use Image Display Option

SciFinder (CAS Registry) Search Results for $C_{14}H_{24}N_4$

-Search with SciFinder and sort by descending associating references
-See top two references, would be hard to tell which isomer by EI
-However, more references could indicate that the isobutyl more likely



ChemSpider Search Results for $C_{14}H_{24}N_4$

-Search with ChemSpider and sort by descending associated references
-Top one consistent with EI fragmentation

Matches any text strings used to describe a molecule.

C14H24N4

Systematic Name, Synonym, Trade Name, Registry Number, SMILES, InChI or CSID

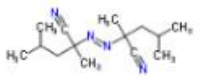
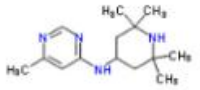
FILTER

Search Hits Limit: 100

Found 14345 results

Search term: **C14H24N4** (Found by molecular formula)

1 2 3 4 5

ID	Structure	Molecular Formula	Molecular Weight	# of Data Sources	# of References	# of PubMed	# of RSC
82654 - 0/2 defined		$C_{14}H_{24}N_4$	248.3672	52	120	85	16
20518212		$C_{14}H_{24}N_4$	248.3672	20	25	0	0

Conclusions

-Using Google search of a molecular formula from accurate mass data with subsequent display by images useful in getting possible candidate structures
-Then use EI or MSMS fragmentation to determine if reasonable
-SciFinder CAS registry search still likely the best, but not a free service
-ChemSpider is free and allows one to sort by relevant database, data sources, references to bring most likely structures to top of list
-Also, ChemSpider can be searched by accurate m/z values +/- error, SciFinder cannot
-Searching by accurate mass useful for higher molecular weight compounds (>500) when unsure of the molecular formula

References

1. Unknown Identification Using SciFinder (CAS Registry)
2. Unknown Identification Using ChemSpider
3. LCGC Article Giving Overview with Examples