

RI Calibration in NIST26 Chromatogram and Applying to Calculating RI in Samples Video/Handout

James Little

Mass Spec Interpretation Services

April 24, 2026

mzinterpretation.com

See **Full Course** on NIST26 with new **Integrated** Deconvolution/Library Searching for
EI GC-MS and **LC-MS/MS**!

Mass Spec (m/z) Interpretation Services

Organic Mass Spectrometry

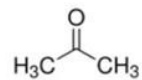
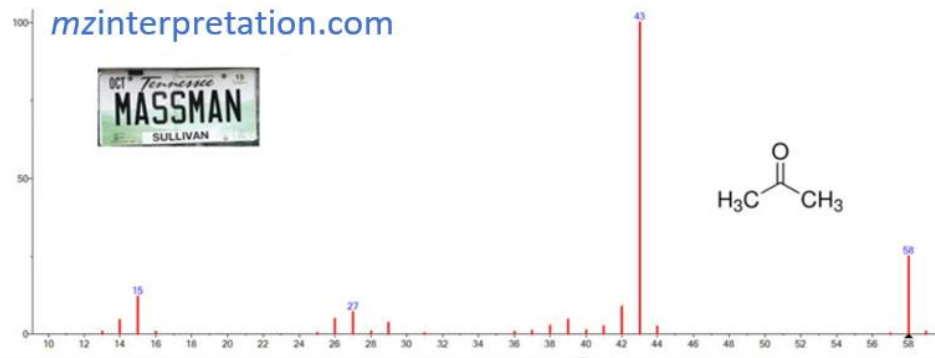


Table of Contents for Retention Index (RI) Video

- Importance of NIST RI and Their Approach
- More Detailed Approach for Stand-Alone AMDIS
- My Approach for NIST26 Integrated AMDIS
- NIST 10 Step Approach
- Purchasing Hydrocarbon Calibration Mixture

I will Discuss only Top 3 Topics, many more details in Handout Only!

Retention Indices (RI) in NIST GC/MS Software *(Summary)*

Retention indices (RI) provide an orthogonal dimension of identification alongside mass spectra, significantly improving confidence—especially for isomeric compounds that produce similar EI spectra.

Over recent library releases, NIST has made a major investment in RI data:

- Experimental RI values are now measured for essentially all newly added compounds
- These RI values are carefully curated and incorporated into the NIST EI library
- Where experimental data are not available, RI values are estimated using predictive models

This combined approach enables both measured and predicted retention behavior to be used during compound identification. Integrated RI Matching within their search software allows filtering and ranking based on RI data. This helps quickly eliminate incorrect isomer candidates.

Retention Indices (RI) in NIST GC/MS Software (*More Detail*)

Retention indices (RI) provide an **orthogonal dimension of identification** alongside mass spectra, significantly improving confidence—especially for isomeric compounds that produce similar EI spectra.

Over recent library releases, NIST has made a major investment in RI data:

- **Experimental RI values** are now measured for a large fraction of newly added compounds
- These RI values are **carefully curated and incorporated into the NIST EI library**
- Where experimental data are not available, **RI values are estimated using predictive models**

This combined approach enables both **measured and predicted retention behavior** to be used during compound identification.

Recent work from NIST researchers demonstrates that:

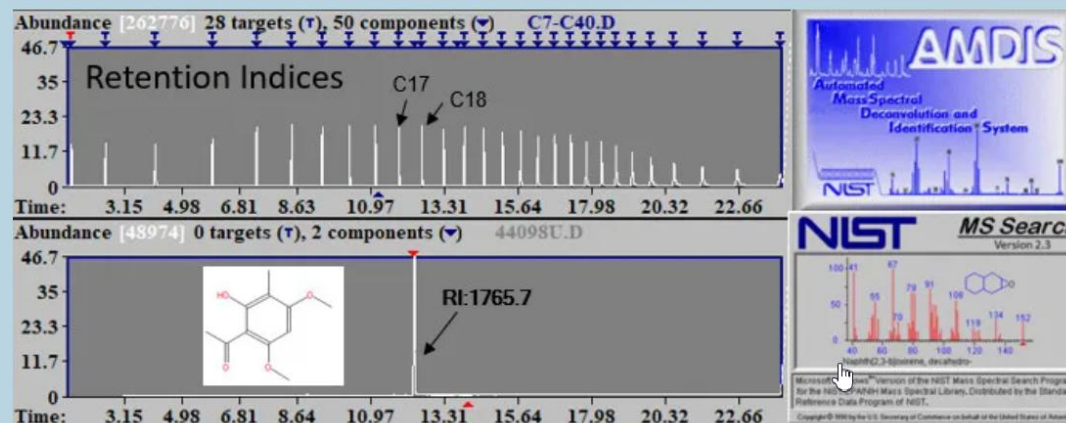
- The NIST RI database provides a **high-quality training set** for predictive modeling
- Modern AI approaches (e.g., deep neural networks) can predict RI values with **errors on the order of ~15–40 index units**

Integrated RI Matching within their search software allows filtering and ranking based on RI data. This helps quickly eliminate incorrect isomer candidates.

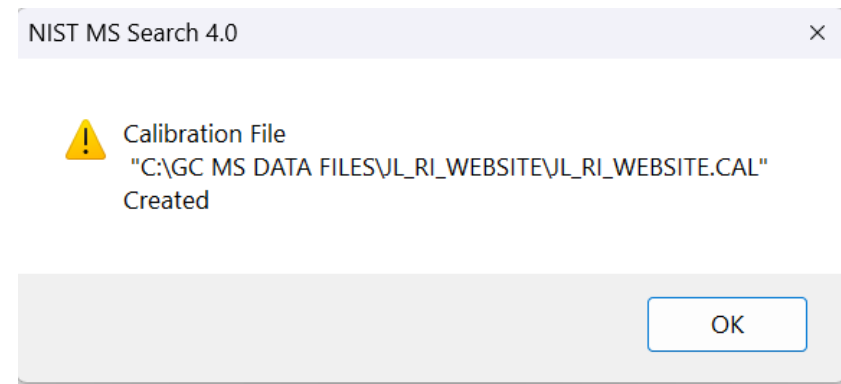
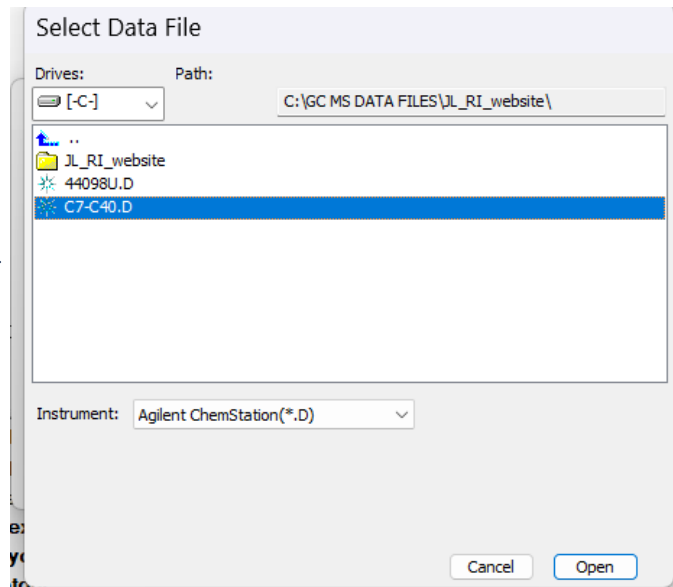
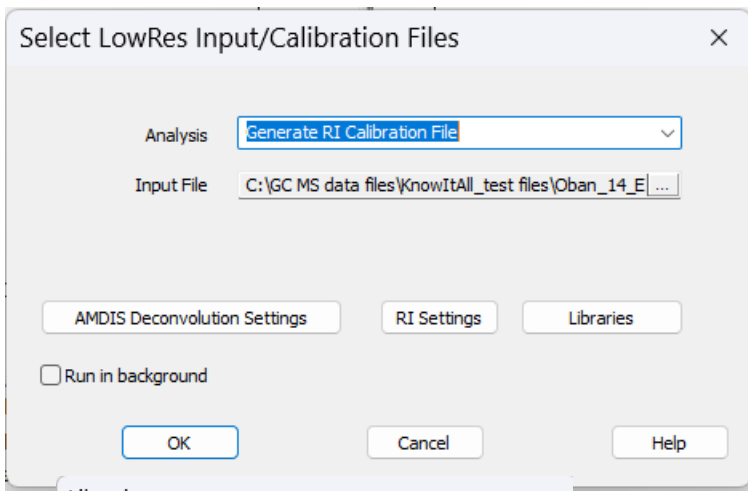
Previous Stand-Alone AMDIS RI Training

- Previously created information on RI calibration in the stand-alone AMDIS
- Included were video, associated handout, test files, and many additional resources
- **Still might be useful** if you want to have a **more in-depth understanding** of the process
- .D data file and associated .D hydrocarbon calibration mix
- **Simpler process** for **integrated** AMDIS **approach** in NIST26 describe here
- Click [this link](#) to access earlier more detailed explanation, header shown below

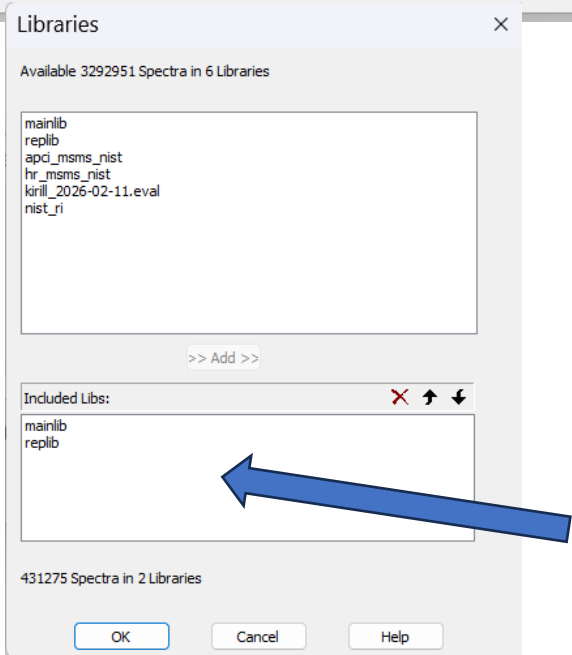
Part VI: Creating and Using Retention Indices in NIST Software



My Approach for RI Calibration in Chromatogram Window Using Files on My Website



- Creates a calibration file
- Returns the location and name
- **keep** in the folder that contains data file of interest



Selected NIST EI Reference Libraries

- Use Notepad to View the Cal File Created
- Found with Folder Name Containing Hydrocarbon Standard
- Missing C39 and C40 for RI 3900 and 4000, respectively
- Find Retention Times for These Peaks and Add to Calibration File
- **Peaks added will vary with calibration file obtained**
- Creating RI calibration routine has trouble with longer chain hydrocarbons because often *no molecular ion* in instrument spectrum and/or library spectrum

Local Disk (C:) > GC MS data files > JL_RI_website >

Sort View ...

Name	Date modified	Type	Size
44098U.D	2/3/2026 1:08 AM	File folder	
additional info	4/27/2026 10:06 AM	File folder	
C7-C40.D	2/3/2026 1:08 AM	File folder	
C7-C40.ELU	5/4/2026 10:08 PM	ELU File	72 KB
C7-C40.FIN	5/4/2026 10:08 PM	FIN File	1,586 KB
JL_RI_WEBSITE.CAL	5/4/2026 10:08 PM	CAL File	2 KB

File	Edit	View
0.675	600.0	96 1766 n-Hexane
0.934	700.0	100 1438 Heptane
1.485	800.0	100 682 Octane
2.561	900.0	100 687 Nonane
4.145	1000.0	100 682 Decane
5.957	1100.0	100 720 Undecane
7.353	1200.0	100 800 Dodecane
8.470	1300.0	100 801 Tridecane
9.437	1400.0	100 815 Tetradecane
10.313	1500.0	100 827 Pentadecane
11.125	1600.0	100 814 Hexadecane
11.886	1700.0	100 816 Heptadecane
12.607	1800.0	100 797 Octadecane
13.292	1900.0	100 789 Nonadecane
13.946	2000.0	100 792 Eicosane
14.569	2100.0	100 798 Heneicosane
15.167	2200.0	100 765 Docosane
15.741	2300.0	100 774 Tricosane
16.291	2400.0	100 746 Tetracosane
16.822	2500.0	100 732 Pentacosane
17.332	2600.0	100 735 Hexacosane
17.824	2700.0	100 706 Heptacosane
18.302	2800.0	100 682 Octacosane
18.763	2900.0	100 645 Nonacosane
19.272	3000.0	100 591 Triacontane
19.875	3100.0	100 545 Hentriacontane
20.607	3200.0	100 502 Dotriacontane
21.502	3300.0	100 447 Tritriacontane
22.607	3400.0	100 402 Tetratriacontane
23.979	3500.0	99 336 Pentatriacontane
25.687	3600.0	99 297 Hexatriacontane
27.781	3700.0	98 233 Heptatriacontane
30.426	3800.0	97 207 Octatriacontane

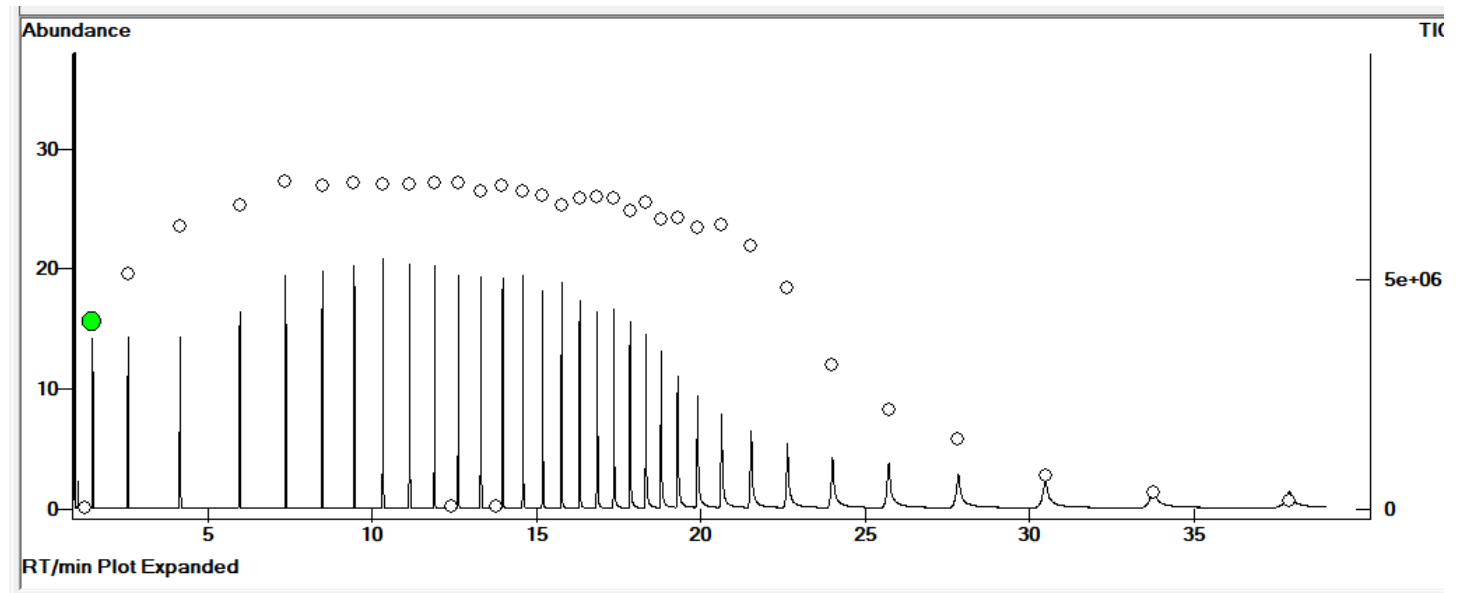
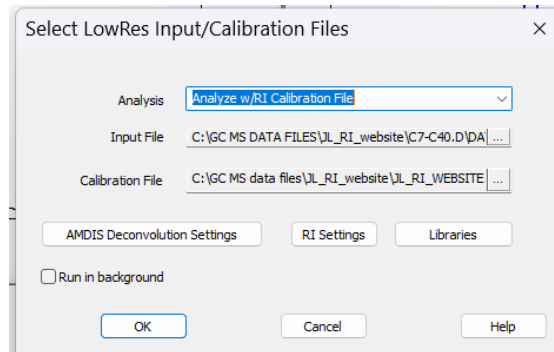
RI 3900 and 4000 missing! ➡

Approach to Determine Retention Time of Missing Standards

- Run calibration obtained vs original file
- Can expand trace to remove low boilers
- Filter with Max2Med seen when Max2Med to remove small peaks but retain C39 and C40
- If multiple peaks marked for C39 and C40
- Set Merge Duplicate = all
- Get retention times for C39 and C40 peaks from list
- Need peaks in sample for RI measurement to be bracketed by hydrocarbon peaks in mix
- Thus, don't necessarily need to calibrate to C40

**Max2Med* – Maximum abundance divided by median abundance of query spectrum

Rule of thumb, library search results tend to be less reliable when Max2Med approaches 10



Manually Edit Calibration Standard Adding RI and Retention Time

77	33.7033	307	3924	24	0.301
78	33.7454	876	3925	25	1.37
79	37.8607	742	4081	-49	0.673

0.675 600.0 96 1766 n-Hexane
 0.934 700.0 100 1438 Heptane
 1.485 800.0 100 682 Octane
 2.561 900.0 100 687 Nonane
 4.145 1000.0 100 682 Decane
 5.957 1100.0 100 720 Undecane
 7.353 1200.0 100 800 Dodecane
 8.470 1300.0 100 801 Tridecane
 9.437 1400.0 100 815 Tetradecane
 10.313 1500.0 100 827 Pentadecane
 11.125 1600.0 100 814 Hexadecane
 11.886 1700.0 100 816 Heptadecane
 12.607 1800.0 100 797 Octadecane
 13.292 1900.0 100 789 Nonadecane
 13.946 2000.0 100 792 Eicosane
 14.569 2100.0 100 798 Heneicosane
 15.167 2200.0 100 765 Docosane
 15.741 2300.0 100 774 Tricosane
 16.291 2400.0 100 746 Tetracosane
 16.822 2500.0 100 732 Pentacosane
 17.332 2600.0 100 735 Hexacosane
 17.824 2700.0 100 706 Heptacosane
 18.302 2800.0 100 682 Octacosane
 18.763 2900.0 100 645 Nonacosane
 19.272 3000.0 100 591 Triacontane
 19.875 3100.0 100 545 Hentriacontane
 20.607 3200.0 100 502 Dotriacontane
 21.502 3300.0 100 447 Tritriacontane
 22.607 3400.0 100 402 Tetratriacontane
 23.979 3500.0 99 336 Pentatriacontane
 25.687 3600.0 99 297 Hexatriacontane
 27.781 3700.0 98 233 Heptatriacontane
 30.426 3800.0 97 207 Octatriacontane

0.675 600.0 96 1766 n-Hexane
 0.934 700.0 100 1438 Heptane
 1.485 800.0 100 682 Octane
 2.561 900.0 100 687 Nonane
 4.145 1000.0 100 682 Decane
 5.957 1100.0 100 720 Undecane
 7.353 1200.0 100 800 Dodecane
 8.470 1300.0 100 801 Tridecane
 9.437 1400.0 100 815 Tetradecane
 10.313 1500.0 100 827 Pentadecane
 11.125 1600.0 100 814 Hexadecane
 11.886 1700.0 100 816 Heptadecane
 12.607 1800.0 100 797 Octadecane
 13.292 1900.0 100 789 Nonadecane
 13.946 2000.0 100 792 Eicosane
 14.569 2100.0 100 798 Heneicosane
 15.167 2200.0 100 765 Docosane
 15.741 2300.0 100 774 Tricosane
 16.291 2400.0 100 746 Tetracosane
 16.822 2500.0 100 732 Pentacosane
 17.332 2600.0 100 735 Hexacosane
 17.824 2700.0 100 706 Heptacosane
 18.302 2800.0 100 682 Octacosane
 18.763 2900.0 100 645 Nonacosane
 19.272 3000.0 100 591 Triacontane
 19.875 3100.0 100 545 Hentriacontane
 20.607 3200.0 100 502 Dotriacontane
 21.502 3300.0 100 447 Tritriacontane
 22.607 3400.0 100 402 Tetratriacontane
 23.979 3500.0 99 336 Pentatriacontane
 25.687 3600.0 99 297 Hexatriacontane
 27.781 3700.0 98 233 Heptatriacontane
 30.426 3800.0 97 207 Octatriacontane

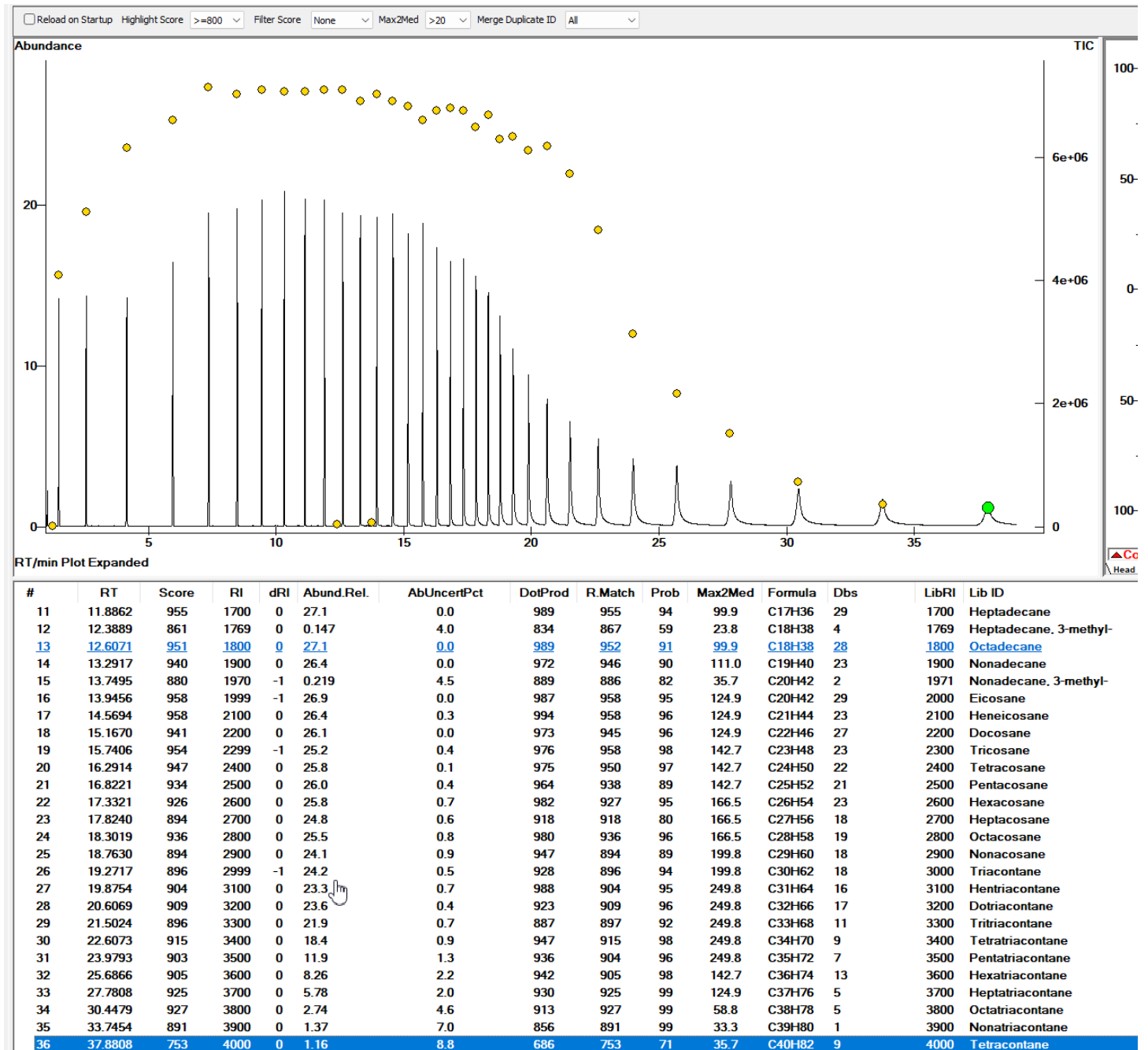
- Open the cal file in Notepad
- In this case, JL_RI_WEBSITE.CAL
- Retention times in first column for C39 and C40
- Add RI index, 3900 and 4000 in the second column
- **Must** be **exactly one space** between these two columns
- No information needed in columns 3 and 4
- Save or Save as with different name

RI 3900 and 4000 with
 Ret. Time Added

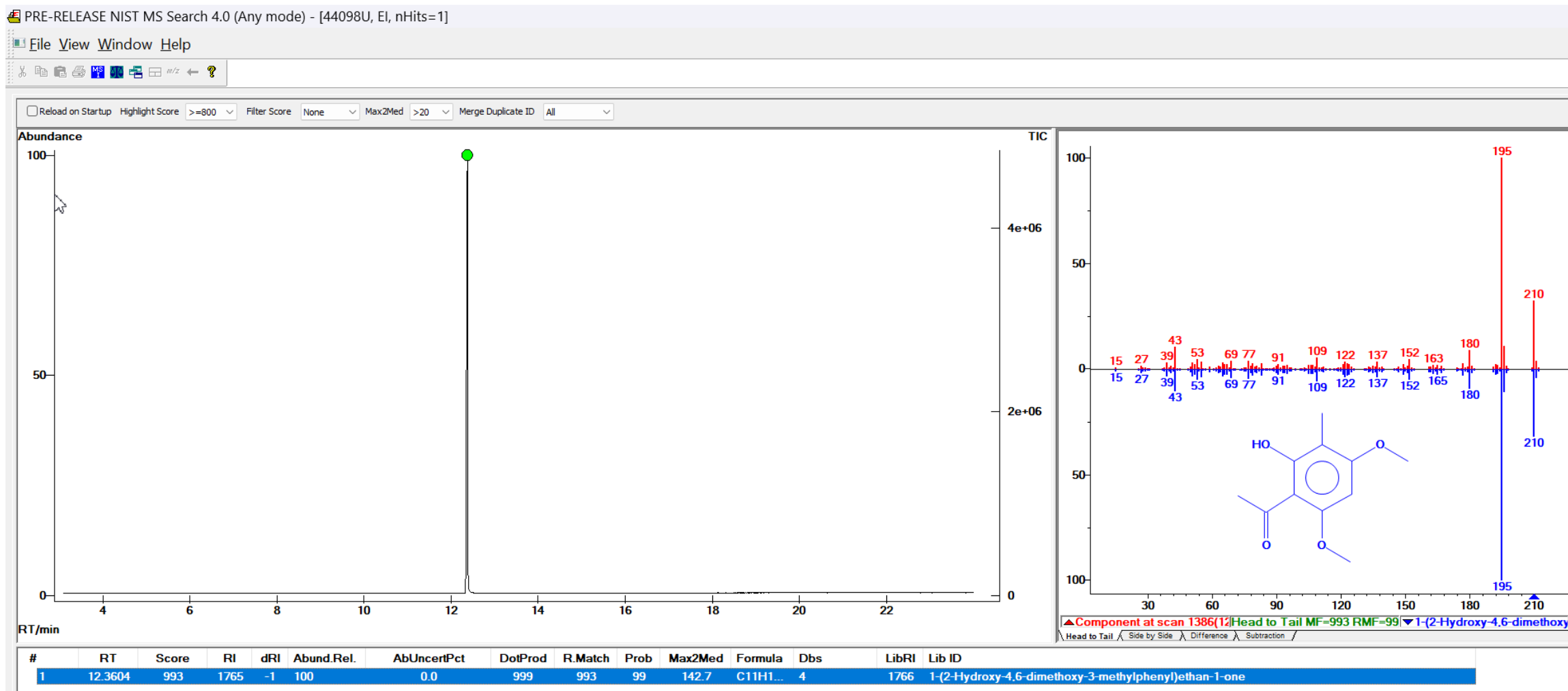


33.745 3900
 37.861 4000

- Checked again with peaks added manually
- Put Merge Duplicate on all
- Max2Med on smallest number to see all the hydrocarbons through C40 (RI 4000)
- Two high boilers were marked twice in the table
- Merge duplicate on all, just shows the top of peak
- Looks good to use now

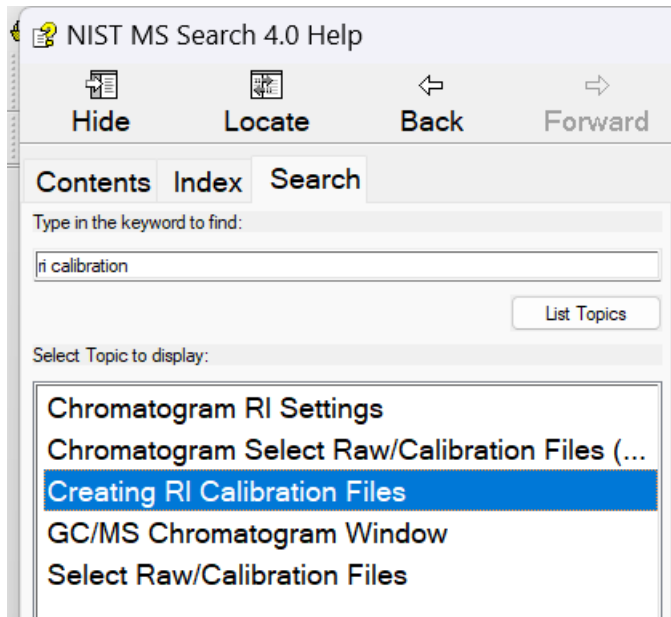


- Now Run Against a **Sample Under Same Conditions** of Hydrocarbon Standard and **Same Folder**
- Note found in NIST library, **dRI** between Current Calibration and RI in NIST Library **-1**
- Often can differ by as much as ~12 for user determination vs. value in current library
- **Remember**, data file and calibration *on my website* if you want **to try to reproduce**



NIST 10-step Process for RI Calibration on Next Few Pages (1 of 9)

- NIST has instructions for this task, locate in help using “Search” tab and “Creating RI Calibration Files”
- See their “10-step” method
- Index tab shows a list of pages
- Search finds text



Creating RI Calibration Files

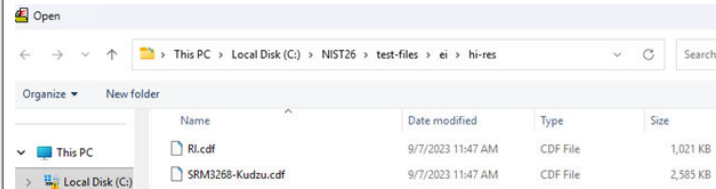
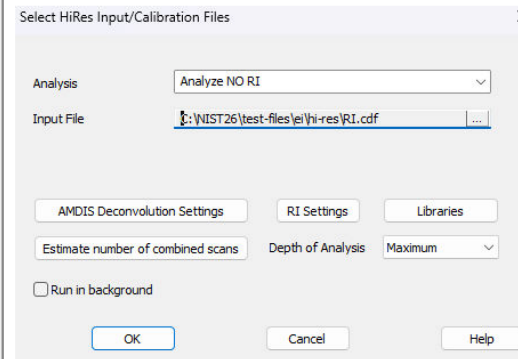
For confident compound identification by GC/MS, known retention indices (RI) must match reference values. Comprehensive Kovats index indices for *Se* under the same conditions at the experiment of interest. This is contained in a Calibration file the must be associated with each RI calibrated run.

Calibration tables are created from the analyses of mixtures of *n*-paraffins. These tables are then used to calibrate runs done under the same analytical conditions. Calibration files end in ".cal".

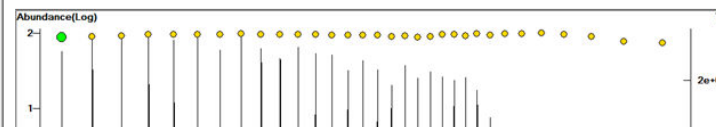
A method for creating such a Calibration file is illustrated below. It uses files provided with this software in folder /test-files/ei in the folder containing the N

While this can be done automatically using the Generate Calibration File, this can be inaccurate for high mass paraffins if they lack molecular ions. A more accurate method which examples are given and described later.

1. Analyze the with the Analyze NO RI setting the high-resolution *n*-paraffin file RI.cdf against the NIST EI library. This is located under in the base folder i



It contains C8 to C40 *n*-paraffins. Here is the resulting chromatogram – note with Abundance is in log units.



NIST 10-step Process on Next Few Pages (2 of 9)

Creating RI Calibration Files

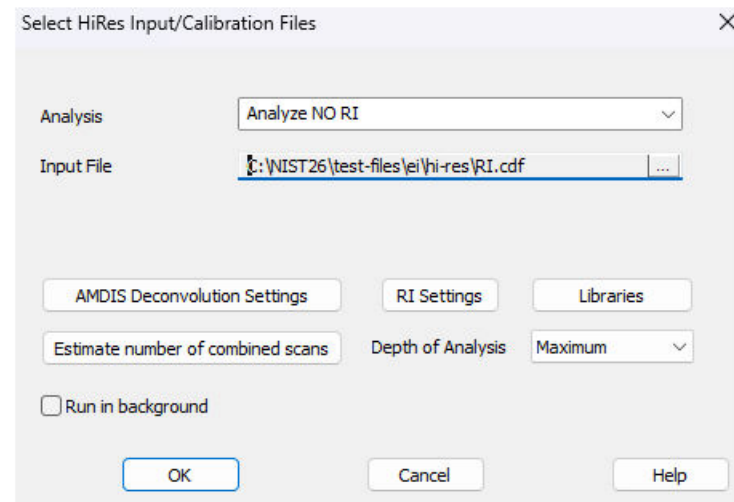
For confident compound identification by GC/MS, known retention indices (RI) must match reference values. Comprehensive Kovats index indices for *Semi-Specific Non-Polar* columns are provided with all spectra in the NIST library. To assign RI values to identifications requires a table that connects retention times with known RI values acquired under the same conditions at the experiment of interest. This is contained in a Calibration file the must be associated with each RI calibrated run.

Calibration tables are created from the analyses of mixtures of *n*-paraffins. These tables are then used to calibrate runs done under the same analytical conditions. These tables simply contain 2 columns – retention times in minutes and corresponding retention indices in Kovats retention units run on *Semi-Standard Non-Polar* columns. Names of these Calibration files end in “.cdf”.

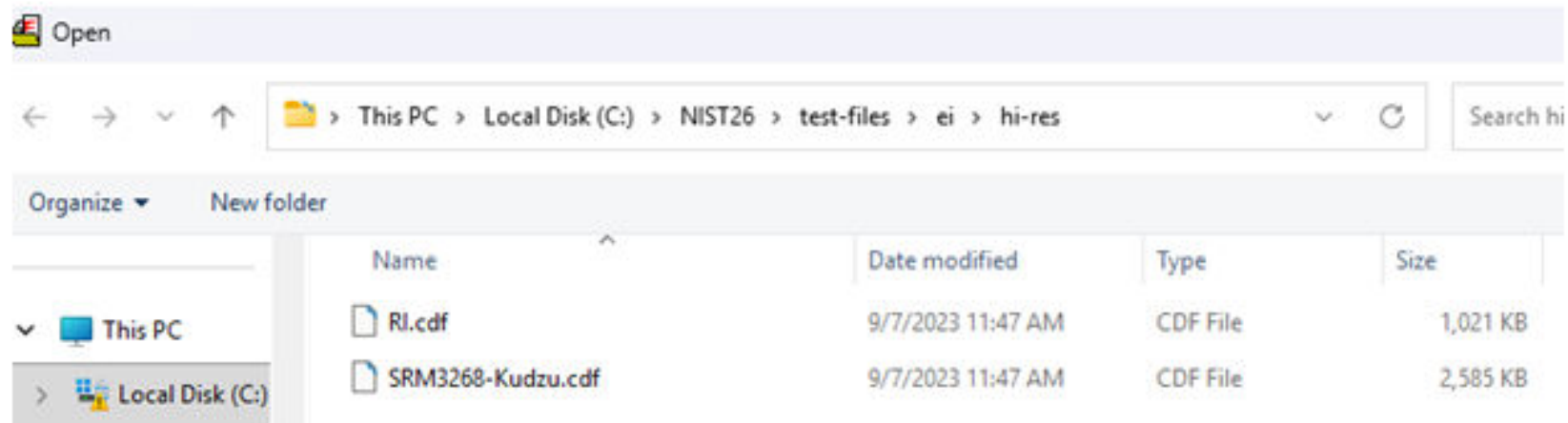
A method for creating such a Calibration file is illustrated below. It uses files provided with this software in folder /test-files/ei in the folder containing the NISTMS program, nistms.exe. Files are in “.cdf” format.

While this can be done automatically using the *Generate Calibration File*, this can be inaccurate for high mass paraffins if they lack molecular ions A more reliable method is now described. It is done through the following 10 steps. This detailed example uses a high-resolution run, which is virtually identical to the method for low resolution runs, for which examples are given and described later.

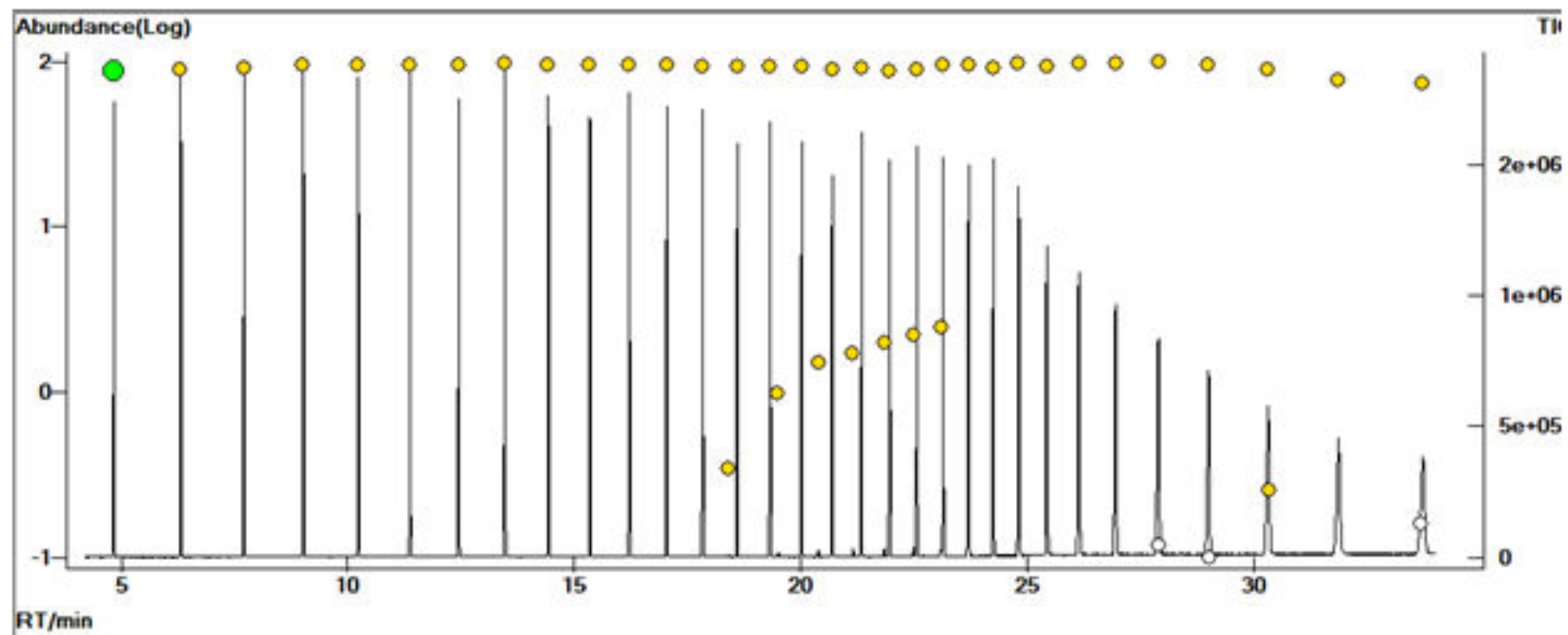
1. Analyze the with the *Analyze NO RI* setting the high-resolution n-paraffin file *RI.cdf* against the NIST EI library This is located under in the base folder in /test-files/ei/high-res/hi-res and selected as follows.



NIST 10-step Process on Next Few Pages (3 of 9)

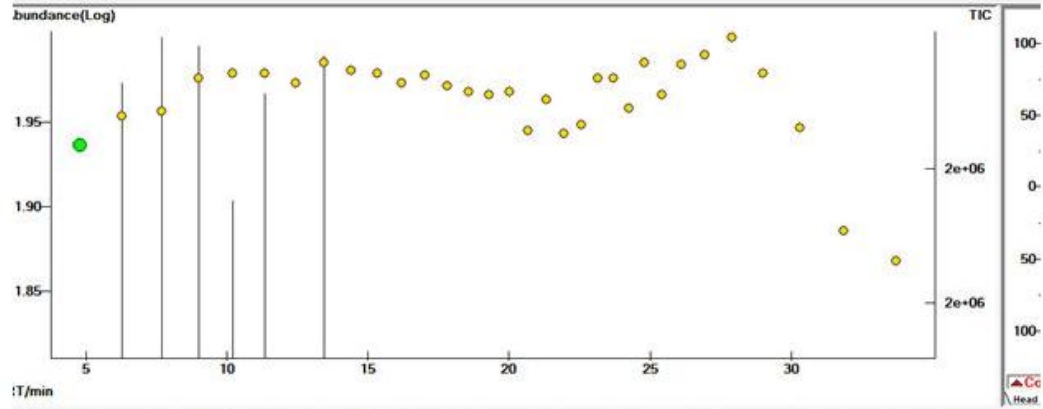


It contains C8 to C40 n-paraffins. Here is the resulting chromatogram – note with Abundance is in log units.



2. Choose standards by dragging mouse (actual IDs don't matter)

NIST 10-step Process on Next Few Pages (4 of 9)



#	S...	RT	Score	Pr...	Sc...	D...	R.M...	Abun...	AbUnc...	#	Pur...	Su...	LibRI	Lib	Formula	D...	Lib ID
1	94	4.8229	963	59	963	997	963	86.3	0.1	2	97.0	3.57	900	mainlib	C9H20	32	Nonane
2	315	6.2945	963	39	963	984	963	89.7	0.0	3	98.0	3.90	1000	replib	C10H22	32	Decane
3	525	7.6935	947	32	947	974	957	90.4	0.1	2	98.0	3.84	1100	replib	C11H24	30	Undecane
4	721	8.9992	948	29	948	970	948	94.5	0.1	3	98.0	3.85	1200	mainlib	C12H26	31	Dodecane
5	904	10.2176	939	20	939	986	939	95.2	0.1	3	98.0	3.88	1300	mainlib	C13H28	31	Tridecane
6	1...	11.3599	946	20	946	971	946	95.2	0.1	4	98.0	3.92	1400	replib	C14H30	33	Tetradecane
7	1...	12.4357	930	14	930	973	933	93.8	0.1	4	95.0	3.95	1600	replib	C16H34	34	Hexadecane
8	1...	13.4514	941	17	941	965	941	96.5	0.2	4	98.0	3.97	1600	mainlib	C16H34	34	Hexadecane
9	1...	14.4139	945	20	945	950	945	95.5	0.1	4	98.0	4.02	1700	replib	C17H36	29	Heptadecane
10	1...	15.3283	930	12	930	949	932	95.2	0.2	5	98.0	4.07	2000	replib	C20H42	29	Eicosane
11	1...	16.1988	933	12	933	955	933	93.8	0.3	5	97.0	4.13	2000	replib	C20H42	29	Eicosane
12	1...	17.0292	940	16	940	943	940	94.8	0.0	5	98.0	4.17	2100	replib	C21H44	23	Heneicosane
13	2...	17.8223	937	13	937	950	937	93.4	0.2	5	98.0	4.23	2100	replib	C21H44	23	Heneicosane
14	2...	18.5808	935	14	935	948	935	92.8	0.1	6	98.0	4.29	2400	replib	C24H50	22	Tetracosane
15	2...	19.3093	940	15	940	976	940	92.4	0.0	5	97.0	4.32	2700	replib	C27H56	18	Heptacosane
16	2...	20.0085	935	13	935	964	935	92.8	0.2	6	98.0	4.39	2100	replib	C21H44	23	Heneicosane
17	2...	20.6783	937	15	937	971	937	88.0	0.1	5	98.0	4.41	2700	replib	C27H56	18	Heptacosane
18	2...	21.3255	941	17	941	974	941	91.7	0.2	5	97.0	4.46	2700	replib	C27H56	18	Heptacosane
19	2...	21.9487	944	24	944	978	944	87.6	0.2	6	96.0	4.55	2700	replib	C27H56	18	Heptacosane
20	2...	22.5492	939	19	939	972	939	88.6	0.3	6	95.0	4.57	2700	replib	C27H56	18	Heptacosane

3. Highlight (ctrl-A) and copy full hit list to the Clipboard

4. Paste in Excel©

NIST 10-step Process on Next Few Pages (5 of 9)

5. Insert empty column (D below) after the RT column

6. Enter **RI** as first entry in empty column (900 in this case) – can verify accuracy of entries by examining successive **RI** differences between paraffins – should be uniform

7. Add 100 to each element of empty column and drag down (assuming all n-Paraffins in standard mix) – can modify depending on mix

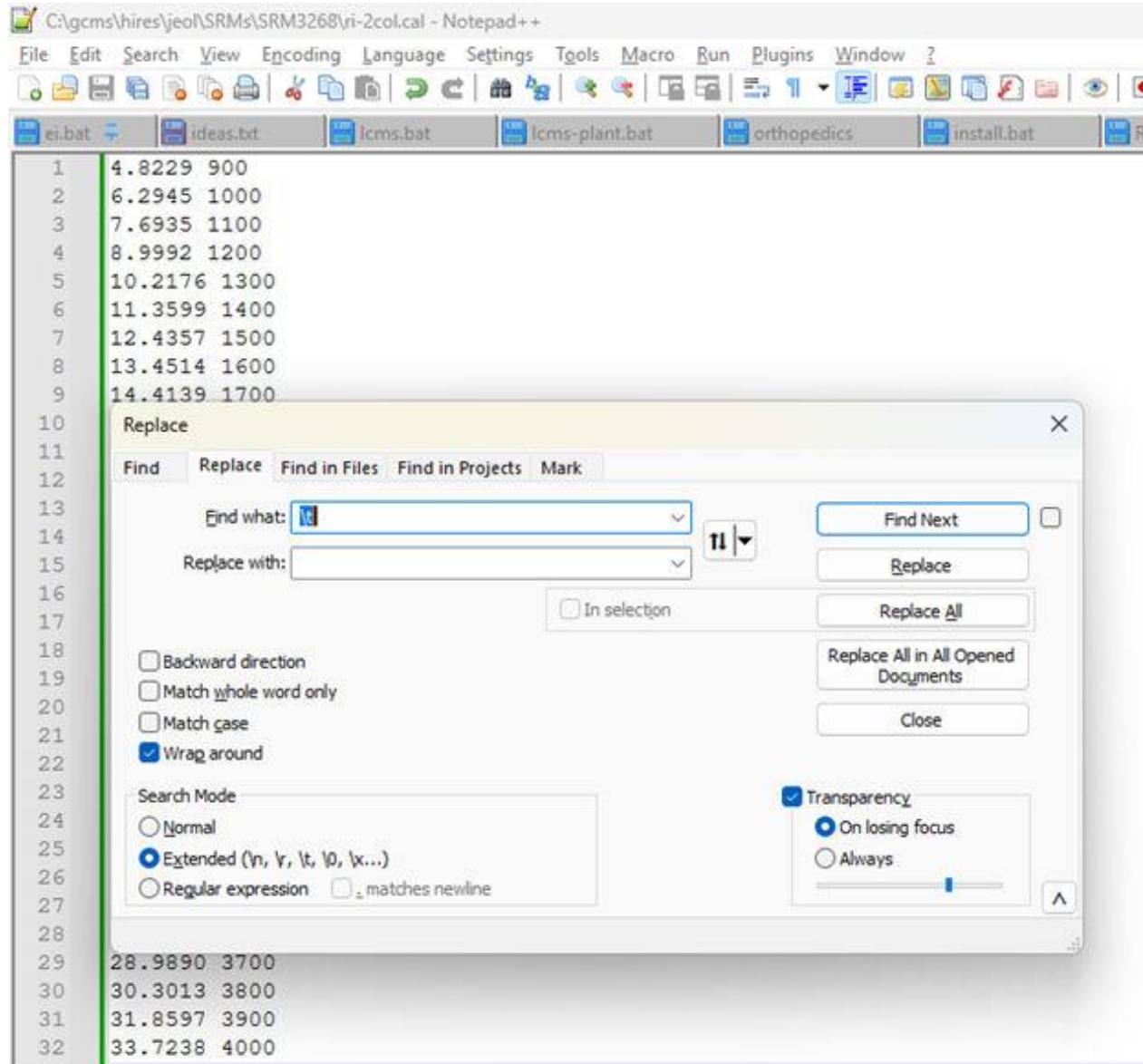
A	B	C	D
#	Scan	RT	RI
1	94	4.8229	900
2	315	6.2945	1000
3	525	7.6935	1100
4	721	8.9992	1200
5	904	10.2176	1300
6	1075	11.3599	1400
7	1236	12.4357	1500
8	1389	13.4514	1600
9	1533	14.4139	1700
10	1670	15.3283	1800
11	1801	16.1988	1900
12	1926	17.0292	2000
13	2045	17.8223	2100
14	2158	18.5808	2200
15	2268	19.3093	2300
16	2373	20.0085	2400
17	2473	20.6783	2500
18	2570	21.3255	2600
19	2664	21.9487	2700
20	2754	22.5492	2800
21	2841	23.1297	2900
22	2925	23.6917	3000
23	3007	24.2355	3100
24	3091	24.7967	3200
25	3184	25.4192	3300
26	3290	26.1223	3400
27	3412	26.9342	3500
28	3554	27.8798	3600
29	3720	28.9890	3700
30	3917	30.3013	3800
31	4151	31.8597	3900
32	4430	33.7238	4000



8. Highlight and copy 2 columns above to Notepad++ (freely available text editor)

NIST 10-step Process on Next Few Pages (6 of 9)

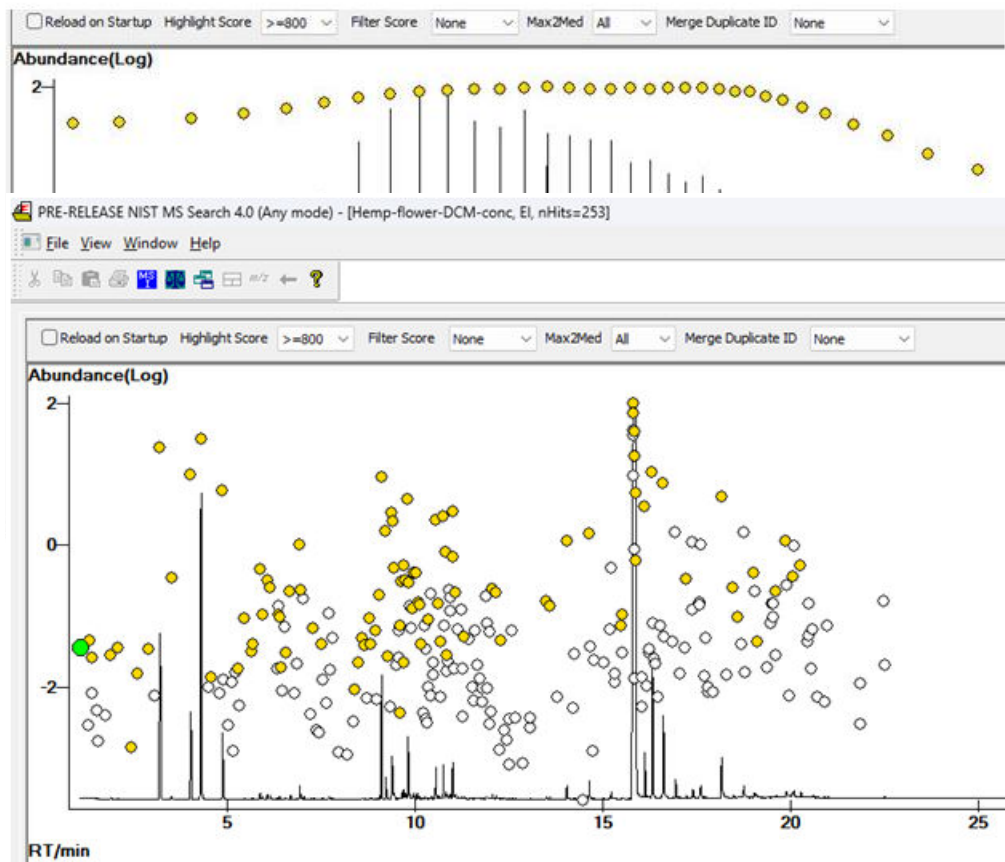
9. Change tabs (t) to spaces (begin by first pressing *ctrl-F* and then replace)



The screenshot shows the Notepad++ application window with a Replace dialog box open. The dialog box is titled "Replace" and has tabs for "Find", "Replace", "Find in Files", "Find in Projects", and "Mark". The "Replace" tab is active. The "Find what:" field contains the character 't'. The "Replace with:" field is empty. The "In selection" checkbox is unchecked. The "Backward direction" checkbox is unchecked. The "Match whole word only" checkbox is unchecked. The "Match case" checkbox is unchecked. The "Wrap around" checkbox is checked. The "Search Mode" section has three radio buttons: "Normal" (unchecked), "Extended (n, \, !, @, |, x...)" (checked), and "Regular expression" (unchecked). The "Matches newline" checkbox is unchecked. The "Transparency" section has two radio buttons: "On losing focus" (checked) and "Always" (unchecked). A transparency slider is visible below the "Always" radio button. The background text in the Notepad++ window is as follows:

```
1 4.8229 900
2 6.2945 1000
3 7.6935 1100
4 8.9992 1200
5 10.2176 1300
6 11.3599 1400
7 12.4357 1500
8 13.4514 1600
9 14.4139 1700
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29 28.9890 3700
30 30.3013 3800
31 31.8597 3900
32 33.7238 4000
```

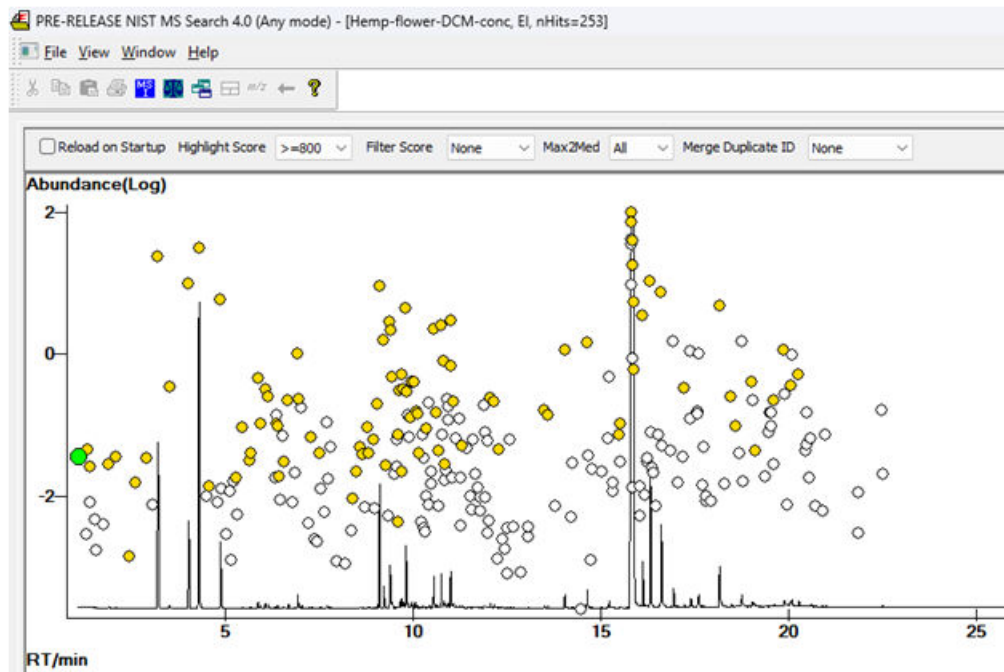

Equivalent low-resolution analysis processing can be done using **calibration** data file *C7-C40.D* in *test-files\ei\low-res*. Then use the derived **calibration** file to run data file *Hemp-flower-DCM-conc.D* (in *.D* format). Chromatograms illustrated below.



NIST 10-step Process on Next Few Pages (8 of 9)

Using this general approach, you can create **Calibration** files from any analysis containing a sufficient number of confident identifications across the entire retention range of interest. This involves selecting choices for reference **RI** identifications and reported library **RI** values. Selections and deletions can be made at any stage in the method – though one must be very careful.

NIST 10-step Process on Next Few Pages (9 of 9)



Using this general approach, you can create **Calibration** files from any analysis containing a sufficient number of confident identifications across the entire retention range of interest. This involves selecting choices for reference **RI** identifications and reported library **RI** values. Selections and deletions can be made at any stage in the method – though one must be very careful.

Chat GPT Suggestions for Purchasing Test Mix (1 of 3)

Yes — and in fact, what you're describing is essentially **exactly what Sigma (MilliporeSigma /Supelco) sells.**

Commercial C7–C40 RI calibration mixes (Sigma/Supelco)

The standard most people (including NIST workflows) use is:

•C7–C40 Saturated Alkane Standard (Supelco 49452-U)


- Full homologous series: **n-C7 → n-C40**
- Typically ~1000 µg/mL each component in hexane
- Sold as a **certified reference material (CRM)**
- Explicitly marketed as a **retention index marker probe**

👉 This is essentially the **Kovats RI ladder** in ready-to-use form.

My ChatGPT Prompt

nist uses a mix of hydrocarbons C7-C40 to do their RI calibration. Does sigma supply such a commercial mix?

https://www.sigmaaldrich.com/US/en/product/supelco/49452u?srsId=AfmBOopEvdSGmrNijfGjN9Ia_Uv5IFw8HObgiGyMEqp_D9BRr1KVTAdu&utm_source=chatgpt.com

49452-U  Supelco.

C7 - C40 Saturated Alkanes Standard

certified reference material, 1000 µg/mL each component in hexane

★★★★★ (0) [Write a review](#) [Ask a question](#)

Synonym(s):Saturated Alkane Calibration Mix, C7 - C40

Chat GPT Suggestions for Purchasing Test Mix (2 of 3)

What's actually in it

It contains the full normal alkane ladder:

- Heptane (C7), octane (C8), ... all the way to tetracontane (C40)
- Each at equal concentration for consistent RI interpolation

That's exactly the same concept NIST uses for RI calibration.

Variants you'll also see from Sigma

Depending on your workflow, Sigma offers several related mixes:

- **C7–C30 alkane mix** (shorter range)
- **C8–C40 calibration mix** (slightly shifted start)
- **Even-carbon mixes (e.g., C10–C40)**
 - cheaper, but less precise for interpolation
- **Performance test mixtures (ISO/DIN methods)**
 - not always full ladders, sometimes sparse coverage

ChatGPT Suggestions for Purchasing Test Mix (3 of 3)

Practical note (based on your NIST work)

For what you're doing (NIST RI calibration / GC-MS identification):

- The **full C7–C40 ladder** is the right choice
- Gives you:
 - consistent **linear interpolation across entire chromatogram**
 - compatibility with **NIST RI libraries (which assume full n-alkane series)**

Using even-only mixtures (C10, C12, ...) will work but degrades RI precision—especially for:

- early eluters (<C12)
- isomer discrimination (which you're clearly emphasizing on your site)

Bottom line

Yes — Sigma absolutely supplies the exact type of mixture NIST uses:

- ✓ Full homologous series
- ✓ Certified reference material
- ✓ Designed specifically for RI calibration

If you want, I can suggest a couple of **lower-cost or DIY alternatives** (Restek, AccuStandard, or making your own mix), or how to tailor the range (e.g., C7–C25 vs C10–C40) for your specific GC method.