

CMC-se

- Structure Elucidation
User Manual
Version 006



Copyright © by Bruker Corporation

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form, or by any means without the prior consent of the publisher. Product names used are trademarks or registered trademarks of their respective holders.

© October 15, 2019 Bruker Corporation

Document Number: 1000031933

P/N: H149082

Contents

1	Introduction	5
1.1	About CMC-se Version 2.7	5
1.2	Significant Changes in CMC-se 2.7	5
1.3	History of Changes	6
1.3.1	Major Changes Compared to CMC-se Version 2.4	6
1.3.2	Major Changes Compared to CMC-se Version 2.3	6
1.3.3	Major Changes Compared to CMC-se Version 2.2	6
1.3.4	Major Changes Compared to CMC-se Version 2.1	7
1.3.5	Major Changes Compared to CMC-se Version 2.0	7
1.3.6	Major Changes Compared to CMC-se Version 1.0	8
2	Getting Started	11
2.1	Program Prerequisites	11
2.2	Basic Program Features	13
3	Structure Elucidation Basics	15
3.1	NMR Spectra for a Structure Elucidation	15
3.1.1	Required Parameter Sets	15
3.1.2	Additional Supported Experiments	17
3.1.3	Data Set Processing	19
3.2	Starting CMC-se	21
3.2.1	Structure Elucidation Project Management	21
3.2.2	Creating a New CMC-se Project	23
3.3	The Correlation Table	33
3.4	Automatic Spectra Analysis	39
3.4.1	Tasks Accomplished by the Automatic Analysis Feature	41
3.4.2	Typical Problems with Automatic Multiplet Analysis	42
3.5	CSEARCH Structure Identification	45
3.6	The Combined Spectra Display Window	46
3.6.1	Manipulating Viewing Options of Combined Spectra Display	50
3.6.2	Actions Possible through the Combined Spectra Display	52
3.6.2.1	Modifying Existing Multiplets	52
3.6.2.2	Adding and Removing Multiplets Through the Combined Spectra Display	57
3.6.2.3	Handling of ^{19}F or $(^{31}\text{P})\text{-}^{13}\text{C}$ Splitting	66
3.6.2.4	Adding INADEQUATE Correlations	68
3.7	Additional Modifications to the Correlation Table	71
3.7.1	Deleting Atoms from the Correlation Table	72
3.7.2	Adding/Deleting Correlations through the Table	74
3.7.3	Disabling Correlations in the Correlation Table	74
3.7.4	Additional Manipulations to Correlations in Table	75
3.7.5	Adding Additional Information to the Correlation Table	76
3.8	Bookkeeping Tools within the CMC-se Module	78
3.9	The Fragment Tool	81
3.10	Setup of Structure Generation Options	88

3.10.1	Structure Generator	90
3.10.2	Execution Control.....	90
3.10.3	Substructures / Ring Rules	91
3.10.4	Correlations.....	92
3.10.5	Chemistry Rules.....	95
3.11	Structure Generation.....	96
3.11.1	Possible Outcomes of Structure Generation.....	99
3.12	User Validation of Generated Structure Proposals	101
3.12.1	Viewing Generated Structures	101
3.12.2	Structure Rating via Chemical Shift Prediction	107
3.12.3	Viewing Correlations on Structures.....	110
3.12.4	Check Assignment for Publication (CSEARCH Robot Referee).....	119
3.13	Process Overview	120
4	The CMC-se Structure Verification Tool.....	125
4.1	Verification of the Structure File on the Disc.....	125
4.2	Verification of Partially Assigned Structure	127
4.3	Verification Report	129
5	Examples	131
5.1	Structure Elucidation for α -Ionone	131
5.1.1	Manual Structure Elucidation of α -Ionone using the Fragment Tool.....	143
5.2	Structure Elucidation for Guaiol	163
5.3	Structure Elucidation for Strychnine.....	180
6	Contact	201
	List of Figures.....	203

1 Introduction

1.1 About CMC-se Version 2.7

Bruker's Small Molecule Structure Elucidation module (CMC-se) is a TopSpin Plug-In first published in TopSpin 3.1.

CMC-se is a part of Topspin, however it requires an additional license. There are two different licenses:

- Full CMC-se license – provides the complete functionality
- CMC-se Classroom edition license – does not include the structure generator.

The CMC-se Classroom edition license is included in the free Topspin academic license.

This documentation describes FULL CMC-se Version 2.7 that is a part of TopSpin 4.0.8.

The full version described here includes automatic structure generation algorithms, as well as the verification routine, which are not included in the Classroom Edition. The available license determines which features will be available to the user. The Examples included at the end of this manual are specifically for the full version of CMC-se, and not the Classroom Edition.

The rich supporting material and Classroom specific training datasets may be downloaded here: <https://www.bruker.com/service/support-upgrades/software-downloads/nmr.html>

CMC-se is based on an automation approach a set of analyzed NMR and the molecular formula are used to determine the 2-Dimensional structure of the investigated sample. Please note that no stereochemistry, including cis vs. trans, information is derived from the data, and structures displayed are only one possible version of the molecular connectivity.

CMC-se implements a correlation based approach that uses only the atom – atom correlations derived from different 2D spectra to assemble the atoms into compounds that are consistent with the data, rather than relying heavily on a database of known fragments.

Chemical shift prediction is automatically performed, and can be used as an initial ranking system to determine the most reasonable structures.

In addition to the tools provided to view these correlations on the generated molecules, which verify the correct structure, there are also bookkeeping tools built into CMC-se to assist the user in working with the data manually.

1.2 Significant Changes in CMC-se 2.7

- NMRReData implementation

NMRReData (www.nmredata.org) is a format for the publishing of NMR signal assignment for small molecules.

CMC-se assignment of NMR signals to atoms in a elucidated structure may be exported in NMRReData format. NMRReData may imported into CMC-se project.

- HOESY spectra were added to the list of supported datasets (display only).
- List of possible charged atom states was updated (e.g. perboric acid)

1.3 History of Changes

1.3.1 Major Changes Compared to CMC-se Version 2.4

CMC-se 2.5 is a new minor release with the following new key features:

3D Structure Analysis with ROESY and NOESY

NOESY and ROESY spectra were added to the list of supported experiments.

This allows the manual estimation of relative stereochemistry in CMC-se.

Correlation Highlighting in Structure Editor

The highlighting functionality discriminates different experiment types and its information content. Correlation highlighting can separate between COSY and HMBC for valid and invalid correlations. As an additional improvement, correlations can be shown between individual drawn protons, especially for correlations from ROESY and NOESY.

Interface to CSEARCH Robot Referee (Structure Assignment Check)

Get an assignment check ahead a publication and ready for the transfer in the manuscript. The result is an email holding URLs to details about the evaluation, the spectrum prediction and the classification of the assigned structure.

Read more information in the following link:

<http://nmrpredict.orc.univie.ac.at/c13robot/robot.php>

1.3.2 Major Changes Compared to CMC-se Version 2.3

CMC-se version 2.4 is a minor update to CMC-se 2.3. Version 2.4 was a new release with the following new key features:

CSEARCH database search

- As a new feature, CMC-se offers a CSEARCH database request to match all possible structures based on C13 chemical shifts search. The request uses the C13 chemical shifts with additional information about hybridization, the chemical formula and the mass after the automatic analysis. The request will be automatically generated and the result is given via email. For more information see [CSEARCH Structure Identification \[▶ 45\]](#)

1.3.3 Major Changes Compared to CMC-se Version 2.2

Version 2.3 was a new release of Bruker structure elucidation suite that brought the following new key features:

Data Storage

- The CMC-se project file storage changed, visible with the new data format *.xcmap* instead of the old ending *.cmap*. The new format stored all analyzed data information (peak and multiplet lists) in the project file. With this, it is possible to have several CMC-se projects working on the same set of NMR spectra.
- The old (CMC-se 2.2 and older) projects are automatically converted to the new format.

Fragment Editor

- Major improvements have been made in the fragment tool ([The Fragment Tool \[▶ 81\]](#)).
- With the combination of the general fragment editor and the CMC-se assignment tool the spectroscopist gets a powerful tool to work manually with the data.
- The new tool allows not only complex definition of fragments for the structure generation, but also draw and assign the whole structure from scratch. This means, that the new CMC-se version may be used as a **structure workbook**.

1.3.4 Major Changes Compared to CMC-se Version 2.1

CMC-se 2.2 is a new release of Bruker structure elucidation suite. It brings following new key features:

Data analysis

- Added ability to “merge split multiplets” for simplified handling of ^{19}F or ^{31}P coupling.
- Added ability to “remove all content” from the correlation table so that the user can start fresh.

Structure generation

- Improved handling of Charged Structures, including native salts, as well as additional charged states from exchangeable protons from the solvent.

Structure Ranking

- The Modgraph prediction routine has been removed as it is no longer necessary. The internal prediction routine that runs automatically during the structure generation process is faster, and produces reliable results without the need for user intervention.

1.3.5 Major Changes Compared to CMC-se Version 2.0

CMC-se 2.1 is a new release of Bruker structure elucidation suite. It brings following new key features:

Automated Data Analysis

- Additional experiments may be easily replaced or added to a fully analyzed CMC-se project. The following analysis of the new data adds the respective correlations to the correlation map. The original correlations are not modified.
- Also manual corrections, functional groups, assignment... stay as is.
- More reliable assignment of ambiguous (fuzzy) HMBC correlations.
- Improved handling of HMBC multiplet overlap.

Structure Generation

- Improved handling of underestimated problems. Presence of several unseen atoms, or atoms with few constraints may lead to huge number of structure proposals. CMC-se 2.1 uses a fast on the fly ranking to keep only the best structure candidates. Currently a fast C13 chemical shift prediction is used to rank generated structure proposals. This allows finding a solution in hopeless cases, where more than 10000 structure candidates are generated.
- General handling of charged atom states. Any functional groups containing charges (e.g. nitro or azide group) and inner salts are now supported.
- Additional correlation types (e.g. TOCSY) are implemented

Verification

- The verification function allows now to handle partially assigned structures.

Fragment Editor

- New built-in fragment editor allows to create or to modify substructure fragments directly from CMC-se window. The editor preserves the CMC-se atom assignments. It shares its user interface with the renewed TopSpin built in structure editor.

1.3.6 Major Changes Compared to CMC-se Version 1.0

In addition to minor bug corrections, and behind the scenes improvement in functionality, there are some relatively major changes and additions that have been incorporated into CMC-se Version 2.0.



One of the biggest improvements to CMC-se in this version is the expanded support of commonly used experiments for structure determination.

- In addition to the previously supported experiments (^1H , ^{13}C , ^1H - ^1H COSY, ^1H - ^{13}C HSQC/HMBC, ^1H - ^{15}N HSQC/HMBC), in this version (CMC-se 2.0), the following experiments are now also automatically analyzed, and when using the Bruker generation algorithm, the correlations used in the structure generation:
 - DEPTq as an alternative to a traditional 1D ^{13}C .
 - 1.1-ADEQUATE.
 - Selective-HMBC. Multiple HMBC spectra can now be used and automatically analyzed in the same project, the spectral regions may overlap.
 - ^1H - ^{13}C H2BC.
 - ^{19}F - ^{13}C HSQC/HMBC (Manual corrections are likely due to extra J coupling).
 - ^1H - ^{31}P - HMBC (Manual corrections are likely due to extra J coupling).
- The following experiments are now supported and used by the Bruker structure generation algorithm, but manual analysis is required:
 - INADEQUATE
 - ^1H - ^{19}F HETCOR
- The following experiments are now supported by CMC-se, but not used in the generation process. Because the correlations are not used in the generation process, they will not be shown in the Correlation Table. However as with any correlation, they can be visualized on the structure candidates for evaluation purposes (see section [Viewing Correlations on Structures \[▶ 110\]](#)):
 - ^1H - ^1H TOCSY.

- ^1H - ^1H NOESY or ROESY.
- ^1H - ^{13}C HSQC-TOCSY.

Because of the expanded support to additional experiments, there have been substantial changes to the look and feel of the Correlation Table, and Combined data display. These windows will be described in detail in the following sections.

- **Improvements have been made to the graphical utility of the fragment tool.** In addition to being able to define fragments for the molecule from known assignments, the experimental correlations can be visualized on these fragments. Not only do these assigned fragments improve the results of the automatic structure generation algorithms, but this is a useful tool for the spectroscopist working with the data manually. This tool is explained in the section [The Fragment Tool \[81\]](#).
- **Another addition to the CMC-se program is the verification tool.** While the main purpose of CMC-se is to generate structure candidates of for unknown molecules, it now can be used to help verify if a previously proposed structure is consistent with the input data (minimum required ^1H , HSQC, HMBC). This tool is described in the chapter [The CMC-se Structure Verification Tool \[125\]](#).

2 Getting Started

This section lists the basic prerequisites for the CMC-se module, and briefly describes the main features of the program. More detailed explanations will be provided in the chapter [Structure Elucidation Basics \[▶ 15\]](#).

2.1 Program Prerequisites

- The CMC-se module will run on any recent PC, but at least 2GB memory is necessary, and 4GB are recommended.
 - Linux, Windows and macOS platforms are all supported by CMC-se.
- In addition to the required TopSpin3 or TopSpin 4 license, a separate CMC-se license is also required to start the software module. As mentioned previously, there are 2 different types of licenses for CMC-se, a Full license (commercial), and a license for the Classroom Edition, which is included in the TopSpin license for academia.
 - If only the Classroom Edition license is available, the automatic generation features will not be accessible.
 - If a full license is available (with or without a Classroom Edition license), then all features of the program will be accessible.
- Unlike other analysis modules provided by Bruker (CMC-assist, Dynamics Center) CMC-se is included with the normal TopSpin install. It is highly recommended that one installs the “NMR-Data” as this will contain the example data that is used in the chapter [Examples \[▶ 131\]](#).

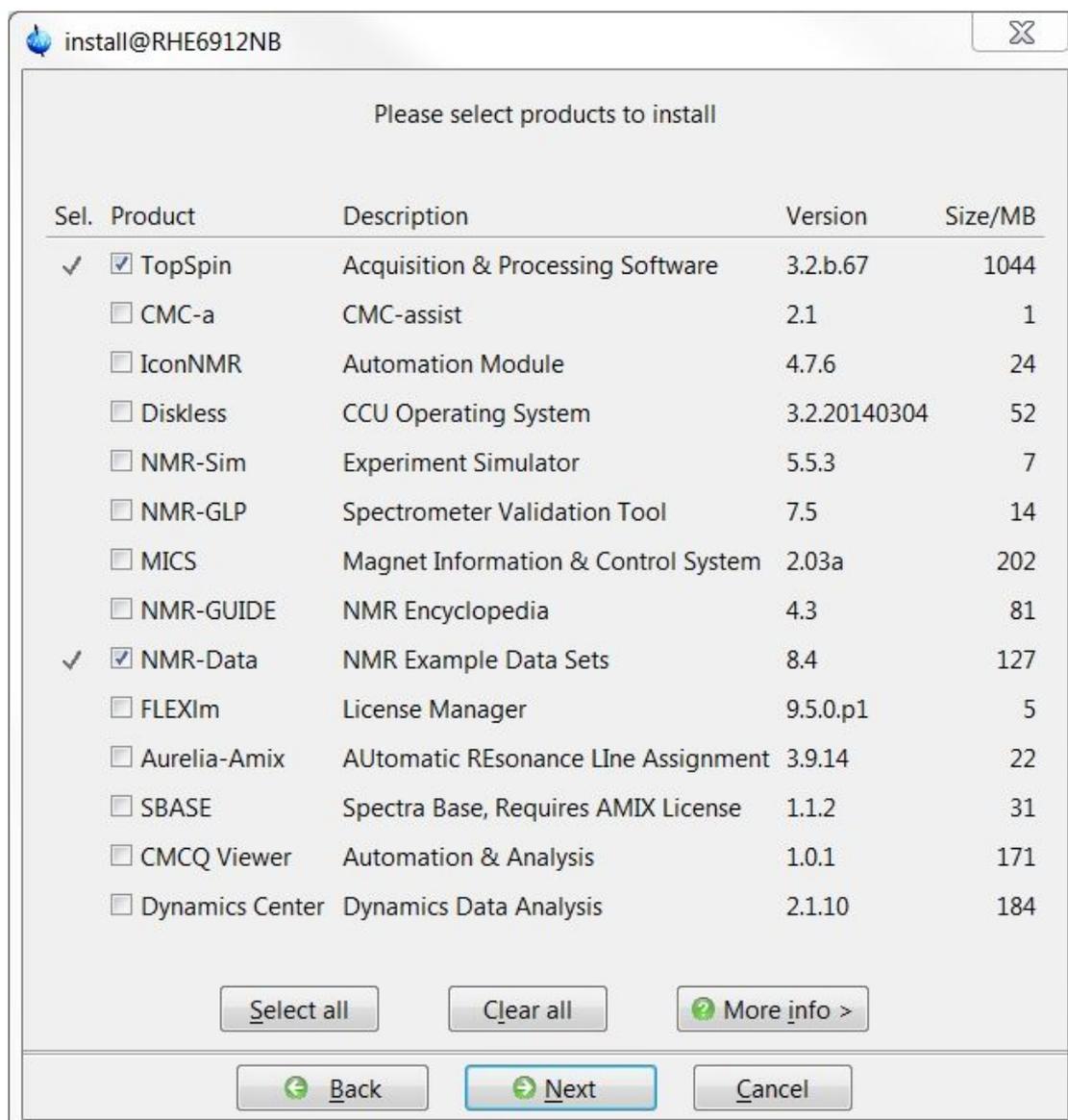


Figure 2.1: Custom Install window highlighting the minimum packages necessary for CMC-se installation.

- The Molecular Formula for the compound of interest must be known. The currently supported atoms are:
 - **H, D, C, N, O, S, F, Cl, Br, I, P, Si, K, Ca, Na, B**



The compounds containing F and P will likely require additional user interaction because the current version of the automated data analysis cannot interpret the possible long range couplings.

2.2 Basic Program Features

- Bruker's Small Molecule Structure Elucidation module (CMC-se) is based on the semi-automatic analysis of a set of NMR spectra (1D- ^1H , $^1\text{H}/^{13}\text{C}$ HSQC, $^1\text{H}/^{13}\text{C}$ HMBC, and, optionally 1D- ^{13}C , $^1\text{H}/^1\text{H}$ COSY, $^1\text{H}/^{15}\text{N}$ HSQC, $^1\text{H}/^{15}\text{N}$ HMBC, $^1\text{H}/^{13}\text{C}$ H2BC, $^1\text{H}/^{13}\text{C}$ ADEQUATE, ^{13}C -INADEQUATE) and the molecular formula, to generate structure candidates based on the correlations in the above data sets. Any additional information, e.g. symmetry, known fragments, might be helpful.
- The data should be acquired using Bruker's respective pulse program, and **parameter set suggestions** (*NMR Spectra for a Structure Elucidation* [▶ 15]), which are optimized for minimal artifacts, and to work best with the automatic analysis component of CMC-se. There are no special requirements for the remaining spectra which are not listed in this chapter.
- The **automatic analysis** tool (see *Automatic Spectra Analysis* [▶ 39]) performs a multiplet and a correlation analysis of the spectra. The result is represented as a table, referred to as the Correlation Table.
- The **Correlation Table** (sections *The Correlation Table* [▶ 33] and *Additional Modifications to the Correlation Table* [▶ 71]), while automatically created and populated by the automatic analysis, should be inspected carefully by the user before proceeding. It is interactive and can be corrected and supplemented. In addition to displaying the experimental correlations in table format, the user can add names and annotations to the Correlation Table. These additions are conveniently displayed in the Combined Spectra Display, thus additionally serving as a **book-keeping tool** for users working with data manually (section *Bookkeeping Tools within the CMC-se Module* [▶ 78]).
- The **Combined Spectra Display** (section *The Combined Spectra Display Window* [▶ 46]) is another CMC-se specific window in which all CMC-se data are displayed in overlay mode, assisting in data interpretation. The Combined Spectra Display is linked to the Correlation Table facilitating modification of the table. Manual peak picking on spectra is possible through this window, and changes are updated in the Correlation Table.
- The CMC-se module includes two **Structure Generation Algorithms** which vary in the handling of the input information (see section *Setup of Structure Generation Options* [▶ 88] for more details). Using the molecular formula and the Correlation Table, as well as adding additional restraint options (see *The Fragment Tool* [▶ 81]); both algorithms generate possible molecular structures consistent with the NMR constraints.
- Using the **Fragment Tool** (section *The Fragment Tool* [▶ 81]), users have the ability define fragments that are to be required or excluded from the generated structures. Atoms of these fragments can be assigned to a particular atom in the Correlation Table, giving the structure generation a building block from which to start from. Part of this tool also now includes extended graphical support for visualizing the experimental correlations on the fragments to assist in the assigning of atoms and connecting functional groups and fragments into full molecules if desired.
- ^{13}C **shift prediction** of generated structures and automatic comparison to experimental data provides an initial ranking system to help determine the most probable of the generated results (section *Structure Rating via Chemical Shift Prediction* [▶ 107]). In addition to a prediction rating for the molecule as a whole, each atom can be color coded according to how well it matched the predicted value. The possibility of providing a user ranking and comment on structures is also possible.
- Several additional tools for **Verifying the Generated Structures** are included (section *User Validation of Generated Structure Proposals* [▶ 101]). The generated structures are displayed in a multi-structure display panel. This panel is linked to the Correlation Table, and the experimental constraints can be highlighted on the generated structures assisting the user in validating the top structural candidates.

- In addition to these tools being useful for verifying the structures that are generated by the program, CMC-se can be used to verify a user proposed structure candidate without having to generate it through CMC-se first (chapter [The CMC-se Structure Verification Tool](#) [125]).

3 Structure Elucidation Basics

A complete structure elucidation with Bruker's small molecule structure elucidation module (CMC-se) is designed to be performed according to a special workflow, which is summarized in a flow chart at the end of this chapter (section [Process Overview \[120\]](#)). Details for each step of this workflow will be presented now.

In general, the steps are the following:

- A suitable set of NMR spectra has to be acquired and processed (section [NMR Spectra for a Structure Elucidation \[15\]](#)).
- A structure elucidation project must be created and the spectra and molecular formula should be defined (section [Structure Elucidation Project Management \[21\]](#)).
- An automatic analysis of the project spectra is performed, and the Correlation Table is populated (sections [Automatic Spectra Analysis \[39\]](#) and [The Correlation Table \[33\]](#)).
- While in many cases, the automatic analysis is sufficient and structures can be generated from this point, a careful inspection of the compiled Correlation Table for errors and inconsistencies is strongly recommended. The populated cells of the Correlation Table can be compared with the actual data, and corrected where necessary using the Combined Spectra Display (section [The Combined Spectra Display Window \[46\]](#)). If possible and desired, further supplemented with other additional information is possible (section [Additional Modifications to the Correlation Table \[71\]](#)).
- Prior to starting the structure generation algorithms, the desired options are set (section [Setup of Structure Generation Options \[88\]](#)), following which the actual generation can be started (section [Structure Generation \[96\]](#)).
- One powerful tool to facilitate the structure generation process is the optional Fragment Tool (section [The Fragment Tool \[81\]](#)), where fragments can be defined and used as filtering elements. Or if assignments are known, they can be used as building blocks for the generation algorithm.
- The initial step of structure evaluation starts with using chemical shift prediction to rate the generated structures assists in determining the most probable structure (section [Structure Rating via Chemical Shift Prediction \[107\]](#)). This ^{13}C chemical shift prediction is automatically performed as part of the generation process.
- Additional tools for evaluating the generated structures are also included in the CMC-se module included such as displaying the correlations from the table onto the structures (section [Viewing Correlations on Structures \[110\]](#)).

3.1 NMR Spectra for a Structure Elucidation

As mentioned before, the Structure Elucidation module requires a suitable set of NMR spectra. The following section describes the supported and required experimental data. To benefit the most from the automatic features of CMC-se, the recommended parameter sets are included in TopSpin 3.1 and greater, and their features are described below. Parameter sets for those experiments (ADEQUATE/INADEQUATE etc.) that were not supported by CMC-se version 1 are only included in TopSpin 3.2 and greater.

3.1.1 Required Parameter Sets

The two included automatic structure generators use HSQC and HMBC correlations as the most important structural constraints. Therefore the following experiments are mandatory:

1D - ^1H – A quantitative 1D proton experiment is required for analysis

- Pulse program = zg30
- Parameter set = CMCse_1H
 - The proton integrals are used as a consistency check for the assigned proton distribution, and for automatic detection of symmetry. Therefore, it is important that this experiment be quantitative. Thus, in addition to using the 30 degree pulse, the relaxation delay has been set to 2 seconds in this parameter set.

$^1\text{H}/^{13}\text{C}$ HSQC – multiplicity edited version is required to determine the number of the protons in the CH_n groups.

- Pulse program = hsqcedetgpsp.3
- Parameter set = CMCse_HSQC
 - It is necessary to use the multiplicity edited HSQC, matched-sweep adiabatic HSQC sequence so that the automatic analysis can correctly assign the proton multiplicity (CH vs. CH_3 vs. CH_2) for the carbon atoms.
 - To distinguish CH from CH_3 , the integrated peak-areas are used. Using J-matched adiabatic-sweep pulses for H-C polarization transfer, makes the HSQC peak intensities, less dependent on the $J_{(\text{CH})}$ coupling constant, and thus more quantitative. Due to the increased bandwidth of these adiabatic pulses compared to rectangular pulses, there is also a minimization of phase artifacts, making the automatic analysis that much more reliable.
 - Note that the integral values from the sensitivity-improved (si) versions of the pulse matched-swept experiments are not as quantitative as the recommended sequence. Therefore the automatic analysis for proton multiplicity/distribution may be affected if these “si” sequences are used.

$^1\text{H}/^{13}\text{C}$ HMBC – a version with 3-fold single-J suppression filter, to maximize the suppression of ^1J artifacts is recommended:

- Pulse program = hmbcetgpl3nd
 - This pulse program was chosen because the echo-anti-echo detection mode results in greater sensitivity in addition to the 3rd order low-pass filter which is more efficient at eliminating the ^1J artifacts.
- Parameter set = CMCse_HMBC
 - The values for J in this parameter set have been selected to optimize for ^2J - ^3J correlations. Any spectra optimized for longer range correlations will be misinterpreted by the structure generation algorithms (see section [Chemistry Rules](#) ▶ 95).

For optimum results with the automatic analysis, **appropriate resolution is crucial**. The parameter sets are set up to use the automatic ^1H SW optimization in ICON-NMR (and thus need to be set up as composite experiments for users – refer to the TopSpin Manual for instructions), or should be optimized manually if possible. Additionally, to accommodate most molecules, TD has been increased to larger values than in the standard Bruker parameter sets; the following are recommended starting points:

- **HSQC**, $\text{TD}^* = 2048$, $\text{TD1} = 400^{**}$ or more
 - Because TD in the ^1H dimension has been increased, it is important to pay attention to the acquisition time (AQ) – if the optimized ^1H SW is small, this might become too long, and could damage the probe due to the ^{13}C decoupling during acquisition. If this is the case, the full 2048 points are not necessary to maintain the resolution, and can be reduced. Because of this, the AU program used in the CMCse_HSQC parameter set (au_getlinv_CMCse) automatically sets the acquisition time to a value between 0.15 and 0.3 seconds, and adjusts the number of points accordingly.

- Because the resolution in the HSQC experiment is more critical than other 2D experiments, the above mentioned AU program also adjusts the TD1 to maintain a constant AQ(F2) of 12 ms so that TD/SI values are set appropriately for the SWH value. This is something that should be noted when collecting the data manually, especially at higher fields.
- **HMBC**, TD = 4096, TD1 = 512 or more
 - It is not absolutely necessary to resolve all signals in the HMBC spectrum, *fuzzy* correlations are allowed in cases where they cannot be resolved (section [Additional Manipulations to Correlations in Table \[75\]](#)) but having multiple unresolved signals greatly complicates the structure generation process.
 - One way the additional resolution could be obtained would be to run additional Selective HMBC experiments. **The use of multiple HMBC spectra is supported in CMC-se version 2.** The Automatic Analysis will recognize and take the correlations from the spectrum with the higher resolution when available, and from the full HMBC in regions where a selective experiment is not available. This would be particularly useful if the ¹H spectrum is severely overlapped as **the CMC-se software can handle EITHER (but not both) “fuzzy” ¹H OR ¹³C assignments for a single correlation.**

If known ahead of time that signal overlap is not a problem with the molecule of interest, the number of points can be decreased without affecting the above stated aspects of the CMCse_* parameter sets.

If necessary, conservative use of linear prediction beyond what is in the standard parameter sets is permissible. However, care must be taken to ensure that the data used in CMC-se does not contain any artifacts introduced by the linear prediction. These artifacts typically arise when the S:N of the starting data set is too low, and the number of predicted points is too aggressive (i.e. predicting 4k points from only 256 real points is not recommended).



While high resolution is desired, the automatic analysis algorithms of CMC-se do **require at least 5% of the SW of noise on both sides of the spectrum for the directly acquired dimension.** Thus optimizing the SW too much is more detrimental than beneficial.

3.1.2 Additional Supported Experiments

To get additional information which can be useful for structure elucidation the following experiments are recommended when for the reasons indicated below. It is HIGHLY REDOMMENDED to include the ¹H-¹H COSY, the ¹³C, and the ¹H/¹⁵N HSQC/HMBC if at all possible. The H2BC, ADEQUATE, INADEQUATE, are supported by CMC-se, but are not always necessary to generate structures. In fact, they were not supported in the initial release of CMC-se. CMC-se parameter sets for these (H2BC, ADEQUATE, INADEQUATE) were not included until TopSpin 3.2 patch level 1.

¹H-¹H COSY – provides additional structural constraints and reduces the number of suggested structures.

- Pulse program = cosygpppmfqf
- Parameter set = CMCse_COSY
 - The automatic analysis requires the magnitude mode (non-phase sensitive) peaks for analysis, and the double quantum filter is used to simplify the spectra. The purge pulse has been included to reduce artifacts which arise due to too short of a recycle delay.
 - Note this was not a standard pulse sequence prior to TopSpin 3.0 – contact your local Bruker representative for a sequence that will work with older versions of TopSpin.

- Just as with the above mentioned HSQC/HMBC parameter sets, the CMCse_COSY parameter set has higher TD values than the regular COSY parameter set.

1D ^{13}C – guarantees most accurate chemical shift determination.

- Pulse program = zgpg30
- Parameter set = CMCse_13C
 - For adequate automatic analysis, it is necessary to have a region that represents 10% of the entire sweep width on both sides of the spectrum that is free from any signals. This is not always the case with the standard Bruker parameter set, and thus a CMC-se specific parameter set has been created to work in most cases. If the molecule of interest has unusual shifts, this may need to be modified further.
 - The 1D ^{13}C spectrum is particularly important for crowded regions of the 2D spectra where identification of multiple signals is not possible.
 - APT, DEPT and DEPTq type spectra are supported and useful in cases where the resolution is not sufficient in the 2D experiments, and a regular 1D- ^{13}C experiment is not available, but not recommended due to the missing (or weak for DEPTq) quaternary carbons.

^1H - ^{15}N HSQC/HMBC – provides additional structural constraints and reduces the number of suggested structures.

- Pulse Program = hsqcetgp / hmbcplpndqf
- Parameter Set = CMCse_15NHSQCf2 / CMCse_15NHMBCf2
 - There are currently no special requirements for the nitrogen experiments. However, considering how useful they can be in the structure elucidation process, there are CMC-se parameter sets available, which are set up to accommodate ^{15}N on the F2 channel of the spectrometer.

^1H - ^{13}C H2BC – helps differentiate between ^2J and ^3J correlations of the HMBC.

- Pulse Program = h2bcetgpl3
- Parameter Set = CMCse_H2BC
 - This experiment is supported due to its popularity. However, it should be used with caution as it suffers from the same problem as the ^1H - ^1H COSY where long range ($>^3\text{J}_{\text{HH}}$) correlations are often present.
 - Because this is a constant time experiment, there is a limit to the number of increments that can be used in F1 before the incremented delay starts becoming negative. If the experiment won't start because a duration becomes negative, then the number of increments in F1 should be reduced. If additional resolution is critical, it is possible to increase the constant time delay (d21), however this will lead to the evolution of smaller J couplings and most likely longer range (invalid) correlations.

^1H - ^{13}C ADEQUATE – unambiguously determines the neighboring ^{13}C atoms for each ^1H .

- Pulse Program = adeq11etgprdsp
- Parameter Set = CMCse_ADEQUATE
 - This refocused version of the 11-adequate is used for the simplicity in its interpretation.
- **^{13}C - INADEQUATE** – When available, unambiguously determines the ^{13}C - ^{13}C backbone of the molecule
- Pulse Program = inadphsp

- This pulse sequence was chosen because of the inclusion of the adiabatic pulses for optimum bandwidth, and the phase sensitive nature of the peaks makes determining real signals from artifacts easier.
- Parameter Set = CMCse_INAD
 - The parameter set included in TopSpin is set up such that the SW in the indirect dimension (F1) is 2x that of the direct dimension (F2). As will be illustrated later (section [Adding INADEQUATE Correlations \[▶ 68\]](#)) there are advanced display options available for analyzing data which has the same SW in both dimensions. This allows for increased resolution by folding the spectrum in the double-quantum dimension.
 - When processing the INADEQUATE, it is likely necessary to manually reference the indirect dimension. The center of the spectrum in F1 should be set to 2 x O1p.



A note about the INADEQUATE and ADEQUATE spectra. While these types of spectra are incredibly information rich, they rely on the rare chance that one ^{13}C is next to another ^{13}C in the same molecule. At natural abundance, this is a very small occurrence, and thus these experiments require highly concentrated samples, high sensitivity probes, and or large amounts of experimental time. This is especially true for the INADEQUATE which is a ^{13}C detected experiment. A general guideline to judge the ability to acquire the INADEQUATE in a single overnight (~14 hours) experiment is that a single scan 1D- ^{13}C experiment needs to have a signal to noise of 100:1. Because the ADEQUATE is a ^1H detected experiment, this requirement is much less, but experimental time is still greater than other 2D indirect detect experiments.

If not listed in this section, there are no special requirements or expectations for the additional supported experiments both those used for structure generation such as ^1H - ^{31}P HMBC, or ^1H - ^{19}F HETCOR, and those experiments that are supported for use of structure verification such as TOCSY, NOESY, and HSQC-TOCSY. And thus any experimental setup is sufficient. The CMC-se software needs to identify the type of experiment to be able to display and use the spectra correctly. The identification is done using the SPECTYP status parameter as described in the next section.

3.1.3 Data Set Processing

To get the most suitably processed spectra the following parameters are recommended:

- It is **crucial** that the HSQC will be phased in such a manner that the signals from CH and CH_3 are positive, while the CH_2 are negative.
- Notes about HMBC and H2BC Processing:
 - The recommended pulse programs are a phase sensitive experiment, to get the necessary magnitude data it is necessary to use the **xf2m** command after executing **xfb**.
 - Proper apodization is to use SINE (F2) and QSINE (F1) functions with SSB=0 in F2 and SSB=2 in F1 (This is very important because other SSB values lead to extra broad lines in F1).
 - The parameter sets should take care of these when processed in automation, but when processing manually, it is important to take note of the extra **xf2m** command.
- Spectrum Referencing:

- The 1D and 2D spectra should be carefully calibrated. It is not necessary that they be absolutely referenced (i.e. TMS = 0.0 ppm), but rather that the calibration be consistent between the different 1D and 2D spectra. For this reason, the **sref** command has been removed from the automated processing AU programs. The manual calibration may be easily done in the topspin multiple display:

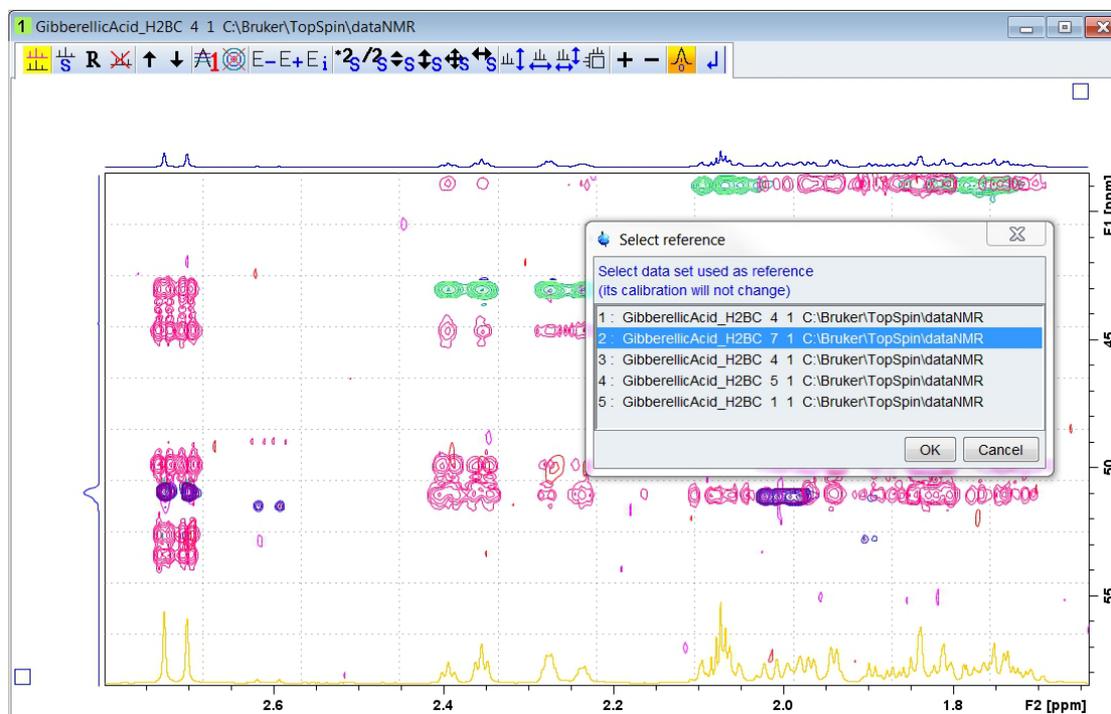


Figure 3.1: Referencing several 1D and 2D data sets in TopSpin multiple display.

The automatic analysis routine has a built in check for non-consistent *SR* values in the ^{13}C dimension. If the different spectra have different *SR* values, a pop-up window will be displayed:

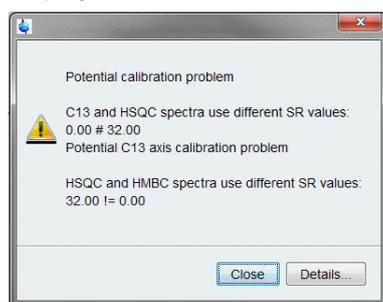


Figure 3.2: Pop-up window indicating a potential problem with the consistency of the *SR* values between different spectra

- If it is intentional that the **SR** values are different (i.e. what was necessary so the referencing between spectra were consistent) this message can be ignored. Otherwise, the analysis should be canceled, and the *SR* processing parameter fixed as necessary prior to proceeding.
- As mentioned above, if including the INADEQUATE data, proper calibration of the reference of the double-quantum (F1) dimension needs to be done manually. There is not a built in check for this, however, it will be apparent when attempting to pick the correlations, as the software will report errors “No ^{13}C found at xxx ppm”.

- It is necessary that the SPECTYP for each spectrum be set correctly, so that the automatic analysis knows how to interpret each spectrum. For the default parameter sets this is done automatically. If not, check the **status** parameter SPECTYP in the edp processing status parameter display. If necessary, type **spectyp** on the TopSpin command line, and select the appropriate type.

The automatic SPECTYP assignment is based on the analysis of Bruker pulse program names and acquired nuclei. The software supports the following types:

- PROTON
- C13 (APT, INEPT, DEPT or DEPTQ are supported but not recommended)
- F19
- HSQC (HMQC supported but not recommended)
- HMBC
- COSY
- HETCOR
- HSQC-TOCSY
- ADEQUATE
- INADEQUATE
- TOCSY

3.2 Starting CMC-se

CMC-se is a TopSpin module, so it will be started as a special analysis option out of the TopSpin main window. It is not necessary to open a file of the dataset to be examined to start CMC-se; however, it is recommended as it will save time in preparing a new project.

There are multiple ways to start the Structure Elucidation module:

- Type `cmcse` into the command line, press return.
- Click **Analysis | Small Molecules | Structure Elucidation [cmcse]** (TopSpin 4).
- Click **Analyse | Structures | Structure Elucidation [cmcse]** (TopSpin 3).

After any of these, the Structure Elucidation Project Selection window will open.

3.2.1 Structure Elucidation Project Management

The Structure Elucidation Project Selection window shows a list of recently used projects which can be opened from this window. Additionally, in the upper right hand corner is a **New** button which is used to start the creation of a new project (see section [Creating a New CMC-se Project \[▶ 23\]](#)):

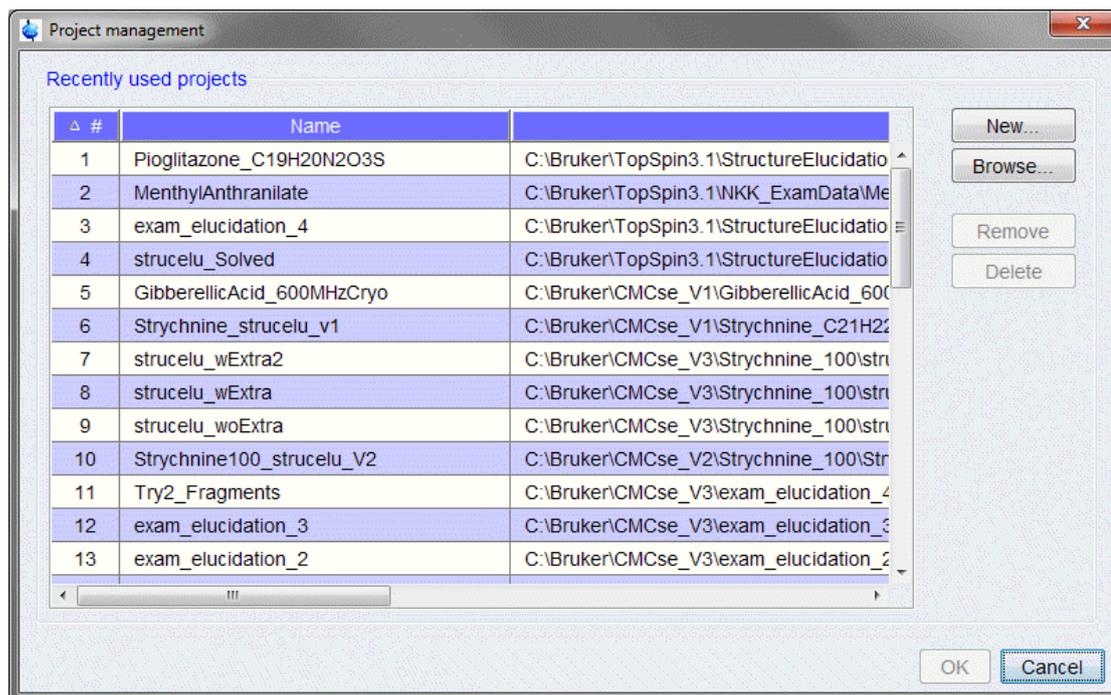


Figure 3.3: Structure Elucidation Project Selection Window

A CMC-se project is an ordinary directory on your disk containing the relevant files for the elucidation process. All necessary information is collected here.

The most important files are:

- `<project name>.xcmap` contains the correlation map (as well as the complete project information including the location of the used NMR data sets).
- `<project name>.cmap` is the old data format (CMC-se 2.2 and older) and will be converted with the first saving process into the new data format. It can be deleted afterwards.
- `<project name>.sdf` holding the generated structures shown in the CMC-se display.
- `*.sdf.bak` files are copies of the corresponding files that are created prior to generating or ranking the structures, and can be used as a back-up if necessary.
- `MetaGenerator.log` is a structure generator log file.
- `<project name>.ini` holds additional settings like the size and position of CMC-se window.
- `FragmentEditor.mol` holds all defined fragments from the Fragment Editor.
- `Fragments.sdf` holds all fragments from the Fragment Filter dialog.

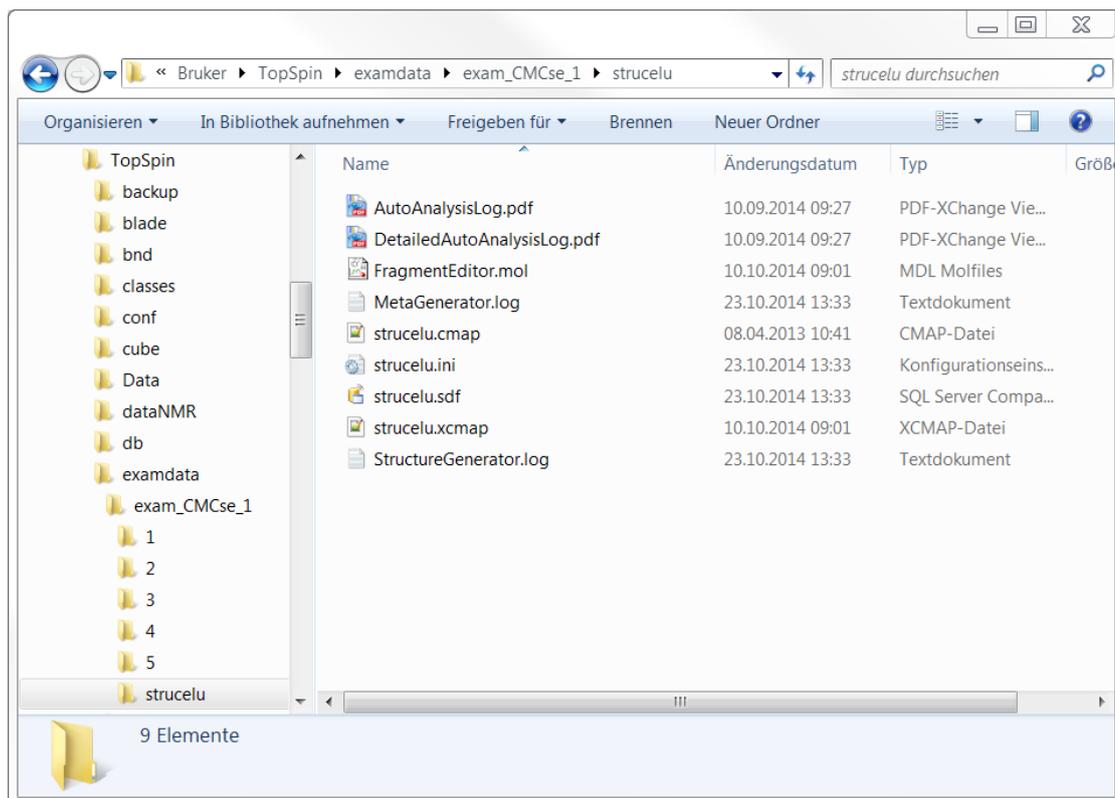


Figure 3.4: The Project Directory

The location of the project directory is not fixed, but it is recommended to keep the default setup, with the project directory created in the dataset tree. For example:

```
/CMCse/<ExperimentName>/strucelu
```

with the experiment number folder in the same path

```
/CMCse/<ExperimentName>/<expno>
```

In addition to keeping the project with the raw data to which it is tied to, if the default path name and location is kept, each time CMC-se is started from within an experiment in that data set, the project is automatically opened.

Even if a CMC-se project is contained within a data set, this window can be brought up at any time using the **cmcse dia** command from the TopSpin command line.

3.2.2 Creating a New CMC-se Project

If CMC-se is being used for the first time on a particular data set, clicking on the **New** button of the “Structure Elucidation Project Management” window ([Structure Elucidation Project Management \[21\]](#)), will bring up the following dialog box which is the basis for creating a new project:

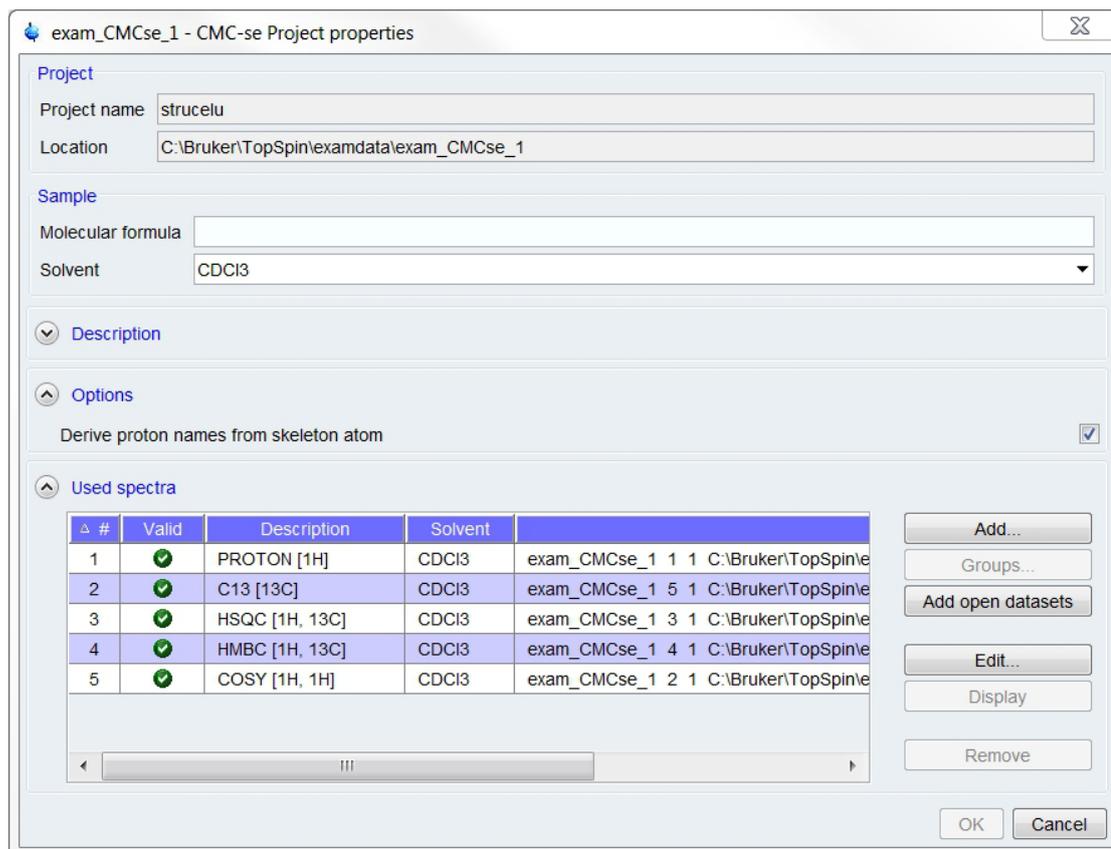


Figure 3.5: Project dialog, illustrating the information automatically filled in upon starting CMC-se with an active data set.

The minimum information needed for the project creation is a project name, project location, molecular formula, solvent and definition of which data sets to be used.

As shown in the figure above, if a project was started while a data set was active in TopSpin, the software fills in many of the required fields automatically.

- The default **Project name** is automatically filled in as **strucelu** and **Location** is the current dataset path.

It is possible to change the name and location of the project; however it must be kept in mind that the project is still connected to the original data defined here during creation. Any changes such as reprocessing, calibration or peak picking will affect *all* projects tied to this data. This is the motivation for keeping the project within the data set tree.

- The **Molecular formula** is the only field that is not filled in automatically and must be entered manually. The molecular formula is not only atom sensitive (see [Program Prerequisites \[11\]](#) for supported atoms), but also case sensitive (e.g. CDCl3 is supported, while CDCL3 is not). The text is automatically validated, a balloon help is shown in cases of non-supported input:

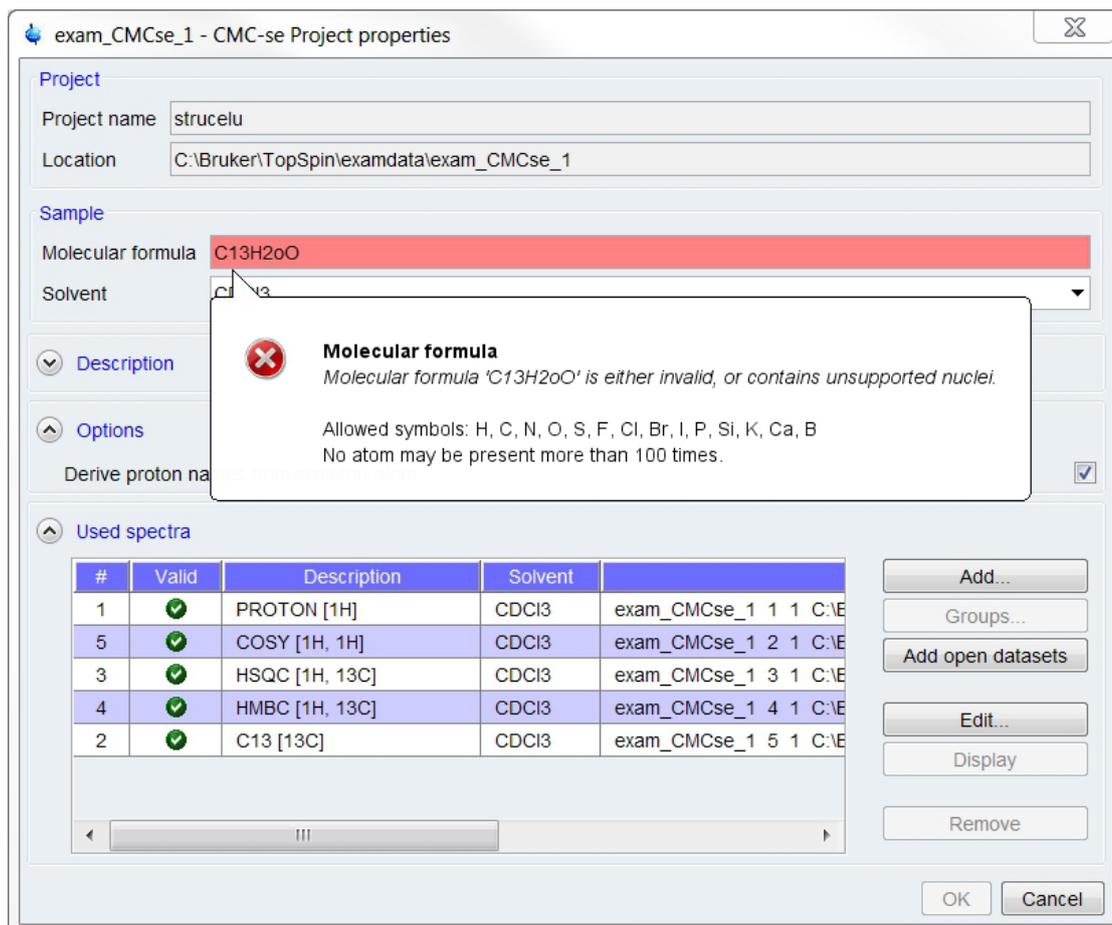


Figure 3.6: Project window illustrating the balloon help pop-up window upon entering an incorrect molecular formula – In this case, a lowercase “o” was used instead of the required “O”.

- The **solvent** automatically populated is the one defined in the corresponding spectra. It is crucial that this be defined correctly since it is used in the automatic analysis to remove peaks from solvent regions, and will be used by the chemical shift prediction routines. The solvents listed in this pull-down are the ones that are currently in the CMC-se data base.

If your particular solvent is not there, choose the one that is closest, and then expect manual corrections (Section [Actions Possible through the Combined Spectra Display \[52\]](#) and [Additional Modifications to the Correlation Table \[71\]](#)) to the Correlation Table to be necessary.

Starting in version 2.0 of CMC-se it was possible to use mixtures of different solvents. Use “+” to specify additional solvents (see the following figure).



Note that when the software is looking for solvents, it has a multiplet pattern and a chemical shift range that it is looking for solvent peaks (i.e. CDCl₃ is a triplet around 77 ppm). If using a mixture of solvents causes either or both solvents to deviate from this position, they will not be recognized as solvent peaks.

exam_CMCse_1 - CMC-se Project properties

Project
 Project name: strucelu
 Location: C:\Bruker\TopSpin\examdata\exam_CMCse_1

Sample
 Molecular formula: C₁₃H₂₀O
 Solvent: CDCl₃+MeOD

Description

Options
 Derive proton names from skeleton atom

Used spectra

#	Valid	Description	Solvent	
1	✓	PROTON [1H]	CDCl ₃	exam_CMCse_1 1 1 C:\E
5	✓	COSY [1H, 1H]	CDCl ₃	exam_CMCse_1 2 1 C:\E
3	✓	HSQC [1H, 13C]	CDCl ₃	exam_CMCse_1 3 1 C:\E
4	✓	HMBC [1H, 13C]	CDCl ₃	exam_CMCse_1 4 1 C:\E
2	✓	C13 [13C]	CDCl ₃	exam_CMCse_1 5 1 C:\E

Buttons: Add..., Groups..., Add open datasets, Edit..., Display, Remove, OK, Cancel

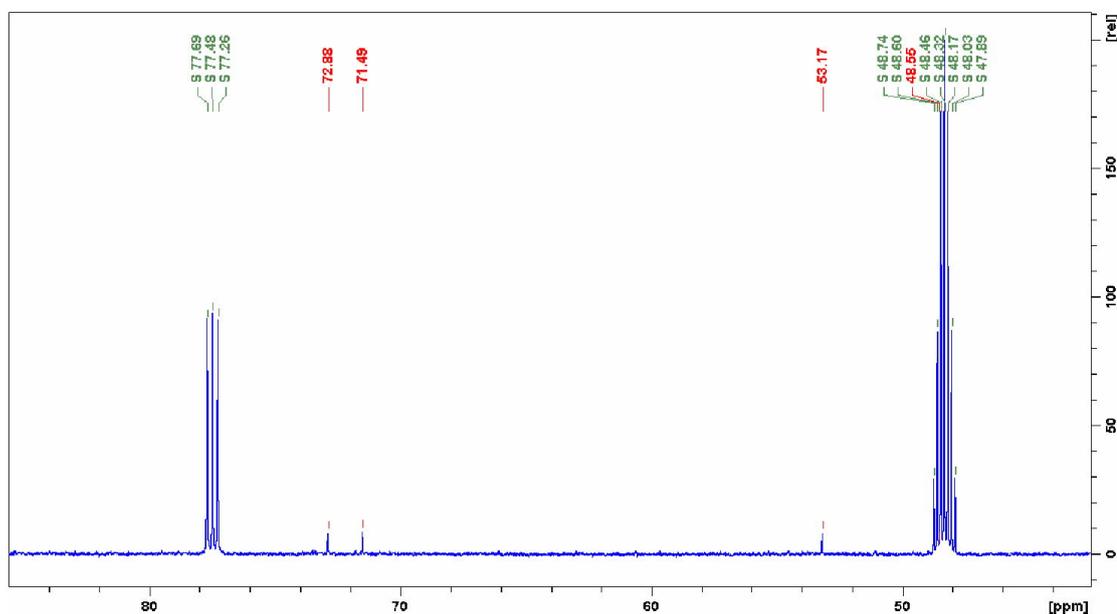


Figure 3.7: Example of using "+" for mixed solvent system. The data analysis detects and excludes solvent peaks from all components.

- The *Options* field of the **Create new project window** is check box providing the user the option to automatically re-name the ^1H signals based on the heteroatom they are attached to. This option is default as the structure generator things of the ^1H atoms not as individual atoms, but being connected to a heteroatom, so by re-naming them automatically, the user is presented with the information that the structure generator is using.
- The *Description* field of the **Create new project window** allows user remarks to be added and stored with the project (for example, the suspected substance).
- The *Used spectra* field of the dialog automatically lists all **processed** spectra (with different experiment and processing numbers) of the current data set (experiment name). The processing numbers widely used for temporary processing such as 999 are ignored.

Note that the **spectype** of the data is set is automatically analyzed and included in the *Description* field of each entry. If this is not set correctly, the project creation canceled until it is correctly fixed.

Any entry in the data list may be edited or removed via any of the following ways if the automatic selection is not suitable.

- **Add:** opens the *Find data* window:

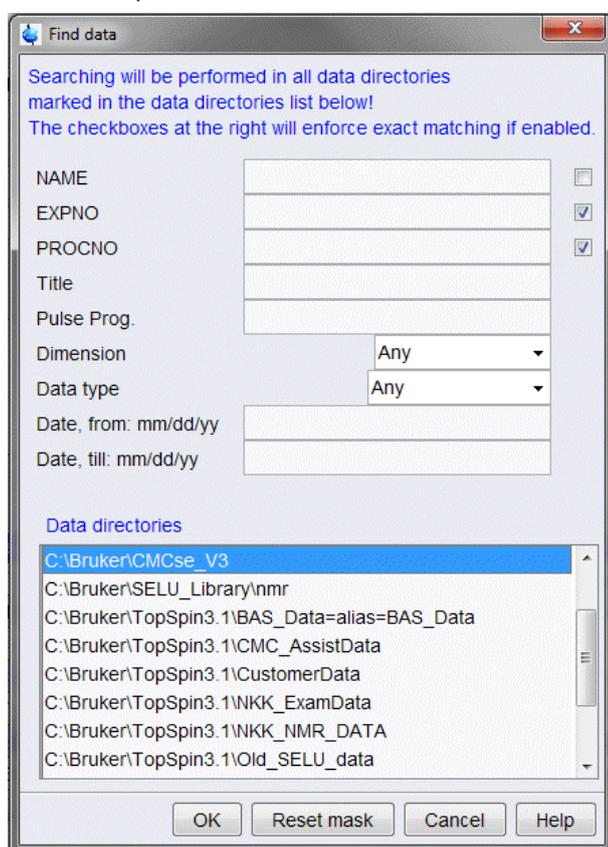


Figure 3.8: Find data window to assist in adding more spectra to the project.

- Spectra can be located according to several criteria. For more help click the **Help** button and the online manual is opened at the find/search topic. After specifying search criteria click **OK** and the search result is displayed in a new window. Mark the datasets desired to be added to the project, and click **OK**.
- Another way to add datasets to a project is by clicking the **Groups** button. If data sets have been previously defined as a group (Refer to the TopSpin Users Guide) in TopSpin, the self-explaining **Select group** window opens and displays a list of data set groups

previously defined in the TopSpin browser. Selecting a group will add all its datasets to the project. This is useful, when not all data for the project was contained within the same data set (experiment name).

- Additionally, it is possible to add the currently displayed data sets in TopSpin to the project. Just click **Add open datasets** and they will be displayed in the *Used spectra* field. This is again useful, when not all data for the project was contained within the same data set name.
- Alternatively click **Edit** and a text window is brought up where the data path can be entered, or deleted manually.

If unsure which of the data sets are to be included in the project, the **Display** button displays the highlighted data set in TopSpin.

To delete one or more entries of the *Used spectra* field highlight the desired one(s) and click **Remove**. This command only removes the entry from the list for this CMC-se project, data files themselves, are not affected.

After adding all information, save and create the project by clicking the **OK** button. A new CMC-se **Project window** will automatically open outside of TopSpin. This window contains the currently empty Correlation Table. Also upon project creation, the Combined Spectra Display window is opened within the normal TopSpin display area, displaying the un-analyzed spectra defined for that project all together. Screen shots of the initial state of these windows are shown below.

While the bulk of the Correlation Table is empty because the auto analysis has not been performed (see section [Automatic Spectra Analysis \[▶ 39\]](#)), the appropriate number of columns and rows have been created based on the molecular formula. As will be described in more detail in Section [The Correlation Table \[▶ 33\]](#), there are multiple tabs and or sections for the various possible correlations (H-C, H-H, C-C). And the right side of the window is the Fragment Tool which will be described in greater detail in Section [The Fragment Tool \[▶ 81\]](#).

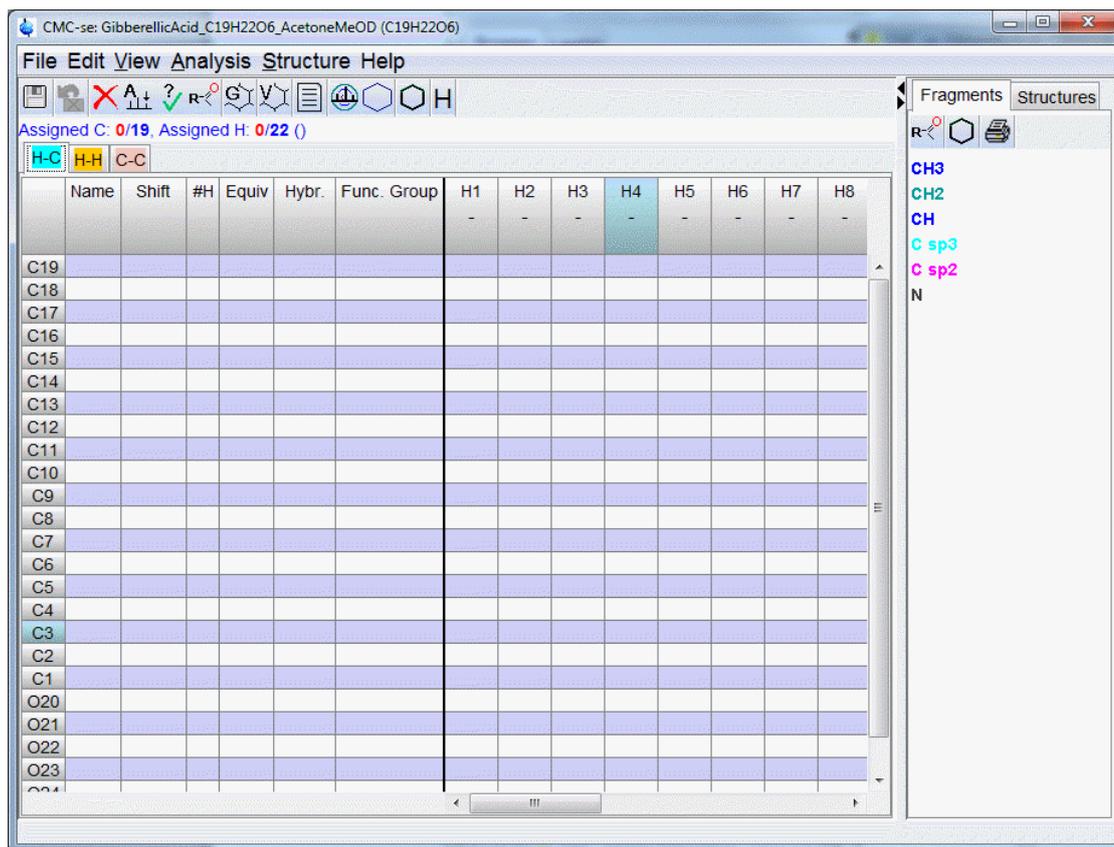


Figure 3.9: New CMC-se project window, containing an empty Correlation Table with appropriate number of rows and columns are present, and the fragment tool on the right side of the window.

If a mistake was made and the formula needs to be changed, it is possible to do this from the **Edit Project Properties** selection of the **Edit Menu**.

Similar to the correlation table, the *Combined Spectra Display* is automatically opened upon creation of the project.

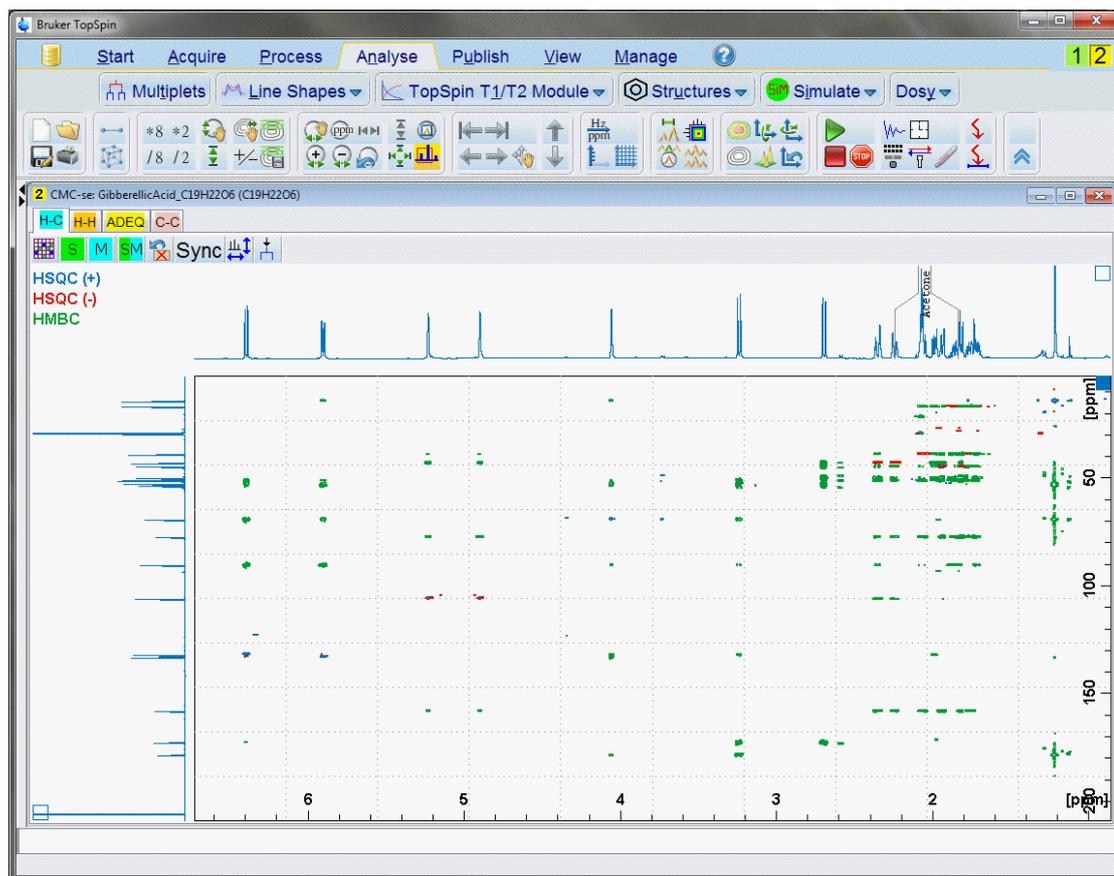


Figure 3.10: Initial Combined Spectra Display within the main TopSpin window.

As seen in the figure above, this initial window contains many of the spectra from the project. In this first (H-C) tab, the HSQC/HMBC are overlaid and the ^1H and ^{13}C projections are on the respective axis. The other tabs have the respective spectra in a similar fashion. More details about the work-flow and manipulating the view of this window are presented in section [The Combined Spectra Display Window \[46\]](#). Because the automatic analysis (section [Automatic Spectra Analysis \[39\]](#)) has not been performed, no peaks are indicated in the spectra.

The CMC-se project is now also visible in the data directories of the TopSpin data browser and may now also be opened using either the Drag & Drop function of TopSpin, or displaying one of the spectra from the project and typing the **cmcse** command:

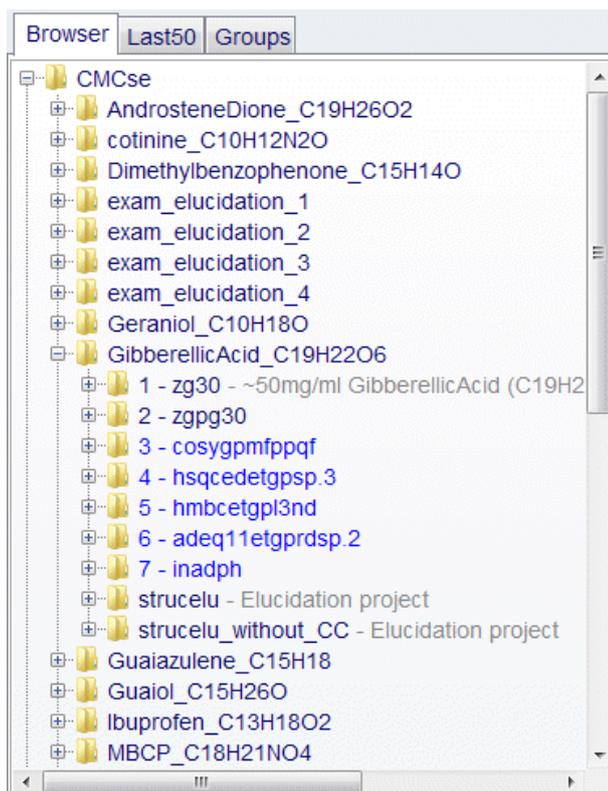
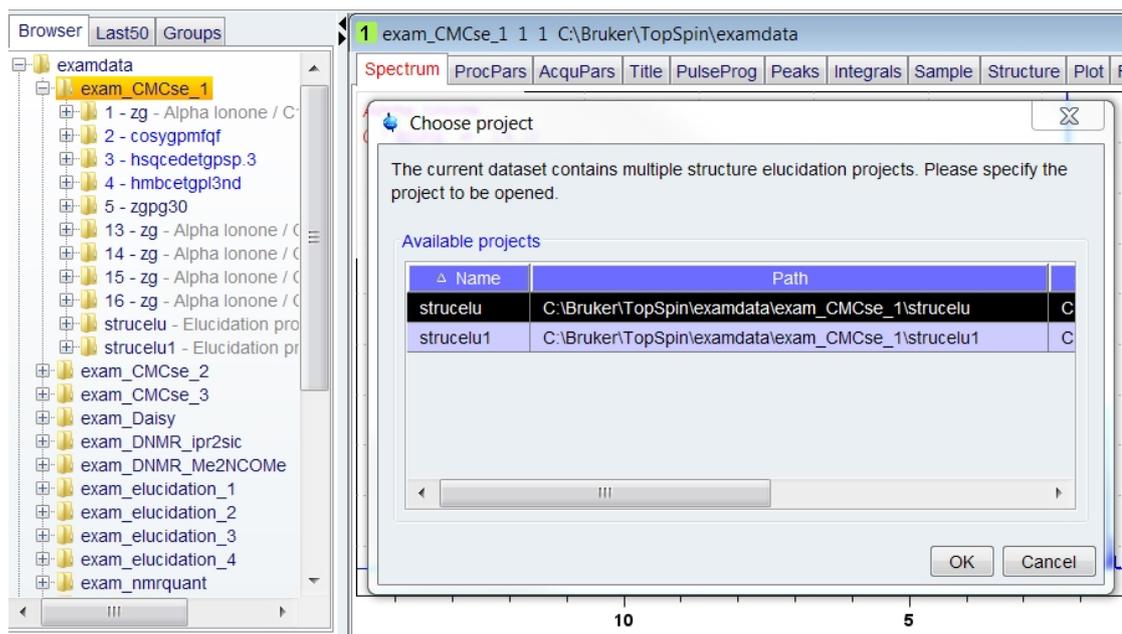


Figure 3.11: Data browser showing the Structure Elucidation (CMC-se) projects in the data tree.

As shown above and in the following figure, a second project tied to the same data has been created (using the **File**→**Save As** function from the menu bar within the CMC-se project window) and now can be opened in the same manner.

As mentioned previously, this new project is still tied to this same data set, but perhaps has a correlation table with different correlations enabled or disabled (section [The Correlation Table \[33\]](#)), or different fragments/assignments assigned (section [The Fragment Tool \[81\]](#)).

If the **cmcse** command is executed and only one project is in the directory, that project automatically opens. However, when there are multiple CMC-se projects in the same data set, as shown above, and **cmcse** is executed from that active data set, a new pop-up window is shown allowing the user to select which project is desired:



selu] Found multiple projects.

Figure 3.12: Project Selection dialog box in the cases that more than one CMC-se project is included with a data set.

Another (quicker) way to build a project is using the newly implemented command:

```
cmcse create
```

A good overview how to use this command with arguments gives the command:

```
cmcse help
```

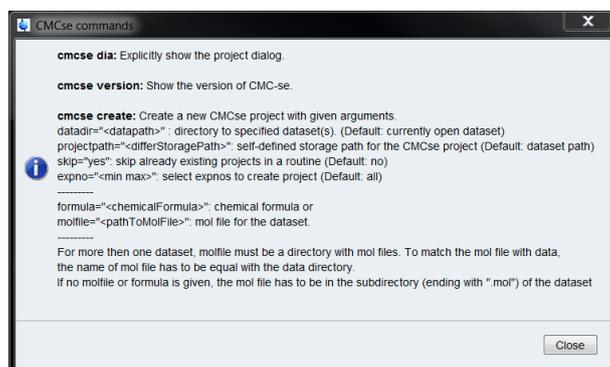


Figure 3.13: Help how to use the CMC-se commands.

The quickest possibility is having datasets in a standard folder with one optional mol file in the directory (for the verification). If one of the experiments is open and the create command is called, all experiments in the directory will be included in the new project as default. The chemical formula will be taken from the argument or from the optional mol file. To differ from default values in the project settings, a combination of arguments are possible, shown in. It is also possible to change settings afterwards with **cmcse dia**, like written before.

For example, any desired experiment in the dataset folder is open (it is not forced that this single experiment must be in the new project) and a new project should be created including only the experiments number 1 till 10, the following command is correct:

```
cmcse create expno="1 10"
```

3.3 The Correlation Table

As described previously, one of the two main windows of the CMC-se module is the *Project window* containing the Correlation Table, and other menu and function options. This is a separate window from the main TopSpin window, and as stated above, is automatically opened when a CMC-se project is created or opened. All CMC-se based functions are accessed through the menu of this CMC-se Project window.

In the figure below, an automated analysis (section [Automatic Spectra Analysis \[39\]](#)) has been performed and the results are filled in to illustrate the possible contents. As shown below, after an automatic analysis, the chemical shifts of the ^1H and ^{13}C atoms (as well as any others such as ^{15}N , ^{19}F , ^{31}P when available) are filled in, and where present, correlations are added as color-coded cells to the table (see description below). The column for attached protons (#H) has been filled in based on the multiplicity edited HSQC (see section [Required Parameter Sets \[15\]](#)), but is left blank for hetero-atoms for which this is unknown. Similarly, if possible based on chemical shift, the hybridization state is also entered into the appropriate column.

CMC-se: exam_CMCse_1 (C13H20O)

File Edit View Analysis Structure Report Help

C13H20O DBE=4.0 C: 13/13, H: 20/20 (4^oCH 2^oCH₂ 4^oCH₃), 32 HMBC, 8 COSY

Information about assigned atoms/correlations

← Tabs for the various types of correlations → Proton Chemical Shift

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17
							6.60	6.03	5.47	2.26	2.23	2.03	1.54	1.44	1.20	0.91	0.83	-	-	-	-	-	-
							C2	C3	C5	C6	C10	C12	C13	C8	C8'	C11	C9	H12	H13	H14	H15	H16	H17
C13		22.72	3		sp3								S+				M*						
C12		22.97	2		sp3							S-					^						
C11		26.75	3		sp3			v		M*						S+ v							
C10		26.89	3		sp3			M*		^	S+			M*		M*							
C9		27.74	3		sp3											M	^	S+					
C8		31.18	2		sp3					M*	<	M		S-	S-								
C7		32.46	0		sp3/sp		M					M		M									
C6		54.26	1		sp3/sp		M	M		S+		M	M	M	M	M	M						
C5		122.62	1																				
C4		131.87	0		sp2		v			v		M	v										
C3		132.30	1		sp2		M*	S+			M		M*										
C2		148.95	1		sp2		S+			M													
C1		198.35	0		sp2		M	M			M												
O14																							

↑ ↑ ↑ ↑ ↑

User added functional group definition (optional)

Hybridization state of hetero-atom

of equivalent carbons (when applicable)

of protons attached to hetero-atom (C,N,O...)

Hetero-atom Chemical Shift

User defined name for atoms (optional)

Figure 3.14: Main project window of the CMC-se module with short explanation of the contents of the "H-C" tab Correlation Table.

The atoms in the table are given a numerical number based on chemical shift, if desired; the user can enter their own name/number in the indicated **Name** column.

As indicated in the figure above, there are multiple tabs for the various types of correlations. The figure above is for the heteronuclear correlations. In the next two figures are the tabs for the homonuclear correlations (H-H, and C-C).

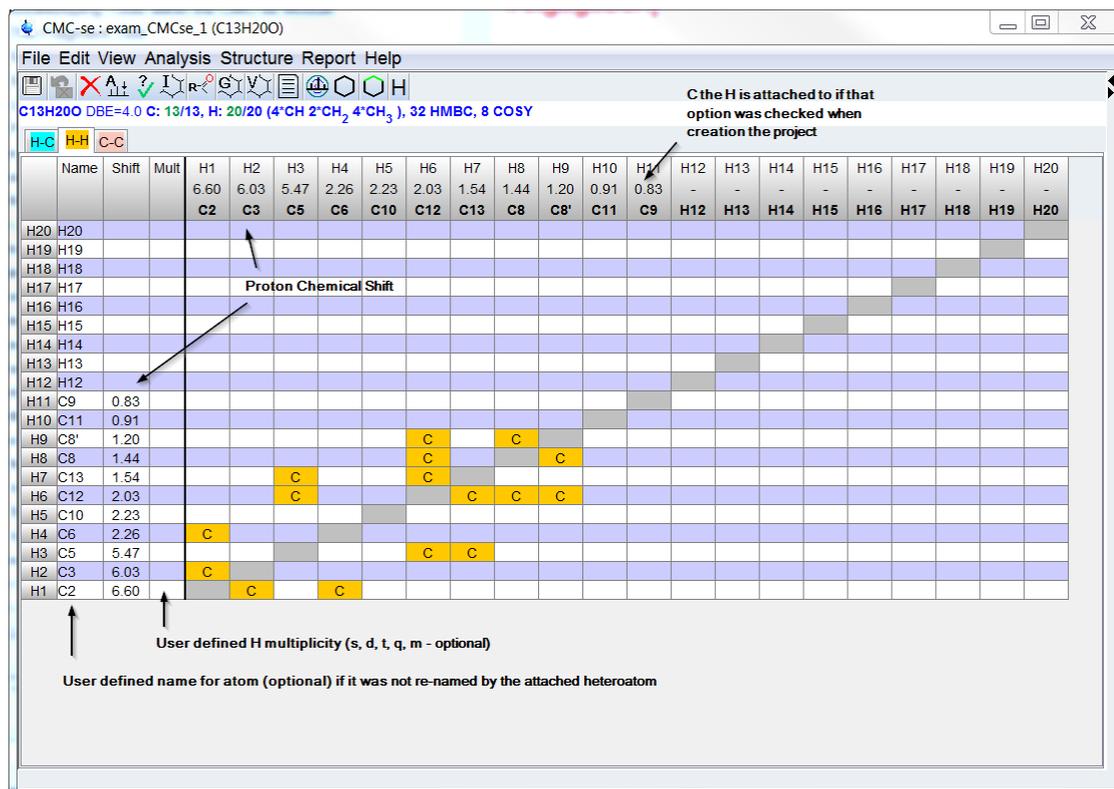


Figure 3.15: A figure with brief descriptions of the “H-H” tab for the homonuclear correlations.

When an INADEQUATE experiment is not included in the CMC-se project, the C-C tab will be available for the user to add C-C or C-X bonds, if they are known through some other mechanism. Shown below is the C-C tab of a project that contains an INADEQUATE spectrum that has been manually picked for the correlations (details on this process will be presented in section [Adding INADEQUATE Correlations \[68\]](#)), as well as the menu option for adding a C-O bond by right clicking in the appropriate cell:

	Name	Shift	#H	Equiv	Hybr.	Func. Group	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	O	
							198.35	148.95	132.30	131.87	122.62	54.26	32.46	31.18	27.74	26.89	26.75	22.97	22.72		
C13		22.72	3		sp3																
C12		22.97	2		sp3																
C11		26.75	3		sp3										I						
C10		26.89	3		sp3					I											
C9		27.74	3		sp3																
C8		31.18	2		sp3			I				I									
C7		32.46	0		sp3/sp																
C6		54.26	1		sp3/sp																
C5		122.62	1		sp2		I														
C4		131.87	0		sp2																
C3		132.30	1		sp2																
C2		148.95	1		sp2																
C1		198.35	0		sp2																
O14																					

Figure 3.16: C-C tab of a project that contains an INADEQUATE spectrum that has been manually picked for the correlations.

After the automatic analysis, information about the number of assigned carbons, protons, and the number of correlations that were assignable by the analysis are shown (see figures above).

The cells in the table(s) have the following meaning:

S+ / **S-** – Positive and negative HSQC correlations

M – HMBC correlation

A – ADEQUATE correlation

I – INADEQUATE correlation

B – H2BC

M A – Cell with both an HMBC and an ADEQUATE correlation (also possible with other types of correlations).

M* – “Fuzzy” HMBC correlation (also possible with other types of correlations, explained in more detail in section [Additional Manipulations to Correlations in Table \[75\]](#)).

^ / **v** / **<** / **>** – Alternative atom assignments for “Fuzzy” correlations

C – COSY Correlation

T – TOCSY Correlation

All necessary tools for working with the CMC-se program are available in the main menu bar of this window. The common tasks are available through the pull down menus or through the buttons. Both are described below.

- The **File** menu allows the user to load, browse for, and save projects.
- The **Edit** menu allows the user to:
 - Undo the last modification.
 - Delete the content of the Correlation Table.
 - Edit the project properties (including molecular formula, solvent, data sets used etc.).
 - Resort the atoms (resorts the atoms based on Chemical Shift, which is useful if additional atoms/correlations have been added).
- In the **View** menu the user can arrange the information displayed on the screen.
 - The user can minimize or restore the correlation window and set the font size of the table.
 - With the **Arrange frames** option, the user can position the TopSpin and Correlation window suitable for their work.
 - It is also possible to view the atom properties and correlations in a table (from the *Show View* option), which can not only be printed, but saved in a variety of formats including .pdf and .csv.



Toggling the layout option will switch the Correlation Table back to a view that was similar to that from the initial version of CMC-se, as shown in the figure below.

- It is particularly useful in this layout to use the  button to hide the protons that are not assigned. Comparing the figure below with the previous figure (brief descriptions of the ^1H - ^1H tab) this action has removed the rows and columns for those numbered protons which don't have chemical shifts. Because multiple protons are connected to the same carbon, the additional cells are unnecessary. A single undefined ^1H is always left free.

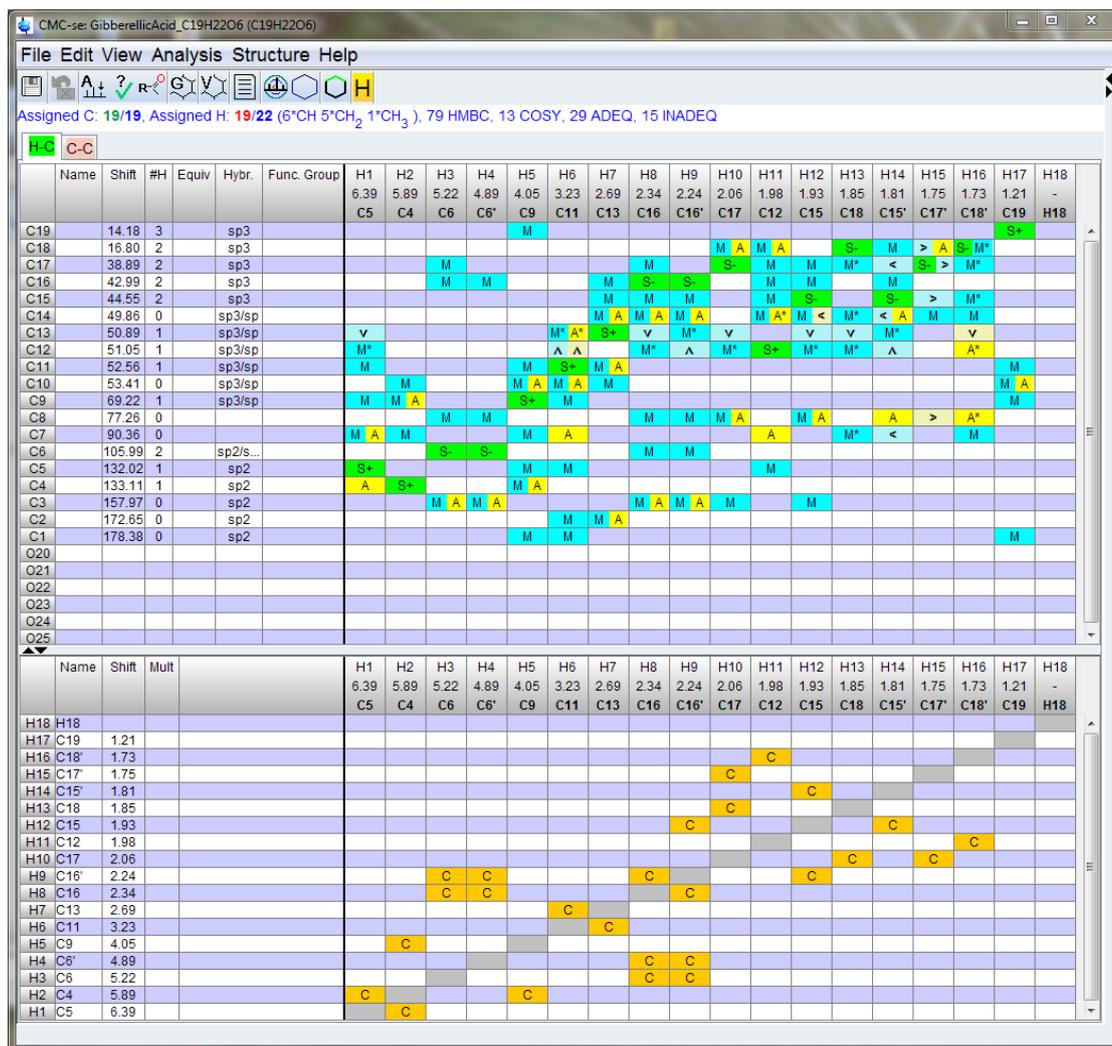


Figure 3.17: Correlation Table after it has been toggled back to the old layout

- The **Analysis** menu offers the tools for the analysis of the spectra set before calculating possible structures. Here the user can start the automatic spectrum analysis and check for consistency. The analysis log file and a detailed technical log file can be opened here.
- The **Structure** menu allows the user to open the main Structure generation window of the CMC-se program. From this menu the user also can display the generated structures, check for invalid correlations and display atoms without any constraints, clear the highlights displayed on the structures, and show the structure generation log file (`MetaGenerator.log`).
- The **Report** menu item generates the summarized Elucidation Report. It involves the elucidated structure with additional common details (like the formula, mass weight, solvent etc.), all detailed carbon and proton spectra and their associated atom tables. Additional information about correct and conflicting correlations, as well as chemical shift correlations, verify the decision for the existing structure. The report based on the generated structure on position one selected in the context menu. If no structure is generated and the Structures tab is empty, the report is generated with the fragment in the Fragments editor.

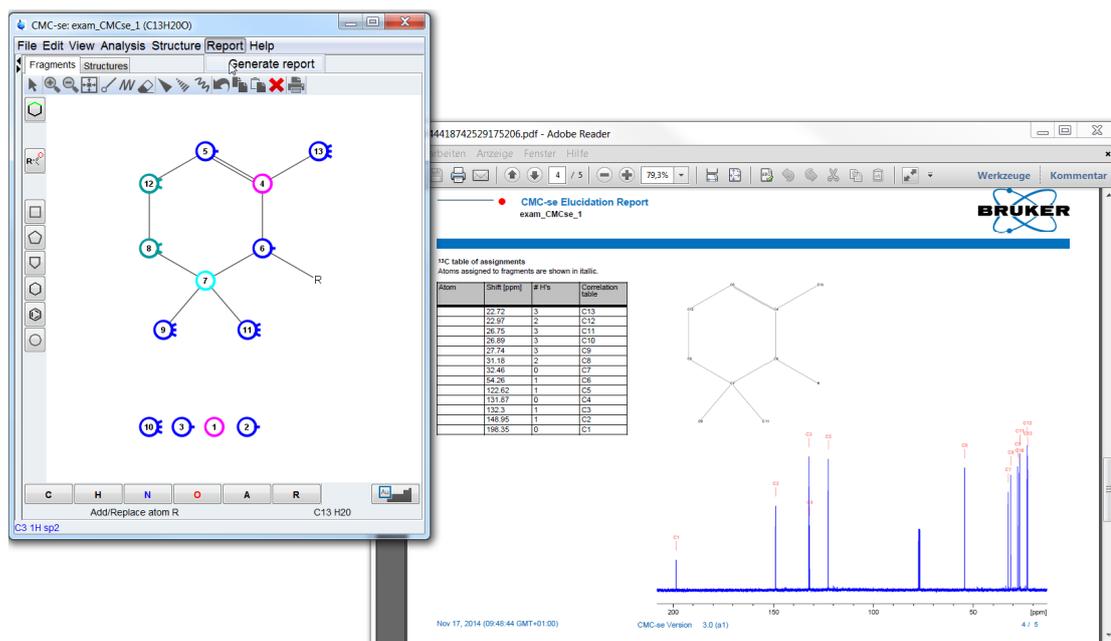


Figure 3.18: Structure Elucidation Report

- The **Help** menu offers access to the Structure Elucidation manual. The databases used for functional group assignment and solvent exclusion can also be viewed from this menu.

The most important tools for a complete structure elucidation are directly accessible from buttons in the toolbar. Here is a brief description of each function, a more detailed description is offered later in this manual:



– Saves the project.



– Undoes the last modifications.



– Deletes the content of Correlation Map.



– Starts automatic analysis of the spectra.



– Displays the results of the Consistency Check of the Automatic Analysis.



– Identify the compound in CSEARCH NMR-Database



– Starts the fragment management tool.



– Generates the structure proposals.



– Verifies a molecular structure against the correlation map.



– Generates the summarized Elucidation Report.



– Re-opens the Combined Spectra Display if it has been minimized.



– Displays the generated structures.



– Clears highlighted correlations from generated structures.



– Removes/displays the extra proton columns.

3.4 Automatic Spectra Analysis

After data collection/processing, and project creation, an automatic spectra analysis is usually the next step in the elucidation process. As stated previously, the PROTON, HSQC and HMBC spectra are required. Other experiments (C13, COSY, ^1H - ^{15}N HSQC/HMBC, H2BC, INADEQUATE, ADEQUATE) are not required, but would significantly improve the reliability of the analysis (especially in the case of the C13) as well as the number and quality of generated structures.

To start the automatic spectra analysis from the menu bar of the Project window of the CMC-se program, choose **Analysis**→**Start Automatic Spectrum Analysis** or click the  button under the menu bar. The **Automatic analysis** dialog window will appear:

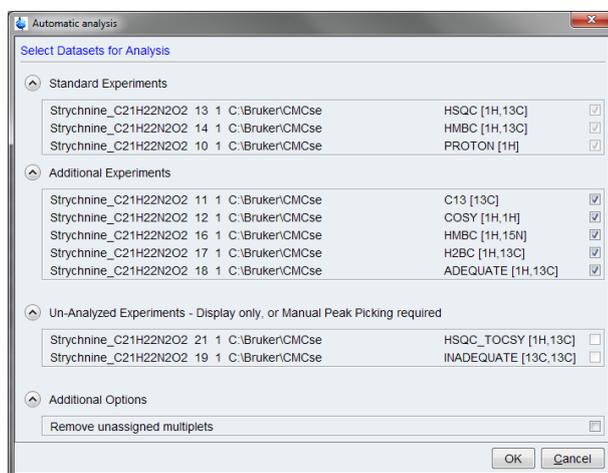


Figure 3.19: Automatic analysis dialog window.

The *Automatic analysis* dialog window displays the datasets of the elucidation project and their types. If these are not defined correctly, it is necessary to return to the data set, and set the SPECTYP parameter as described in section [Data Set Processing \[19\]](#), and then use the *Edit project properties* pull-down menu to re-define the data sets, and the changes to the CMC-se project.

- The required datasets (PROTON, HSQC and HMBC) will always be analyzed and are therefore displayed as Standard Experiments without the possibility to deselect.
- “Additional Experiments” (for example shown here: C13, COSY, ^1H - ^{15}N HMBC, H2BC, and ADEQUATE – but also when present, the selective HMBC, ^{19}F experiments) are grouped together. The user can decide to include those individual experiments in the analysis or not.
- Those data sets which are not automatically analyzed are sorted in the Un-Analyzed Experiments.
 - For some experiments only the first available experiment will be selected for the automatic analysis. In this case further experiments of this type will sorted to Unused Experiments with appropriate information.
 - Some experiments (INADEQUATE, ^1H - ^1H TOCSY) are used by the structure generation algorithms, but require manual peak picking.

- Other experiments (HSQC-TOCSY, ^1H - ^{19}F HETCOR) are for display only, and not used in the generation process.
- Depending on user preference, there is the option to *Remove unassigned multiplets*. Checking this box will simplify the spectra in the Combined Data Display, because any multiplets that could not be assigned to atoms in the Correlation Table will not be shown. However, it is sometimes beneficial to be able to visualize these peaks which were picked, but not assignable, as they may be useful.
 - It is also possible to remove these peaks in a separate step after the automatic analysis, thus providing the user the option to view these un-assignable correlations prior to removing them.

In this version of CMC-se is an improved *Automatic analysis* dialog window (see following figure). This includes the possibility to keep already analyzed datasets (green marked datasets) through further analysis.

Results from previous analysis steps (including manual corrections, functional groups, fragment assignments etc...) will not be modified when adding and automatically analyzing new data sets. This allows the user to try running the structure generation with minimal data, and then if necessary come back at a later date and add additional data (^1H - ^{15}N HMBC, H2BC, ADEQUATE etc.) for better results, without having to undo the previous work. This also can be useful in cases of samples with impurities where the automatic analysis is less than perfect due to extra peaks in the spectra.



With this feature, it is possible to manually peak pick ([Adding and Removing Multiplets Through the Combined Spectra Display](#) ▶ 57) one or more of the spectra and then have the automatic analysis use those chemical shifts and correlations as a basis for automatically peak picking the remaining data sets.

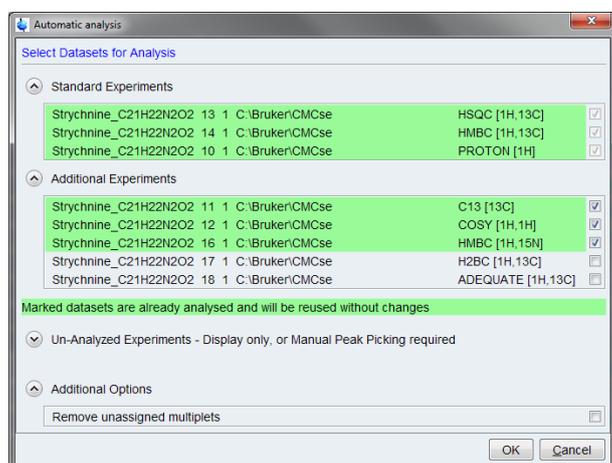


Figure 3.20: Automatic Analysis Dialogue box when data sets have previously been analyzed.

In the figure above, the green highlighted data sets will not be re-analyzed, only the H2BC and Adequate correlations will be picked based on the current correlations from the other datasets.

Deselecting an already analyzed dataset in the dialog would remove the existing correlations and multiplets.

For an analysis from scratch it is necessary to delete the current content of the Correlation Map. This is accomplished with the  button in the Correlation Table. The same function also allows selectively delete multiplets and correlations for chosen experiments, (see following figure.). It is not possible selectively delete already analyzed 1D-¹H, or ¹H/¹³C – HSQC, in this case all correlations must be deleted.

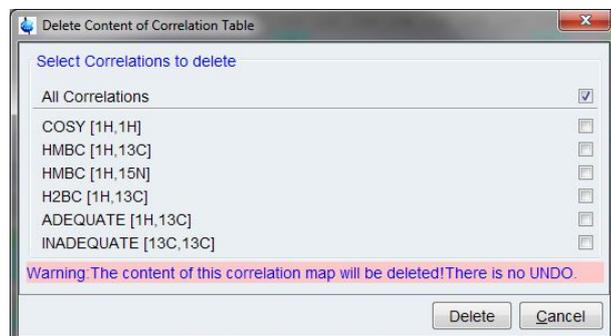


Figure 3.21: Delete Content of Correlation Map Options discriminate between “all correlations” or multiplets and correlations from selected experiments.

Clicking the **OK** button will start the automatic analysis. This process might need several seconds to a minute to complete depending on number and complexity of the spectra, and speed of the computer. During the analysis a pop-up window is displayed, which allows the user to cancel the process.

When the automatic analysis is finished, an analysis protocol is displayed as a PDF document. This can be inspected now or closed because it is saved in the CMC-se project directory, and may also be viewed at any time using the pull down menus **Analysis** → **Show Analysis LogFile** of the Correlation Table window. Additionally, a more detailed log file can be displayed using **Analysis** → **Show Detailed (technical) LogFile**.

After the automatic analysis is completed, the Correlation Table window itself will be updated with the identified atoms and correlations as shown in the figure above, and the analyzed multiplets and correlations are show on the spectra of the Combined Spectra Display.

3.4.1 Tasks Accomplished by the Automatic Analysis Feature

The automatic analysis for a structure elucidation project consists of the following tasks:

- Peak picking and multiplet analysis of the spectra.
- Collecting the information from all data sets.
 - The peak and multiplet lists are evaluated. One common HSQC and HMBC correlation list is built. Information from the 1D ¹³C spectrum is also used here when available.
 - Then the number of protons will be assigned to the C-atoms based on the integral value of the peaks in the HSQC.
 - The hybridization state of the C-atoms is assigned in most cases by using the chemical shift information.
 - Hybridization state and assignment of protons to X-nuclei like N or O is not completed automatically, and must be done manually if desired (see section [Adding Additional Information to the Correlation Table \[76\]](#)). Note however, that while it can improve the results of the generation process; it is not necessary to complete the hybridization state.
 - Symmetry detection is executed (i.e. identification of equivalent atoms/groups) based in the relative intensities of the 1D-¹H spectrum in conjunction with the HSQC and ¹³C spectra.

- Updating the Correlation Table.
 - As shown above in the figures, the results of the peak picking are automatically transferred to the Correlation Table.
 - It is possible that some correlations will be picked by the automatic analysis, but they can't be readily assigned to a given ^1H , or ^{13}C atom, and thus will be marked as an impurity in the Combined Spectra Display (or not shown at all, if the *Remove unassigned multiplets* option was checked) and not included in the Correlation Table. As discussed in the following sections of this chapter, there are a variety of methods available to add them to the Correlation Table after the initial automatic analysis (see sections [Actions Possible through the Combined Spectra Display \[52\]](#) and [Additional Modifications to the Correlation Table \[71\]](#) for more information).
 - Once the automatic analysis is completed, the header of the Correlation Table is also updated to include how many of the $^1\text{H}/^{13}\text{C}$ atoms in the molecular formula were assigned, the number of $\text{CH}/\text{CH}_2/\text{CH}_3/\text{XH}$ groups that were found, as well as the total number of constraints that were assignable (see figure in the section [The Correlation Table \[33\]](#)).
- Updating the Combined Spectra Display.
 - The picked peaks are now indicated in the Combined Spectra Display. More details about this window will be provided in the section [The Combined Spectra Display Window \[46\]](#).

For an effective structure elucidation it is important to have properly analyzed spectra. Although the automatic analysis is a powerful tool it is **strongly recommended** to inspect the results of the analysis carefully.

3.4.2 Typical Problems with Automatic Multiplet Analysis

The automatic multiplet analysis suffers from some typical problems especially for complex spectra or spectra obtained from non CMC-se parameter sets, or non-optimal processing. These can include, but are not limited to:

- Incorrect assignment of ^1H distribution (CH versus CH_3).
- Incorrectly assigned peaks due to spectral overlap.
- Calibration problems.
- Signals hidden in solvent regions.
- Peaks introduced by phase errors.
- Artifacts (COSY-like peaks in HSQC spectra, residual ^1J correlations in HMBC etc.).
- Folded peaks due to insufficient spectral widths used for data collection.

To avoid calibration problems and phase error introduced signals a careful processing of the spectra is strongly recommended. For more help see **Processing Commands & Parameters**, one of the TopSpin online manuals.

To eliminate incorrectly assigned correlations due to artifacts or to find and consider hidden signals the user has to carefully inspect the results from the multiplet analysis. The preferred method for doing this is with the CMC-se specific Combined Spectra Display, and described later (section [Actions Possible through the Combined Spectra Display \[52\]](#)).

While it is not possible for the software to detect all of these problems, the CMC-se module does have a built in **Consistency Check**. The  button at the top of the main project window will display the results of this basic consistency check. The results of which indicate potential problems, and provides the user a starting point for focusing their more critical

evaluation of the data described in the sections below. The results of such a consistency check are stored to the disk as a PDF file *ConsistencyCheck.pdf*. Highlights of this file are shown below.

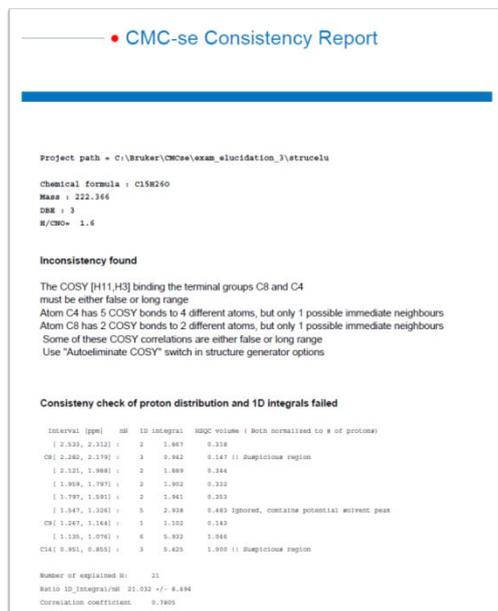


Figure 3.22: First page of the Consistency Check Report

The first section of this report is information about the project and molecule itself. Things such as project path, Chemical Formula, Molecular Mass, Double Bond Equivalent, and Proton : Hetero-Atom ratio.

The next section of the report is under the heading “*Inconsistency found*”. There is a standard ^3J limit for the COSY correlation (section [Chemistry Rules \[95\]](#)), thus a result of too many COSY correlations for a given atom, as shown above, indicates either long range or falsely assigned COSY correlations for that atom, and the user should inspect the COSY spectrum to possibly eliminate this contradiction manually if possible (section [Actions Possible through the Combined Spectra Display \[52\]](#) or [Additional Modifications to the Correlation Table \[71\]](#)), or in most cases with the Auto-elimination option (section [Correlations \[92\]](#)) as suggested in the report.

As mentioned earlier (section [Tasks Accomplished by the Automatic Analysis Feature \[41\]](#)), the ^1H distribution is derived from the integral value of the multiplicity edited HSQC. In cases of equivalent atoms, and or extreme overlap, this is sometimes not accurate, and mistakes can be made. The 3rd part of the consistency check is to determine a deviation between the assigned number of protons for a given region, and the integral value from the 1D- ^1H spectrum. When there are potential problems as shown above, the results of this check are under the section “*Consistency check of proton distribution and 1D integrals failed*” These following lines simply compare integral regions (defined in the first column) with the number of protons assigned to that region (2nd column) the integral value from the 1D- ^1H spectrum (3rd column), and the volume from the HSQC spectra for that same region (4th column). Note in cases of HSQC peaks overlapping (in F2), the regions are grouped together and treated as one large region. When the number of assigned protons does not match the integral values, it is noted as a “Suspicious region”. Also illustrated in the figure above is the fact that this consistency check does ignore regions of the ^1H spectrum that contain solvent peaks and common impurities (such as H_2O in DMSO).

The next section of the report contains this proton distribution check information as a graph as shown in the figure below. Most of the peaks fall on a linear line where the x-axis of “Number of Protons” correlates well with the “1D integrals”. However there are multiple inconsistencies with the peaks which were assigned to 3.0 protons.

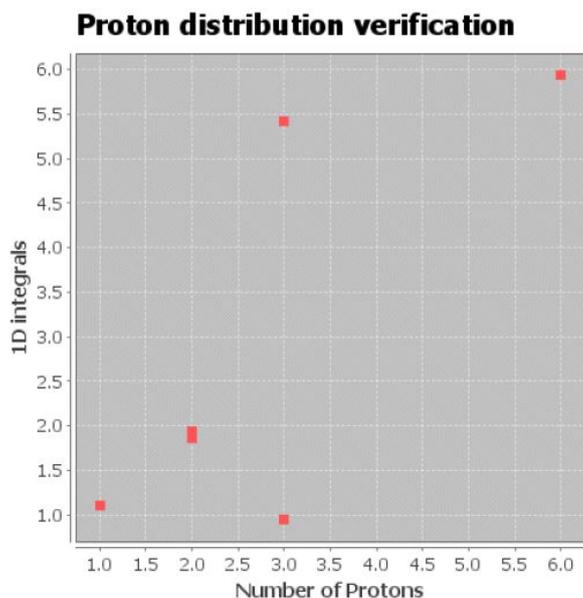


Figure 3.23: Graphical representation of a potential failed analysis of proton distribution.

These problematic regions are also highlighted on the Combined Spectra Display:

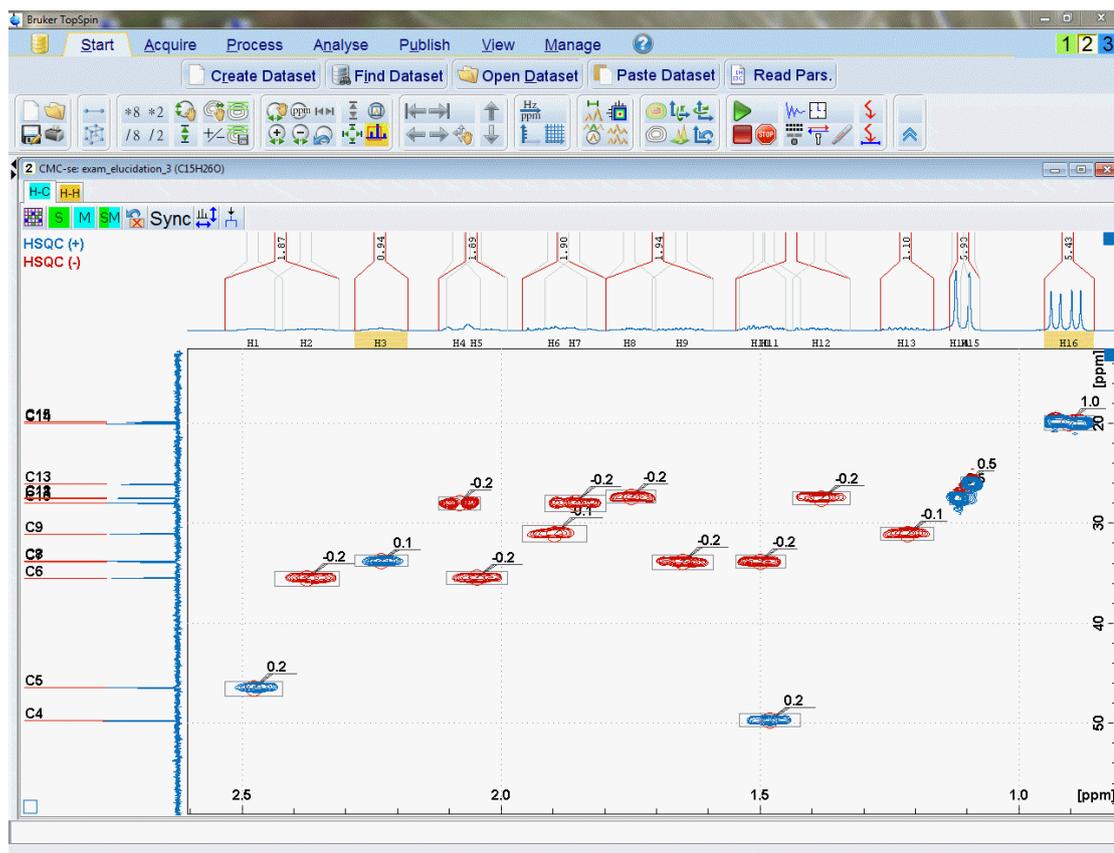


Figure 3.24: Combined Spectra Display highlighting the failed ¹H Integral consistency check

As illustrated in this figure, the automatic analysis integrates the HSQC peaks (integral volumes shown in 2D spectra), and then also integrates the 1D spectra of this same region (grey integral brackets in ^1H projection). The 1D integral values that are used for the consistency check are shown in red. Note that in instances where multiple HSQC peaks (grey brackets) overlap, the red bracket combines multiple ^1H resonances into one region. It is this red bracket value that is compared with the number of assigned protons in the table (column 3 and column 1 respectively) to obtain the "Correlation Coefficient". As seen in this figure, the "suspicious regions", are highlighted in yellow in the Combined Spectra Display to aid the user in resolving these inconsistencies.

Note that areas surrounding those contained in the solvent base (solvents and known impurities) are ignored for accuracy reasons. Thus, these highlighted regions are not the only places there might possibly be errors. In this particular example (exam_CMCse_2 as described in [Examples \[131\]](#)), there are also errors in the region around 1.5 ppm, which is close to where H_2O would be in the CDCl_3 of this example.

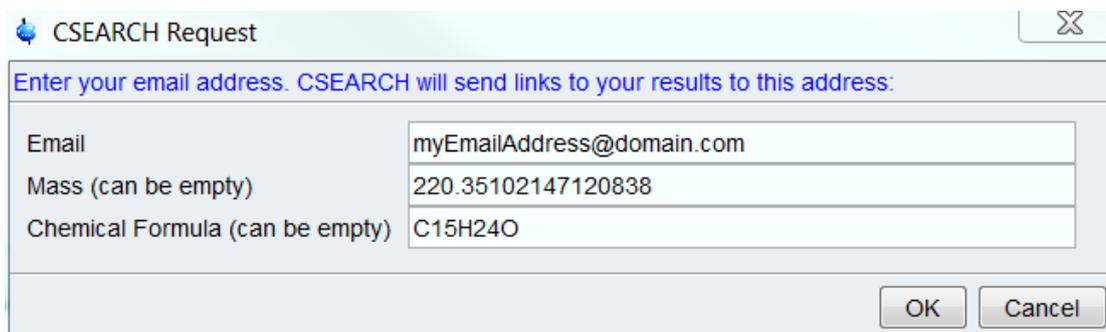
The reliability of such check is directly related to the quality of 1D proton spectra. Using extremely short relaxation times makes the 1D spectrum not quantitative. This is the reason the CMCse_1H parameter set uses a D1 of 2 seconds, and in some cases, this may not be long enough for such analysis to be accurate. Additionally, any phasing errors or baseline distortions in the proton spectrum make such analysis less reliable.

Also, in cases where the sample under investigation contains impurities, these additional signals which overlap with resonances from the molecule of interest can also cause false failures in the consistency check.

3.5 CSEARCH Structure Identification

After automatic spectra analysis there is the possibility to match result with compounds in the CSEARCH database. The search based on $\text{C}13$ chemical shifts and uses additional information about hybridization, chemical formula and the mass. Start the request by pressing

the button , CMC-se automatically generates a request file with the given information (see the following figure):



Enter your email address. CSEARCH will send links to your results to this address:	
Email	myEmailAddress@domain.com
Mass (can be empty)	220.35102147120838
Chemical Formula (can be empty)	C15H24O
<input type="button" value="OK"/> <input type="button" value="Cancel"/>	

Figure 3.25: CSEARCH request dialog

In a next step the default email program starts with an automatically filled email. Following the instruction in the email text, it is necessary to attach the data file located in the given file path. With the Control V key the filename entry can be pasted in the mail client attachment browser, because CMC-se copied the filename into the clipboard. The email subject is always the project name, ending with a number coding the current date and a time.

The result will be send to the given email address after few minutes. There could be a handful of emails including links to possible results. The following figure shows an example for the structure search:

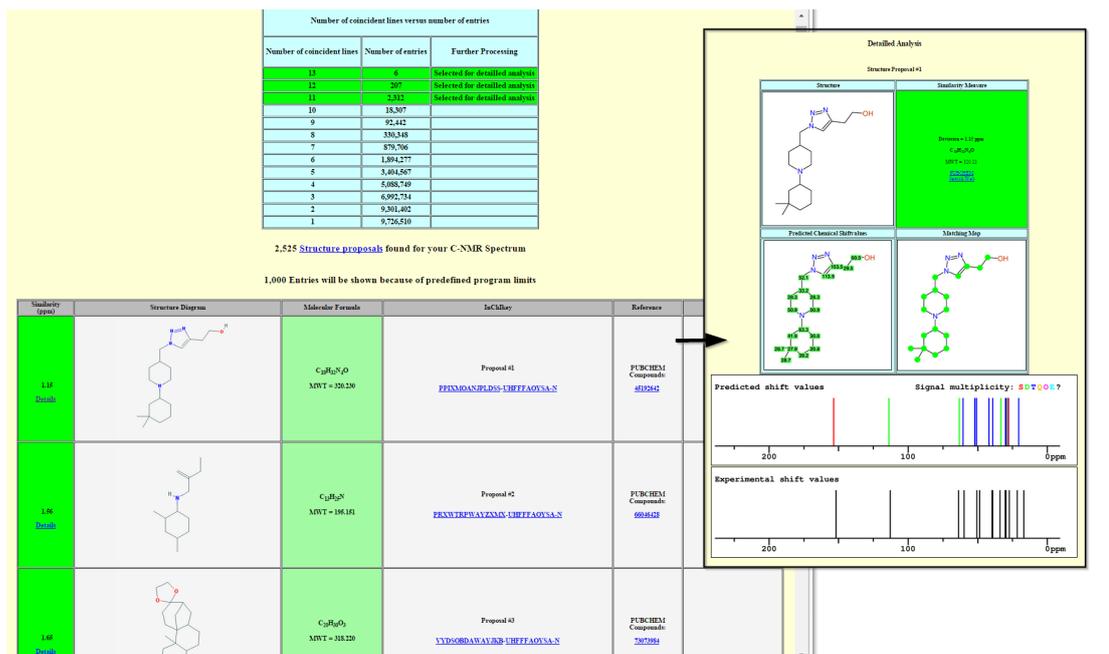


Figure 3.26: CSEARCH result example

For more information about CSEARCH use the following link:

<http://nmrpredict.orc.univie.ac.at/>

3.6 The Combined Spectra Display Window

It is recommended practice to inspect the Correlation Table and the analyzed spectra using the Combined Spectra Display shown in the figure below. This is a strong recommendation because this CMC-se specific module was designed to contain and work with all of the data together, instead of on an individual basis. All peaks that are picked and included in the Correlation table are highlighted in the Combined Spectra Display, and as shown below, these two windows are directly linked:

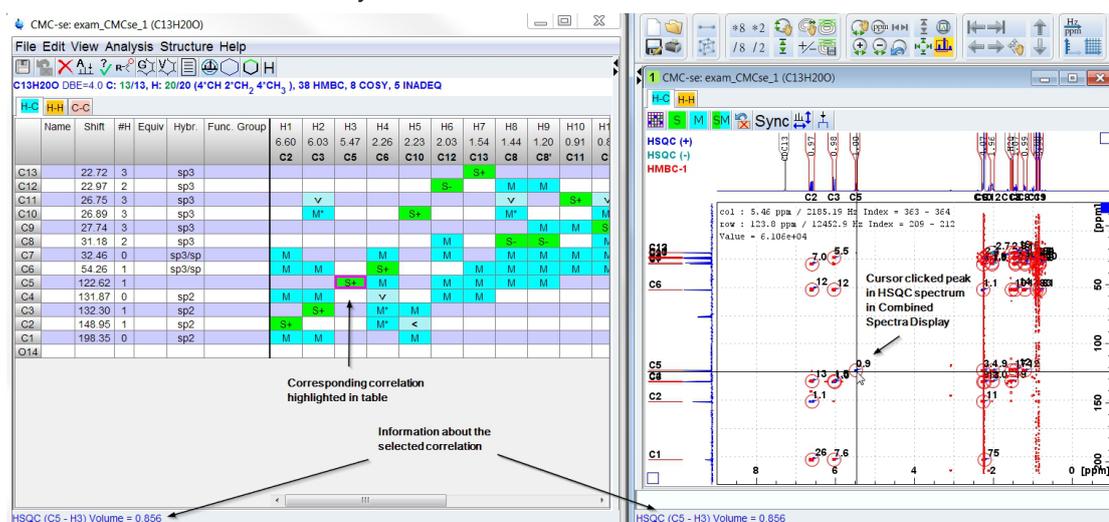


Figure 3.27: Combined Spectra Display (right), illustrating how the cursor is tied the Correlation Table (left).

Due to the importance of this Combined Spectra Display window, whenever a CMC-se project is open, both the Correlation Table and this Combined Spectra Display window will automatically be opened. Closing one with the  button will automatically close the other.

If the Combined Spectra Display is ever minimized, or behind another window in TopSpin,  in the

toolbar of the CMC-se Project window brings the Combined Spectra Display to the forefront in the main TopSpin window, and maximize the TopSpin window if necessary. Conversely the  button of the Combined Spectra Display will bring the Correlation Table to the forefront if it is ever minimized.

One of the most useful aspects of the Combined Spectra Display window is that it is linked to the Correlation Table. This provides a convenient and highly useful mechanism for not only viewing the correlations that are used for generating structures, but also modifying them as necessary.

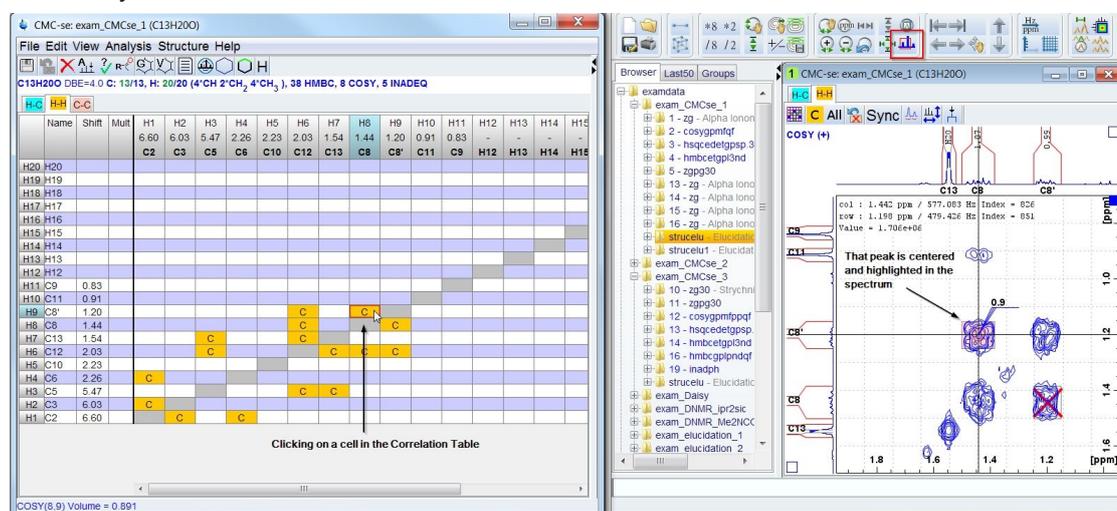


Figure 3.28: Clicking on a cell in the Correlation Table will result in that peak being highlighted in the data.

As shown in the figure above, if the **Keep Zoom** option is not selected, so the COSY peak is zoomed in on and centered in the window. If the **Keep Zoom** is selected in TopSpin () the zoom of the Combined Spectra Display does not change, but the cursor will move to that peak (visible only for peaks in the current zoom level).

The two windows are also tied together in such that clicking on one of the tabs will automatically jump to that tab in the other. For example, as shown in the figure above, the cursor is on the H-H tab in the Correlation table, and the Combined Spectra Display automatically jumps to show the H-H Spectra.

Part of the powerful utility of the Combined Spectra Display is the ability to not only work with the data in combination with the Correlation Table, but also with all of the different spectra at once. Thus it is possible to toggle the layout to switch it back to a layout that is similar to the one used in the initial version of CMC-se, where all the data is on a single screen as shown in the figure below. Just like toggling the layout of the Correlation Table (see [The Correlation Table \[33\]](#)), this is accomplished through the **Toggle Layout** option of the **View** pull down in the CMC-se Project Window.

When in this view layout, by default, the upper left hand spectra contains an overlay of the $^1\text{H}/^{13}\text{C}$ HSQC and HMBC. The lower left hand spectra is the $^1\text{H}-^1\text{H}$ spectra, and on the right hand side of the window are any additional spectra (in individual tabs) that were included in

the project, shown in the figure below is the ^1H - ^{15}N HSQC/HMBC and ADEQUATE. Also as indicated, when present the high-resolution 1D spectra are always included as projections on the relative spectra.

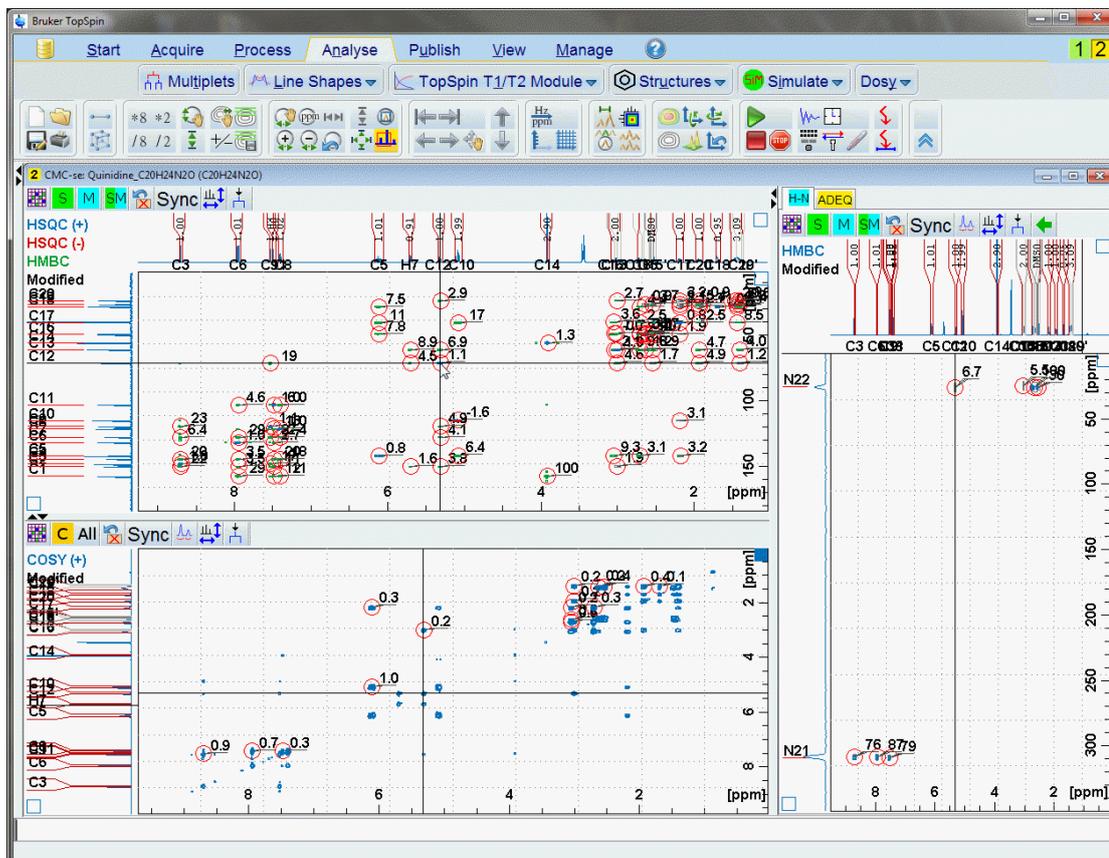


Figure 3.29: Screen shot of the Combined Spectra Display in the “old” layout. In this mode, more types of spectra are visible at once.

The bars with the ▲▼ and ►◀ icons can be used to manipulate the size of the various spectra windows, or close individual windows completely.

Also, the  icon can be used to move the spectra in the top tab of the right hand side of the window, to the lower left hand side of the window (beneath the HSQC/HMBC section) as shown in the figure below. This allows for better examination of certain spectra in combination with the HSQC from which the skeletal atoms are mainly defined.

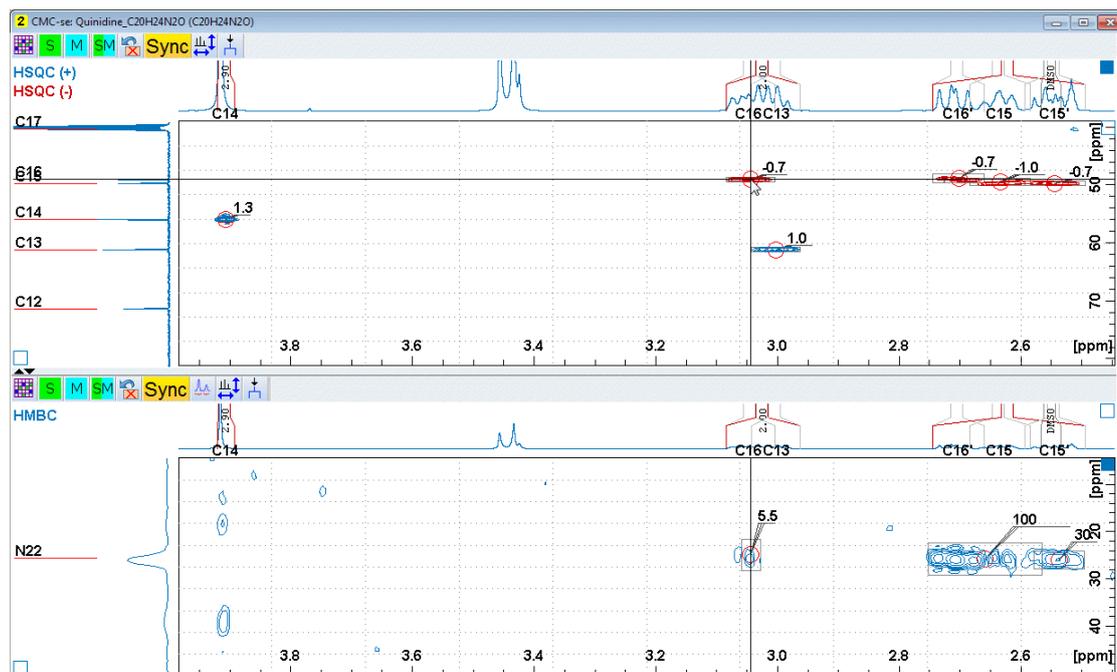


Figure 3.30: Illustration of how the “Old layout” can be modified to allow closer examination of multiple types of spectra at once, but yet simplified so as not to be viewing too much information at once.

As alluded to above, and will be described later (section [Manipulating Viewing Options of Combined Spectra Display \[50\]](#)), the exact display of these windows can be modified by the user depending on what exactly is necessary to be seen for the given task.

Regardless of which layout is used, there are multiple reasons for using this Combined Spectra Display window during a Structure Elucidation Project:

- All correlations picked by the automatic analysis are displayed in the Combined Spectra Display including those determined as impurities, and not included in the Correlation Table. While the picked peaks might be visible outside this CMC-se specific window, the differentiation is not. In the Combined Spectra Display, picked peaks are represented with a grey box around the multiplet.
 - Those peaks that are included in the Correlation Table are indicated by a circle .
 - Those not assignable (and thus not in the Correlation Table), or manually disabled in the table are marked with a red cross .
 - A peak that is picked manually (section [Adding and Removing Multiplets Through the Combined Spectra Display \[57\]](#)) will be noted with a dashed red circle .
- The signals in the spectra of this Combined Spectra Display window are tied not only to each other (e.g. COSY to HSQC) via the correlated cursor, but also more importantly, to the Correlation Table. As shown above in the first figure, clicking on a correlation in the table moves the cursor to that location in the Combined Spectra Display. Similarly, moving the cursor around in the Combined Spectra Display highlights the corresponding correlations in the table.
- This Combined Spectra Display window is also still correlated with all other TopSpin windows, so evaluation of this window in the context of other windows which may or may not be included with the CMC-se project is still possible.

- As will be described in the section [Actions Possible through the Combined Spectra Display](#) [52], manual modifications to picked peaks in the displayed in the Combined Spectra Display, are automatically updated in the Correlation Table.

3.6.1 Manipulating Viewing Options of Combined Spectra Display

In the first figure of section [The Combined Spectra Display Window](#) [46], the Correlation Table and the Combined Spectra are arranged side by side. This is only one possible way to arrange the combined data view. Each user can choose their preferred arrangement using the **Arrange Frames** option under the **View** menu of the Project Window, or by manually moving the windows on the screen.

As mentioned previously when in the old layout the relative sizes of the individual windows can be modified as necessary using the arrow icons, and dragging the horizontal bars in the

middle with the  arrows as necessary, and the  icon can be used to move the various spectra around in these 3 windows.

While the overlay feature is often necessary to evaluate the contents of the various, depending on the task, and or complexity of the data, this might not be desirable. It is possible to toggle each of them off and on by pressing the corresponding buttons in the toolbar of the spectra window. In all cases, the buttons only affect what is displayed in the 2D spectra, the high resolution projections remain unchanged.

For example, in the illustration of the old layout in section [The Combined Spectra Display Window](#) [46] above, and the figure below, the HMBC correlations have been toggled off, and the only HSQC correlations are shown:

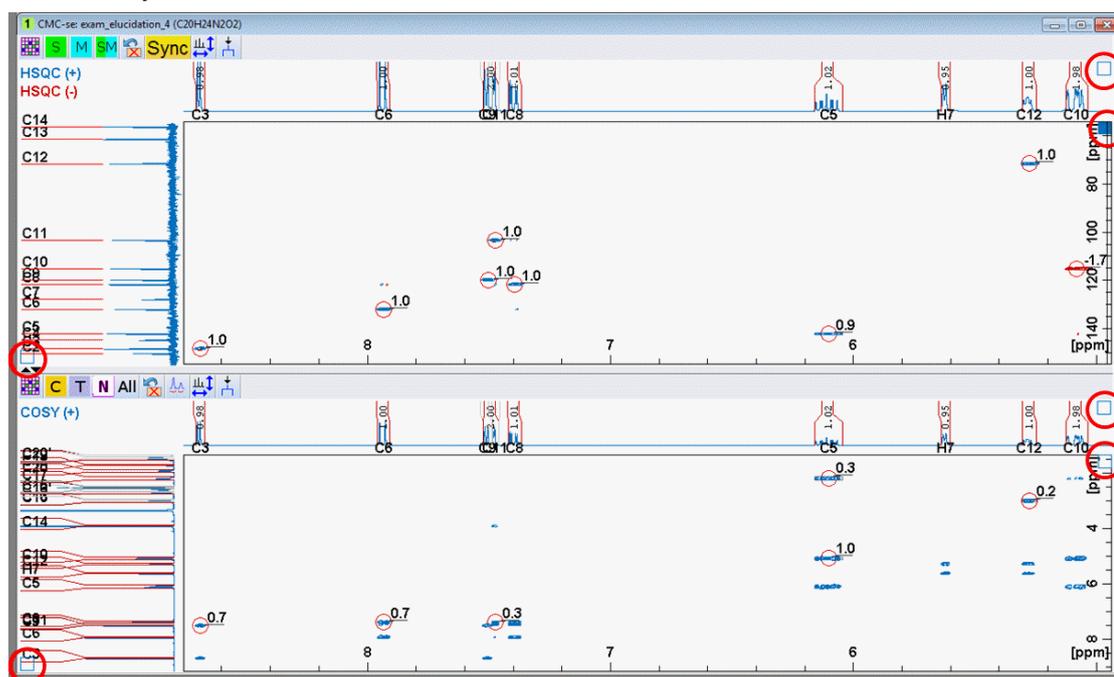


Figure 3.31: Combined Spectra Display window enlarged to highlight features used for manipulating the display options.

 - Displays ONLY the HSQC spectrum.

 - Displays ONLY the HMBC spectra (all that are included with the project).

M2 - Displays ONLY the 2nd (selective) HMBC spectrum if available.

SM - Displays BOTH the HSQC and HMBC spectra.

When included with the project, the following buttons are used to modify the display of the additional experiments when they are included with the project.

C - Displays ONLY the COSY spectrum.

T - Displays ONLY the TOCSY spectrum.

N - Displays ONLY the NOESY or ROESY spectrum.

All - Displays ALL ¹H-¹H spectra that are included in the project.

B - Displays ONLY the H2BC spectrum.

B+S - Displays the overlay of the HSQC + H2BC.

A - Displays ONLY the ADEQUATE spectrum.

A+S - Displays the overlay of the HSQC + ADEQUATE.

T - Displays ONLY the HSQC-TOCSY spectrum.

T+S - Displays the overlay of the HSQC + HSQC-TOCSY.

The manipulation of the spectra in the dual display window is very similar to that of a regular TopSpin window, in terms of zooming/expanding and adjusting contour levels. Refer to the TopSpin manual for more information if necessary. Similar to regular TopSpin windows of 2D data sets with projections, control over which spectrum is being modified (**Zoom In, Show All** etc.) is indicated by the filled in square of either the projection or the 2D spectrum. These squares are indicated by a red circle in the figure above. An open square indicates an inactive spectrum.

For example, in the figure above, the window is set up to modify the zoom/scale of the HSQC. To modify the COSY spectrum, the COSY portion of the window would be clicked on, filling in the blue square.

To modify the projections of the ¹H/¹³C experiments the square of the upper window must also be checked in addition to the appropriate projection (¹H, ¹³C or both), while modification of the ¹H projections on the COSY spectrum require the square button of the COSY spectrum to be checked in addition to the appropriate projection.

If the ¹³C projection has the filled in square as well as the COSY, the ¹³C projection will not be modified until the HSQC is clicked, and that square is filled in instead of the COSY.

Note that when in the “new” layout and only one type of spectra is shown in a given tab, this is automatically set for the user.

To adjust the contour levels of the one of the overlaid spectra but not the other, for example, adjust the HSQC, but leave the HMBC un-modified:

- Make sure the square of the ¹H or ¹³C projections is **not** filled in but that of the 2D spectrum **is** filled in.

- Use the  button so that only the HSQC is shown, and then adjust the levels of the HSQC, via the middle scroll wheel or the respective buttons in the TopSpin window.
- Alternatively, use the  button to adjust the HMBC contours.
- Then use the  button to return to the overlay view.

The  button in the COSY window toggles the horizontal 1D proton projection on and off. In the figure above, this has been turned on.

The zoom level of multiple spectra can be synchronized with the **Sync** button. When this button is highlighted yellow  (as shown in the figure above), all 2D spectra are synchronized, and zooming in on one will result in the other(s) zooming to that same ^1H region. To zoom them independently, uncheck the  button to be grey.

When in the “new” tab layout, this button is still present, and has the same effect of controlling the zoom of the spectra on the other tabs, and zooming to a desired area on one spectra and then switching tabs will result in viewing that same synchronized zoom, just as if the spectra were displayed at the same time.

Right clicking in any 2D spectra of the Combined Spectra Display window (but not while cursor is on a peak) brings up an additional menu of display options:

- *Toggle Peak Grid Display* will toggle on and off the ^{13}C projection lines.
- *Toggle Annotation Display* will toggle on and off the annotations for the 2D cross peaks (section [Bookkeeping Tools within the CMC-se Module \[78\]](#)).
- *Toggle Volume Display* will toggle on and off the displayed peak volumes.
- *Toggle Cursor Info Display* is a convenient way to toggle on and off the cursor information displayed in the upper left hand corner just as in a normal TopSpin window.
- *Clean Multiplets* which can be used to clear the peaks which have been picked, but could not be assigned to the table. Note that this can also be done if desired when the automatic analysis is first performed.

3.6.2 Actions Possible through the Combined Spectra Display

In addition to being a convenient tool for examining the all the data used in the structure elucidation project, there are multiple actions possible through the Combined Spectra Display that will improve the results of the structure generation process. These are described in the following sections.

3.6.2.1 Modifying Existing Multiplets

Using the Combined Spectra Display it is possible to inspect and modify all signals in the spectra. Right clicking on a picked correlation peak will open the following pop-up window:

Properties ...
Resize multiplet box
Disable or enable current correlation
Is folded in F1
Is inverted
Normalize Volumes ...
Delete

Figure 3.32: Peak-actions pop-up window

All of these options will modify the current correlation in the following way(s):

- *Resize multiplet box* – the box shape changes in a selection box allowing the user to change its size and position. The frequency of corresponding atoms (e.g. HSQC multiplet), or atom assignments are changed accordingly.
- The multiplet can be resized by selecting this option, and then dragging the sides of the box accordingly (center image below). Alternatively, it can be moved by clicking and dragging from the center of the box (right image below).

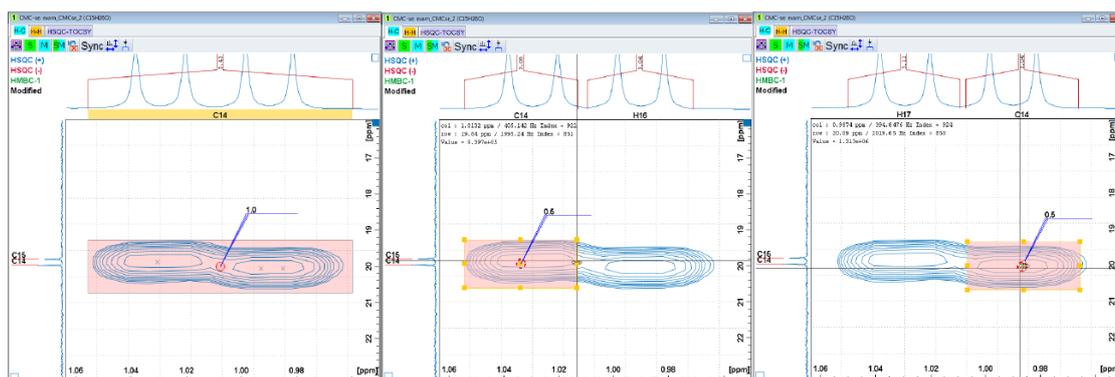


Figure 3.33: Resize multiplet box function to change box size or position.

- *Disable or enable current correlation* -- Changes the icon of an active peak to  and disables that correlation in the Correlation Table (section [Disabling Correlations in the Correlation Table \[74\]](#).) If any correlations are suspected to be incorrect or long range, it is advisable to *Disable* them. They are still picked, and can be used for evaluation purposes (Section [Viewing Correlations on Structures \[110\]](#)), but they are not used as constraints in the generation process.



The *Disable or enable current correlation* functions acts as a toggle, and if the selected peak is currently not assigned in the table, selecting this action will try to add the signal to the Correlation Table.

For example, if the *Remove Unassigned Multiplets* option was not checked in the Automatic Analysis Step (section [Automatic Spectra Analysis \[39\]](#)), or deleted manually as described above, and the peaks are picked but just not assigned, it is possible to use this *Disable or enable current correlation* toggle to add correlations to the table as seen in the figure below.

However, this should be used with caution, and not every peak should be added to the table unless there is good reason. As discussed the section [Correlations \[92\]](#), there are rather strict definitions to correlations (COSY = 3J correlation, HMBC = 2 or 3J), and **weaker peaks are often longer range correlations and will actually be detrimental to the generation process.**

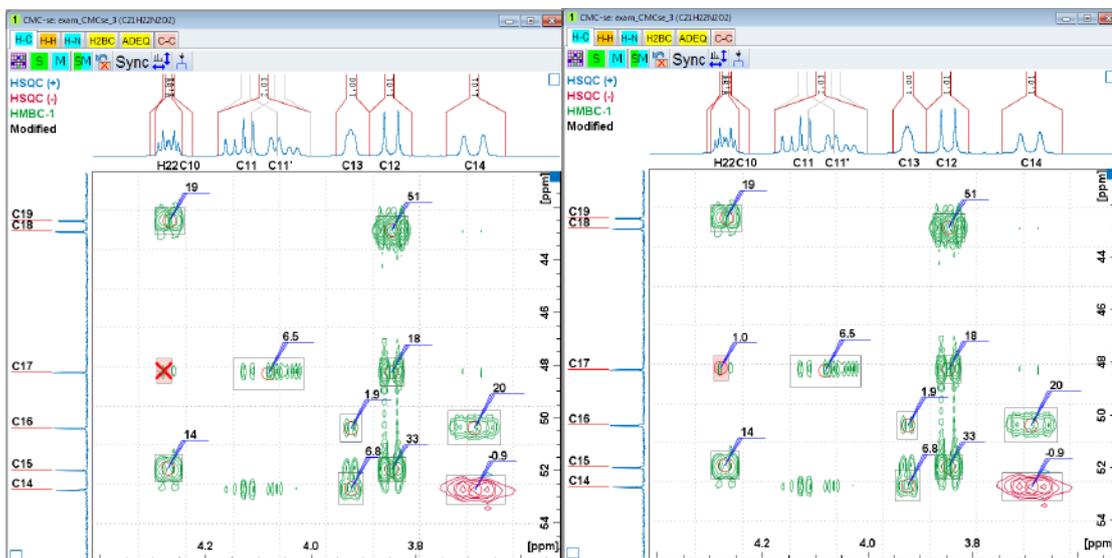


Figure 3.34: Using the “Disable or enable current correlation” to add a currently non-assigned correlation to the table.

The *Disable or enable current correlation* will not only add the correlation, but also the corresponding atom, if there is room in the table:

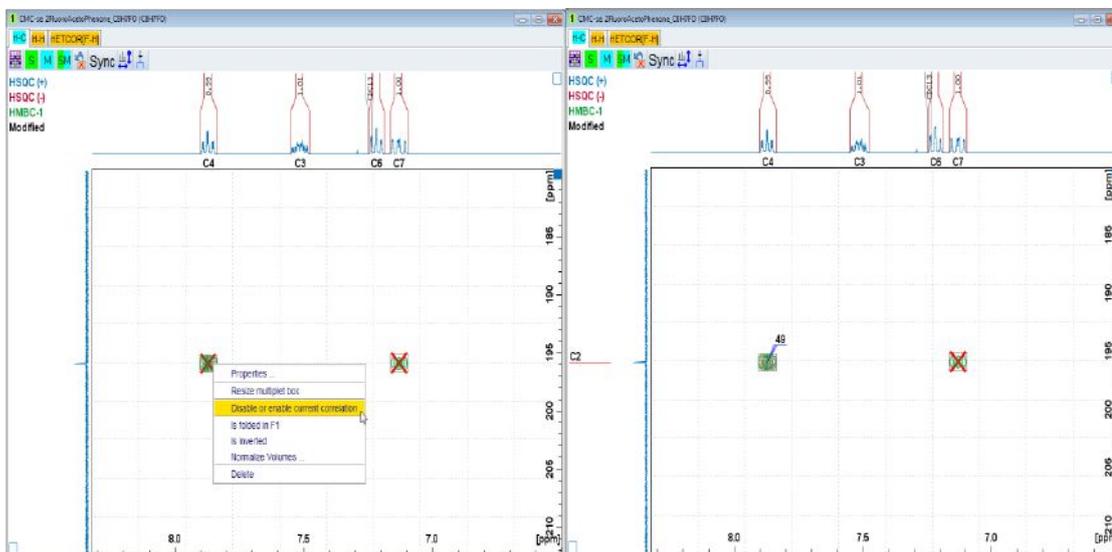


Figure 3.35: Right clicking to enable a picked, but not assigned correlation will add both the atom and the correlation to the table.

If the correlation cannot be assigned by the software, the following pop up message will appear. When this occurs, and the selected peak is truly part of the molecule, it is possible that current contents of the Correlation Table that are incorrect need to be deleted. Or the correlation can possibly be added to the table manually (such as the case for an HMBC or

COSY correlation that does not define an actual atom like an HSQC correlation does). See section [Additional Modifications to the Correlation Table \[71\]](#) for more details on modifying the correlation manually.

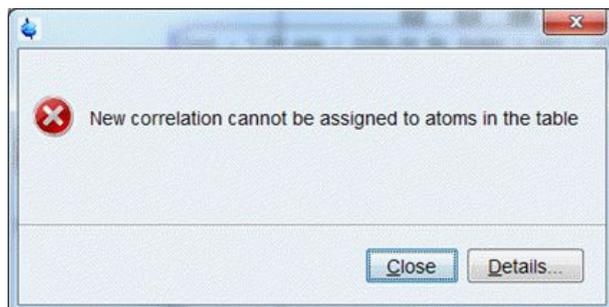


Figure 3.36: Pop-up window indicating that the selected correlation can't be assigned to the current Correlation Table.

Other actions from this pop-up menu after right clicking on a correlation will modify that chosen signal as follows:

- Identified as *folded in F1* – If a peak in an HSQC or HMBC spectrum was outside the SW of the experiment, CMC-se can automatically calculate the correct chemical shift of this correlation, and update the accordingly. After being identified as folded in F1, in the spectrum the peak will no longer be identified as an impurity, and will have the following icon:



- Identified as *Is inverted* – As stated in section [Data Set Processing \[19\]](#), the HSQC should be phased such that signals from CH and CH₃ groups are positive and CH₂ groups as negative. In instances where this is not possible (two examples would be, an unusual coupling constant, or a non-edited HSQC was used), a peak can be designated as “inverted” and the software will evaluate the integral value as such, and define the attached proton count accordingly.
- The *Normalize volumes* option allows the user to set a particular integral value to the selected peak, and scale the remaining peaks in that spectrum (HSQC, or HMBC, or COSY) relative to that value.
 - Note during the automatic analysis the volumes of the various correlations automatically normalized as follows:
 - HSQC: mean volume of one proton correlation is 1.
 - HMBC: the most intense correlation has a volume of 100.
 - COSY: the most intense correlation has a volume of 1.
- *Delete* – this action removes the selected peak from the correlation table (if it was present), and no longer “picked” in the Combined Spectra Display.
- The *Properties* option, brings up a window (see following figure) showing the multiplet properties of the correlating signal, and more detailed information pertaining to the signal of interest:

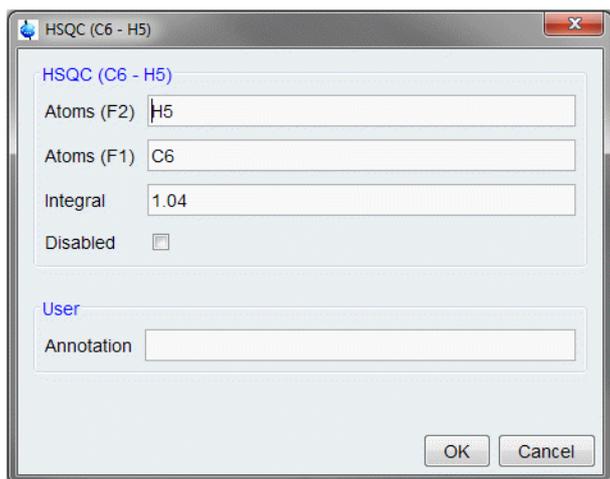


Figure 3.37: Multiplet Properties window for an HSQC correlation.

As will be described further in sections [Additional Manipulations to Correlations in Table \[75\]](#) and [Bookkeeping Tools within the CMC-se Module \[78\]](#) below, if desired, a user defined annotation can be added to the signal, which will automatically be displayed on the spectra in this Combined Spectra Display window.

Because an HSQC correlation relates directly to the ^1H and ^{13}C atoms, *double clicking* on an HSQC peak in the Combined Spectra Display will bring up an additional window where Atom Properties can be modified instead of just the multiplet properties. This window is shown below:

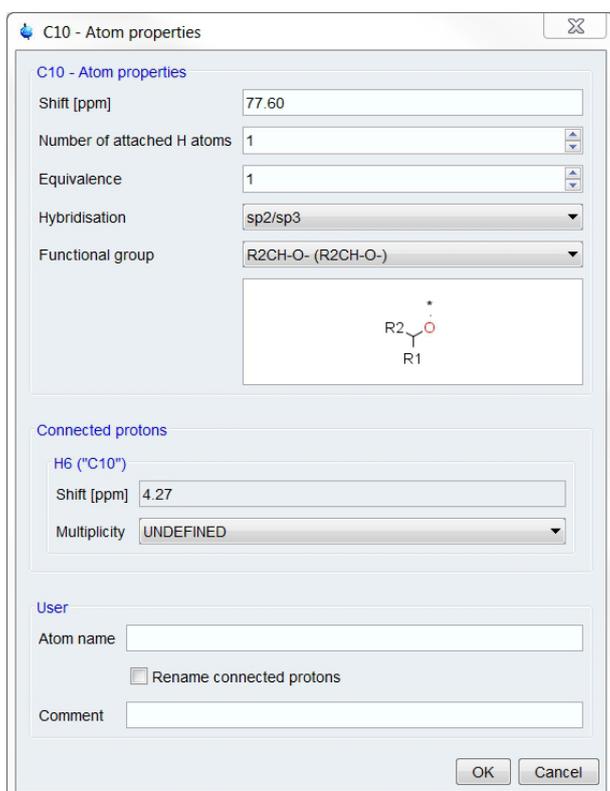


Figure 3.38: The Atom Properties Dialog Box

The Atom Properties dialog box shows several important pieces of information regarding the atoms in the Correlation Table can be either modified or added. More information on this box is found in [Adding Additional Information to the Correlation Table \[▶ 76\]](#).

The information that is entered or modified here are things that would be found in the Correlation Table such as

- Chemical shift
- Number of attached H atoms
- Hybridization of the atom
- Functional group
- Multiplicity of the proton
- Atom Name and/or Comment

Because these entries are found in the Correlation Table, and not directly on the Combined Spectra Display, they are explained in more detail in [Adding Additional Information to the Correlation Table \[▶ 76\]](#). However, sometimes modifying these entries is more reasonable when looking at the data and thus accessing it through the double click action describe above can be useful.

3.6.2.2 Adding and Removing Multiplets Through the Combined Spectra Display

It is possible, especially in cases of signal overlap or additional signals from contaminants that not all correlations will be suitably picked by the automatic analysis and should be (re)-picked and added to the Correlation Table manually.

To explain a manual addition to the Correlation Table an example is described below.

This is a common case of overlapping peaks in the HSQC. As shown in the figure below, from the ^{13}C spectrum the 15th carbon was found, and is included in the table however, the automatic analysis could not assign a proton to this carbon atom (top of the figure). The bottom of the figure is the Combined Spectra Display with the HMBC correlations toggled off and expanded to the region likely to find the missing proton in the HSQC (see section [Manipulating Viewing Options of Combined Spectra Display \[▶ 50\]](#)). As seen in the figure, the ^{13}C spectrum shows clearly separated signals, but the HSQC was not resolved enough for the automatic analysis to detect these as two peaks. Also, the ^1H projection for this region is highlighted yellow indicating an inconsistency found with the ^1H distribution consistency check.

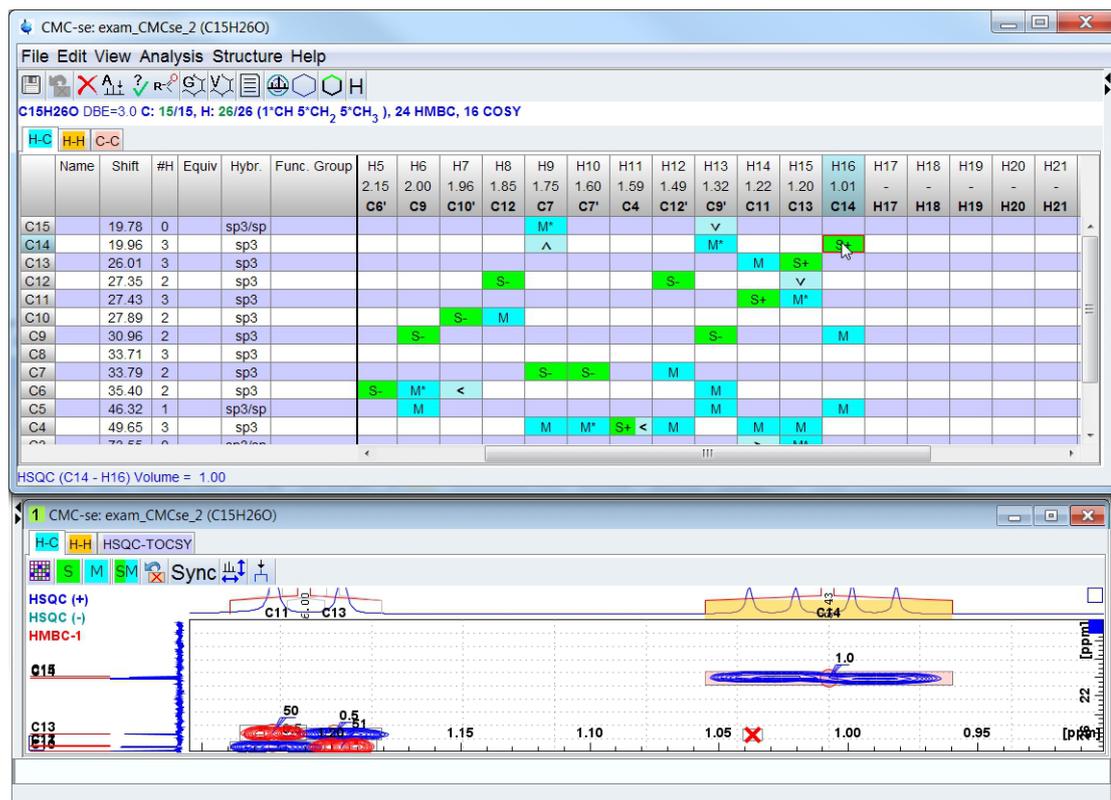


Figure 3.39: Combined Spectra Display prepared for manual addition of correlations to table

To correct this, it is necessary to delete this incorrect correlation first. Right click in the grey rectangle to display the pop-up window shown in [Modifying Existing Multiplets \[52 \]](#) is displayed. Click **Delete** to delete this current correlation and the grey rectangle disappears, as well as the cell in the Correlation Table (see the next figure).

Note that this deletion, only deleted the HSQC correlation in the table, not the rows and chemical shift assignments for the ^{13}C atoms, or the HMBC correlations that had been assigned to that proton. This was the desired effect in this situation as the Carbon atoms in the table are both correct. An example is shown in figures (section [Deleting Atoms from the Correlation Table \[72 \]](#)) below where the entire atom (Row/Column in the table, not just the HSQC Correlation) is incorrect.

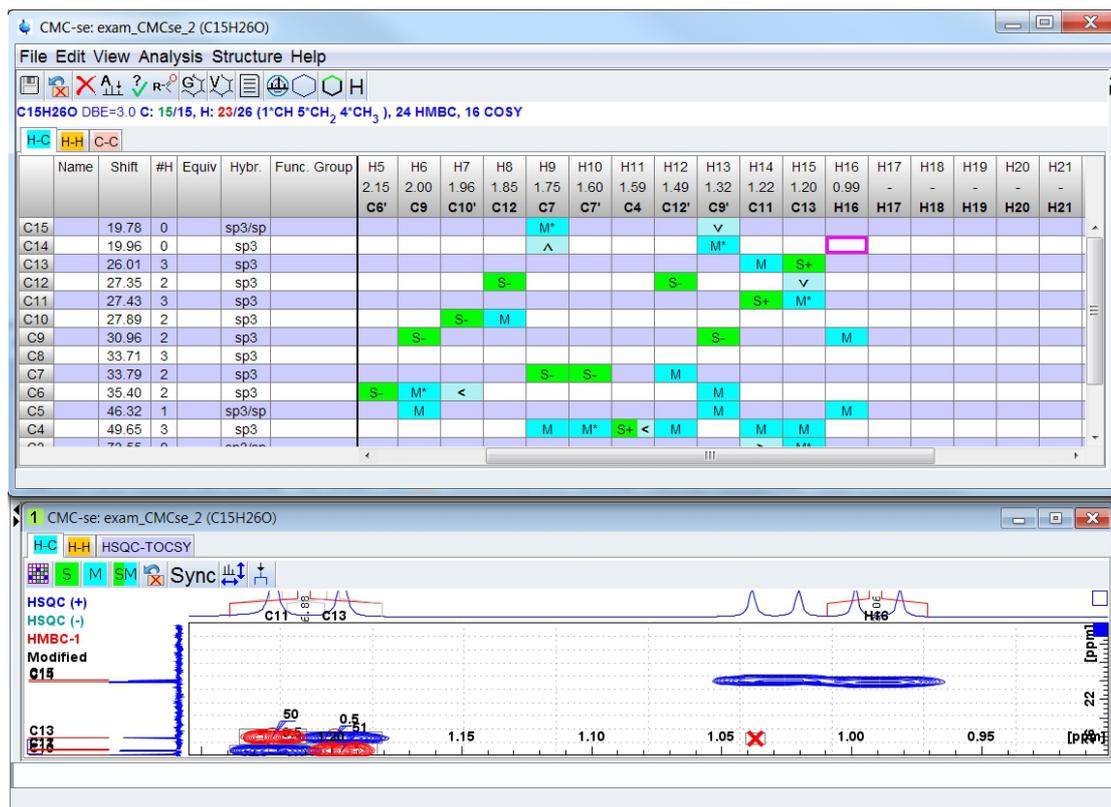


Figure 3.40: Correlation Table, and Combined Spectra Display, after deleting the incorrect HSQC correlation.

Rather than deleting the correlation, it is also possible (as described in the figure above) to use the *Resize multiplet box* to change the size and position of the existing multiplet. The corresponding cell **S+** in the Correlation Table will be updated accordingly.

Either after deleting, or resizing this multiplet, a new multiplet needs to be defined for the other HSQC or both of the HSQC correlation(s).

Click the  button from the toolbar of the Combined Spectra Display (or right click and select the *Define Multiplet* option), and the two black lines for region definition change to red colored ones (as compared to black as shown in the figure above). Click and drag with the left mouse button and define the region of the first desired multiplet.

If more than a single type of spectra are shown, the *Select multiplet type* pop-up is displayed (see figure below). Selecting the HSQC option results in the updates to the Correlation Table and Combined Spectra Display as shown in the figure below.

NOTE that if only the HSQC spectra is visible, as shown in Figure above, the software will not prompt for which type of correlation, as it already knows that it should be an HSQC.

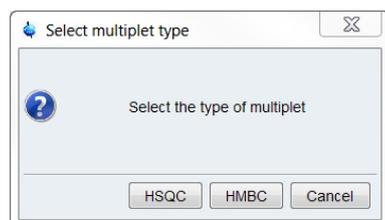


Figure 3.41: Manual definition of a multiplet in the Combined Spectra Display window

Now that this peak has been defined in the Combined Spectra Display, it is also automatically updated in the Correlation Table with a new **S+** in a new column of the table (^1H # 17 at 1.03 ppm).

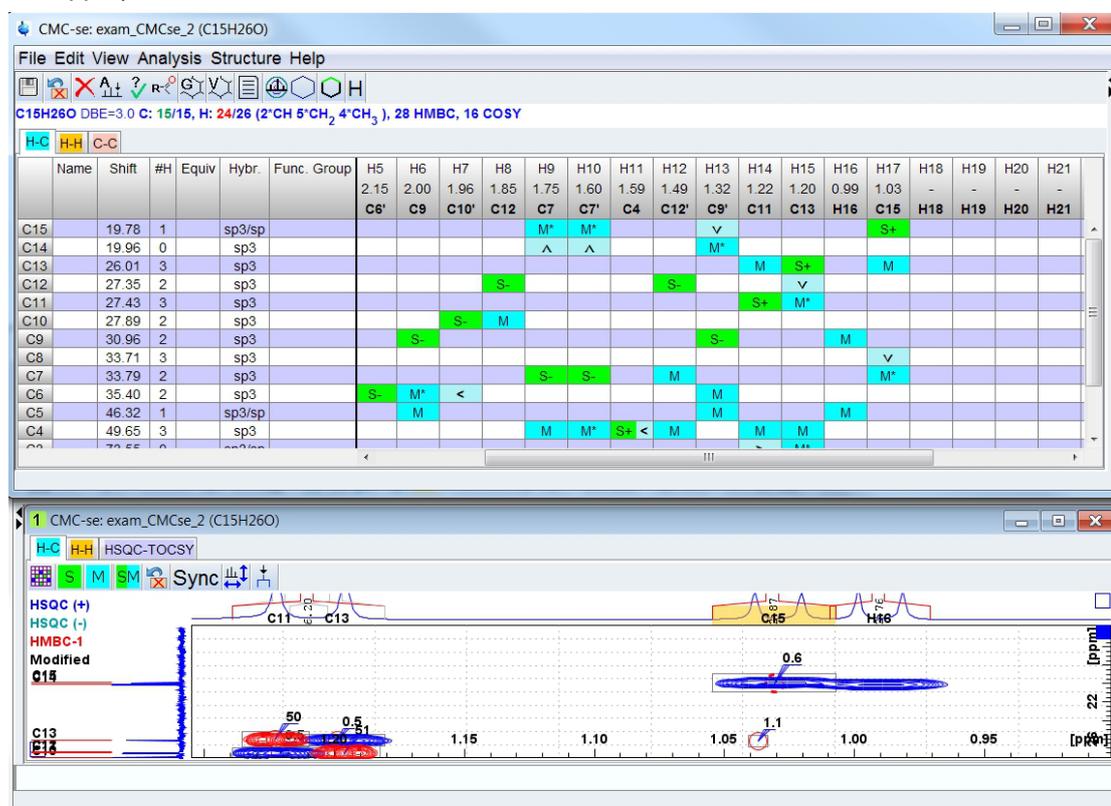


Figure 3.42: Correlation Table, and Combined Spectra Display after manually adding the HSQC peak that corresponds to C15 that was previously missing.

Because it is a new proton chemical shift to the table, it has been added to the right of existing columns, this can be adjusted at any point by selecting the *Resort atoms* option from the *Edit* menu, which will sort the columns back into numerical order based on the chemical shift of the protons.



Also because this new proton chemical shift has been added to the table, previously unassignable HMBC correlations can now be automatically assigned to this new ^1H . Part of adding a new HSQC correlation is to search the un-assigned correlations to see if they can now be assigned to the new signals.

However, please note the presence of the yellow “warning” on the ^1H projection indicating that the ^1H count is not consistent between the 1D and the HSQC. This is indicative that the attached proton count for the ^{13}C atoms is incorrect. This will need to be fixed, and is described in the section [Adding Additional Information to the Correlation Table](#) [76].

The same process for manually defining multiplets can be repeated for any and all correlations as necessary. Note the  button must be selected each time a new peak is to be defined.

- When evaluating the automatic analysis, and manually picking correlations, it is important to keep in mind that the automatic analysis routine purposely attempts to only include the stronger correlations, because the weaker ones are in many cases long-range and thus problematic. Therefore, peak picking EVERY peak in the data for use in generating structures is not usually advisable.
- However, in some cases where there is a small or unfavorable J-coupling constant value, these weaker peaks are not long range and can actually be necessary for the generation algorithms to piece the molecule together.

If at any point a mistake is made, it is possible to undo the last action with the  button in either the Combined Spectra Display or the Correlation Table. The software is able to record up to 20 modifications, making more than 20 modifications starts to remove the older ones from the memory.

It should again be pointed out that in the Combined Spectra Display window; the software distinguishes between peaks picked through the automatic analysis routine (noted by a solid circle, or x) and those that are handpicked by the user (a dashed circle). The Multiplet Properties Box will also note the difference, however in the Correlation Table, there is no difference between the cells in the Correlation Table.

Right clicking in the Combined Spectra Display in a region *without* picked peaks will open the following pop-up window:

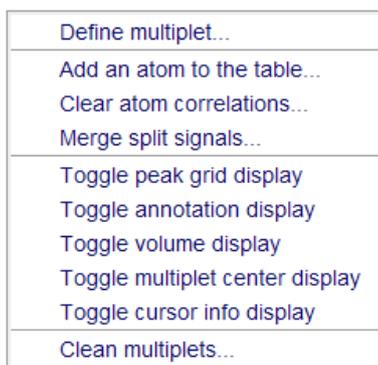


Figure 3.43: Additional Modification Actions Table.

From this menu, one can not only toggle the display of the Combined Spectra Display as described above, but also modify the contents of the Correlation Table.

- *Define Multiplet*, will have the same action as the  button and will allow the user to manually add additional multiplets (See the previous paragraphs in this section for details.)
- *Merge Split Signals*, Is particularly useful in cases of either ^{19}F or ^{31}P where there is manual correction necessary due to the X- ^{13}C coupling. There is an example of this in section [2.5.2.4].
- *Clear Atom Correlations*, Is used to either remove all correlations for an atom, or also removing the atom entirely from the table. The user should place the cursor over the corresponding chemical shift to remove (but not on the peak) right click select the corresponding action.
- The user is then prompted for which type of atom to remove

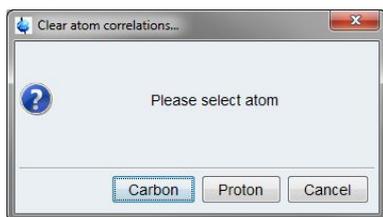


Figure 3.44: Manual remove options

And then the user is asked what exactly is desired to remove:

- *Clear all atom information* removes not only the correlations, but the atom entry from the Correlation Table, including chemical shift, hybridization etc.
- *Clear correlations only* leaves the atom in the table, and only removes the correlations associated with the atom.
- *Remove the Atom Completely* will not only delete it from the correlation table, but from the molecular formula. This can be used when the incorrect formula was given upon project creation.

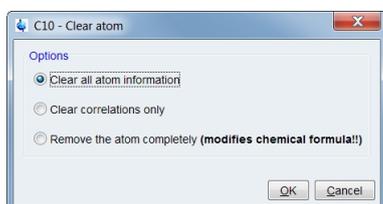


Figure 3.45: Dialogue box asking the user the extent of the information to be removed through this action.

- *Add an Atom to the Table*, can be used to add an atom to the table if there is an additional atom in the molecular formula that is not accounted for in the table.

To explain the use of these options to correct the Correlation Table an example is described below. In this example, the Automatic Analysis mistakenly picked the solvent peak instead of a carbonyl signal.

As highlighted in Figure below, the automatic analysis, in attempt to complete the necessary ^1H and ^{13}C count for the molecule, picked the solvent signal instead of a quaternary carbon. This can be the case of either a non-quantitative 1D- ^1H , or non-recommended HSQC sequence, or a sample with many contaminants.



Figure 3.46: Starting situation where the automatic analysis has picked the solvent peak (indicated by the cursor) but missed HMBC signals from a quaternary carbon (bottom of the figure).

Because this atom is completely incorrect, it is necessary to delete it and any corresponding correlations from the table. So the above option of *Clear Atom Correlations* from right clicking away from any correlations, and then selecting the first option of *Clear all atom information* will remove this ^{13}C from the Correlation Table, as shown in Figure below. Because this atom has an HSQC correlation, option to *Clear all Atom information* will also remove the corresponding ^1H from the table.

Structure Elucidation Basics

CMC-se: AndrosteneDione_C19H26O2 (C19H26O2)

File Edit View Analysis Structure Help

C19H26O2 DBE=7.0 C: 19/19, H: 25/26 (6*CH 8*CH₂ 1*CH₃), 86 HMBC, 17 COSY

H-C H-H C-C

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	
							6.47	3.33	3.26	3.24	3.24	3.11	2.98	2.83	2.73	2.70	2.56	2.50	2.43	2.42	2.3	
							C3	C7	C13	C9	C12	C13'	C12'	C10	C15	C16	C11	C14	C9'	C17	C16	
C19		14.63	1		sp3/sp																	
C18		18.19	3		sp3									M						M*	<	
C17		21.18	2		sp3													M			S-	
C16		22.62	2		sp3				>	M*	<			M		S-	M	M				S-
C15		31.73	2		sp3				M*	<	<	M			S-		M					
C14		32.33	2		sp3													S-		>	M*	
C13		33.09	2		sp3		M		S-			S-			M*	<	M					
C12		34.90	2		sp3		M				S-	S-	M							>	M*	
C11		35.68	1		sp3/sp				>	M*	<	M			M		S+		>	>	M*	
C10		36.39	2		sp3							M	S-									
C9		36.56	2		sp3					S-							M			S-		M
C8		39.55	0		sp3/sp		M					M	M	M			M		M*	<		
C7		41.03	1		sp3/sp			S+	M													
C6		48.17	0		sp3/sp											>	M*					
C5		51.35	1		sp3/sp				>	>	M*				>	M*	M	M				M
C4		54.46	1		sp3/sp										M*	<	M*	<	M*	<		
C3		124.56	1				S-		M*	<	<	M	M									
C2		171.93	0		sp2				M*	<	<	M		>	M*	<						
C1		220.74	0		sp2				>	>	M*			M	>	M*						
O20																						
O21																						

CMC-se: AndrosteneDione_C19H26O2 (C19H26O2)

File Edit View Analysis Structure Help

C19H26O2 DBE=7.0 C: 18/19, H: 24/26 (5*CH 8*CH₂ 1*CH₃), 85 HMBC, 17 COSY

H-C H-H C-C

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	
							6.47	-	3.26	3.24	3.24	3.11	2.98	2.83	2.73	2.70	2.56	2.50	2.43	2.42	2.3	
							C3	H2	C13	C9	C12	C13'	C12'	C10	C15	C16	C11	C14	C9'	C17	C16	
C19		14.63	1		sp3/sp																	
C18		18.19	3		sp3									M						M*	<	
C17		21.18	2		sp3														M		S-	
C16		22.62	2		sp3				>	M*	<			M		S-	M	M				S-
C15		31.73	2		sp3				M*	<	<	M			S-		M					
C14		32.33	2		sp3													S-		>	M*	
C13		33.09	2		sp3		M		S-			S-			M*	<	M					
C12		34.90	2		sp3		M				S-	S-	M							>	M*	
C11		35.68	1		sp3/sp				>	M*	<	M			M		S+		>	>	M*	
C10		36.39	2		sp3							M	S-									
C9		36.56	2		sp3					S-							M			S-		M
C8		39.55	0		sp3/sp		M					M	M	M			M		M*	<		
C7																						
C6		48.17	0		sp3/sp											>	M*					
C5		51.35	1		sp3/sp				>	>	M*				>	M*	M	M				M
C4		54.46	1		sp3/sp										M*	<	M*	<	M*	<		
C3		124.56	1				S-		M*	<	<	M	M									
C2		171.93	0		sp2				M*	<	<	M		>	M*	<						
C1		220.74	0		sp2				>	>	M*			M	>	M*						
O20																						
O21																						

Figure 3.47: Removing incorrect atoms from the Correlation table, with the “Clear Atom Correlations” through the Correlation Table.

Now the user can go back and add the correct quaternary carbon. This could be done by “enabling” each of the HMBC correlations as shown in the figure above, or with the *Add an Atom to the Table*, option which will not only add the atom, but enable all of the HMBC correlations at once as shown in Figure below.

This option will add an atom wherever the cursor is placed, so it is important that the cursor be at the correct chemical shift for the desired atom to be added, just not on a current peak.

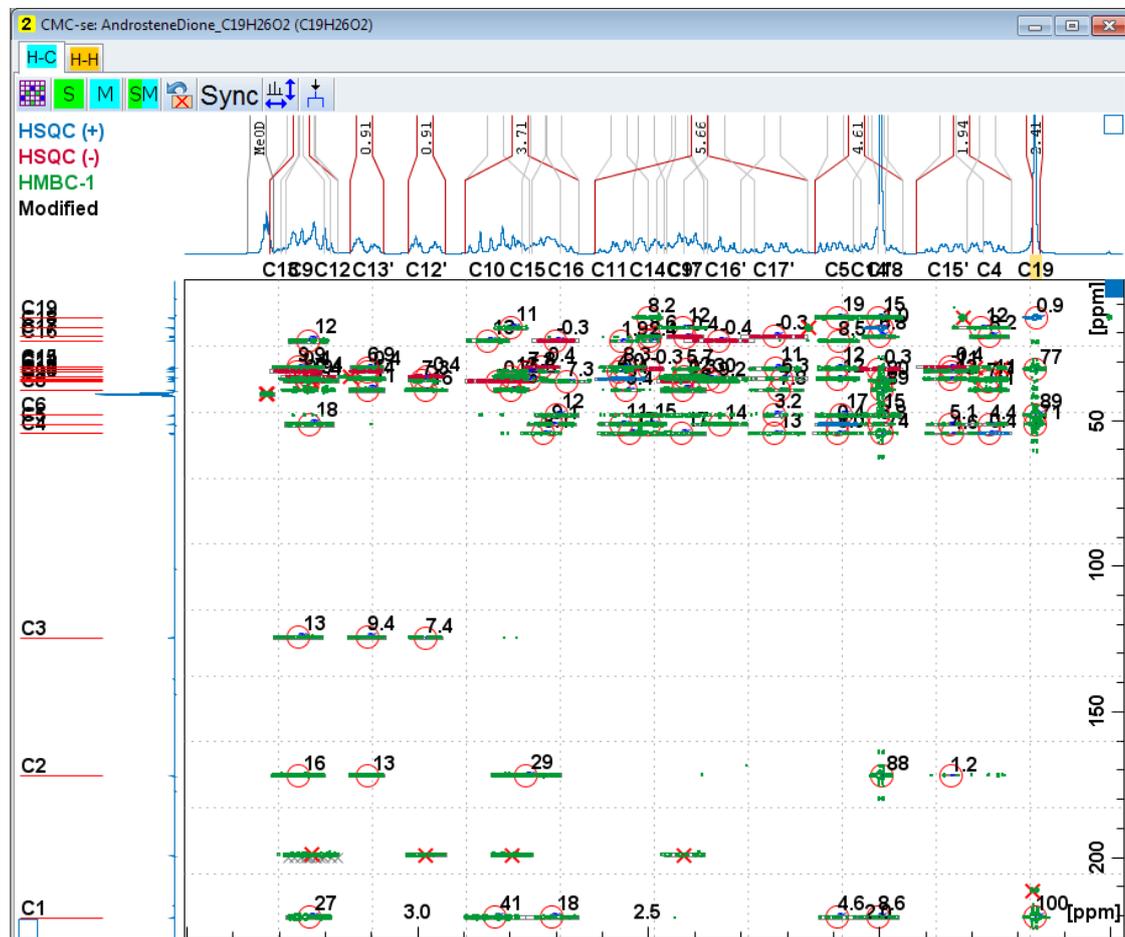




Figure 3.48: The manual addition of an atom and corresponding HMBC correlations to the table through the “Add an Atom to the Table”, option in the Combined Spectra Display.

3.6.2.3 Handling of ^{19}F or (^{31}P)- ^{13}C Splitting

As indicated previously, in molecules containing ^{19}F and ^{31}P likely will require manual corrections to the Correlation Table. This is due to the unknown extent and magnitude of these heteronuclear couplings.

The previous sections describing the addition and removal of atoms is quite general and may be used in this case if desired. However, CMC-se 2.1 introduced a new tool tailored to handle the visible ^{13}C - ^{19}F or ^{13}C - ^{31}P coupling.

As seen in the figure above, the popup menu in the HSQC/HMBC window of the Combined Spectra Display (which is obtained by right clicking away from a peak) contains the option to *Merge split signals*.

The Figure below illustrates an example where ^{19}F coupling is observed in both the ^1H and ^{13}C dimensions. The Automatic Analysis picked each of the individual ^{13}C multiplets as separate carbon atoms, thus not enabling the quaternary carbon to be included. The *Merge split signals* option to merge these 2 carbon atoms to a single atom with the chemical shift in the center of the multiplet.

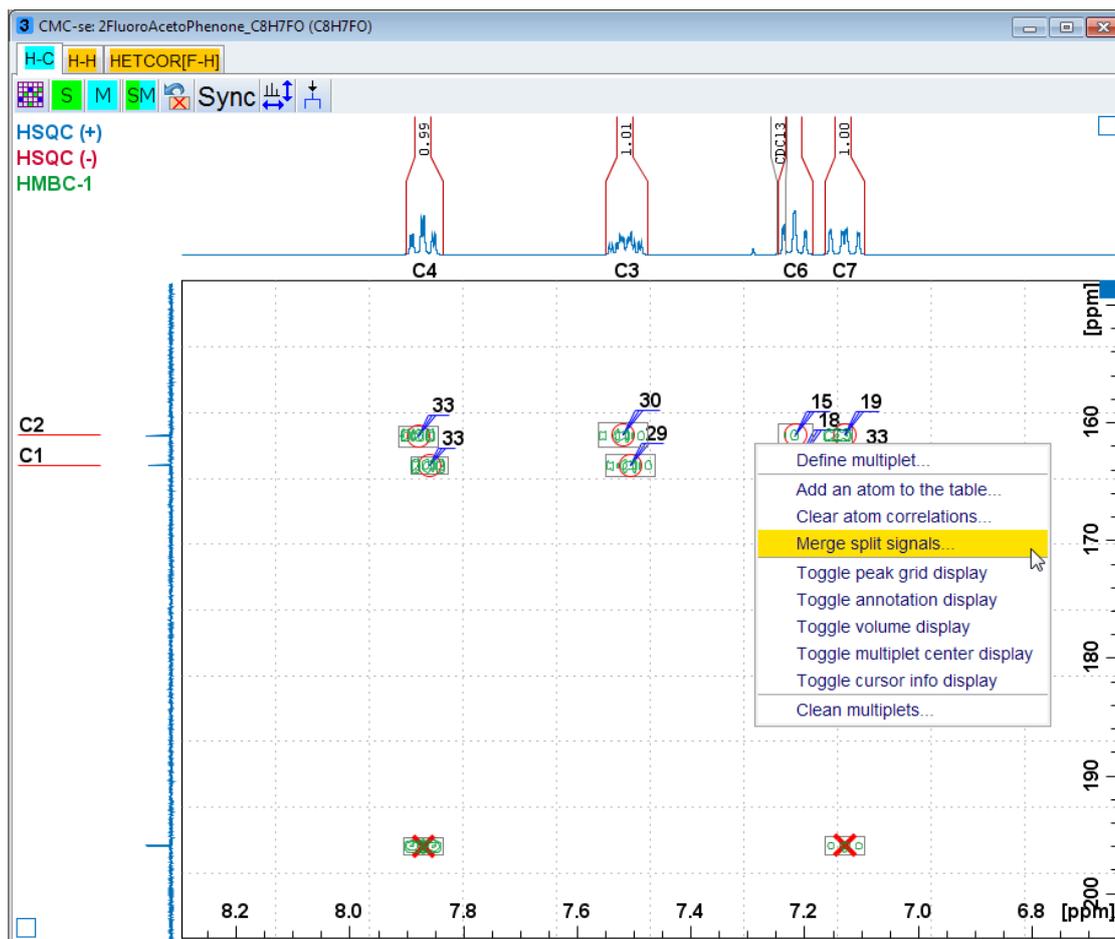
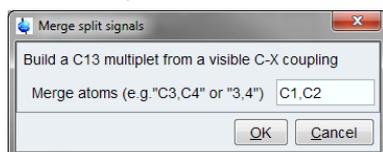


Figure 3.49: Situation of large ^{19}F - ^{13}C coupling which requires manual modification.

In the figure above the *Merge Split Signals* option can be used to merge the individual multiplets into a single atom in the Correlation table.

The software selects a pair of nearby carbon signals, and requests confirmation (or correction) from the user.



Confirming the dialog merges the specified carbon signals and all corresponding HMBC and HSQC correlations into one atom in the correlation table. The original signals are removed:

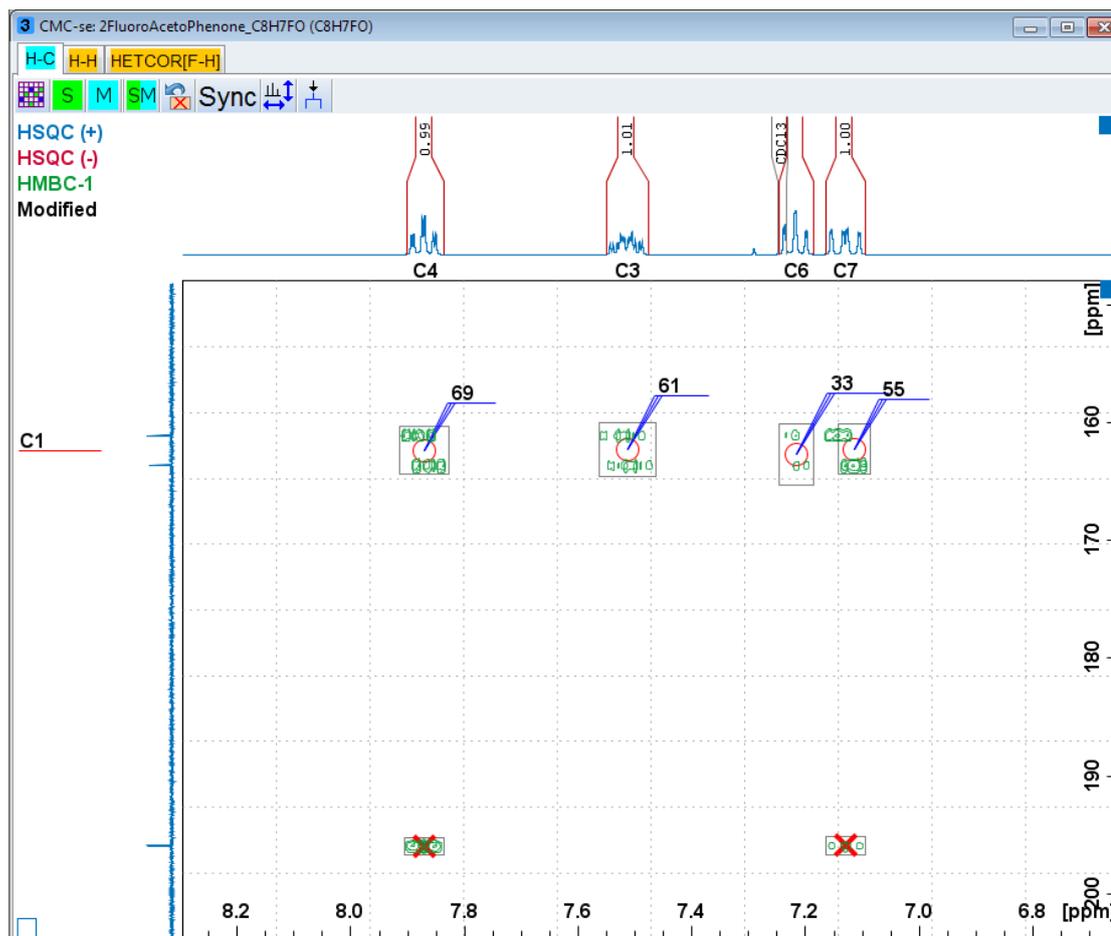


Figure 3.50: The result of merging the split multiplets of a ^{13}C coupled to ^{19}F .

It would now be necessary to add the desired quaternary carbon to the Correlation Table as was done above.

3.6.2.4 Adding INADEQUATE Correlations

The only type of correlation that cannot be picked by the method described above in the previous section, are the INADEQUATE correlations.

As mentioned in section [Additional Supported Experiments \[17\]](#), the CMCse_INAD parameter set is set up such that the SW in F1 is 2X that of SW in F2. However, it is often beneficial to set these to be equal and thus fold the spectrum to gain additional resolution. The INADQUATE used in this section, which is the one that is included in Example 3 (section [Structure Elucidation for Strychnine \[180\]](#)) has been folded.

It was also stated in section [Additional Supported Experiments \[17\]](#), but it is important enough to reiterate here, no matter how the data was collected (folded or not), it is crucial that the spectrum be referenced manually prior to attempting to pick the correlations. As shown in the figure below, the proper reference is that F1 (the indirect, double quantum dimension) be set to 2x that of F2 (the direct chemical shift dimension).

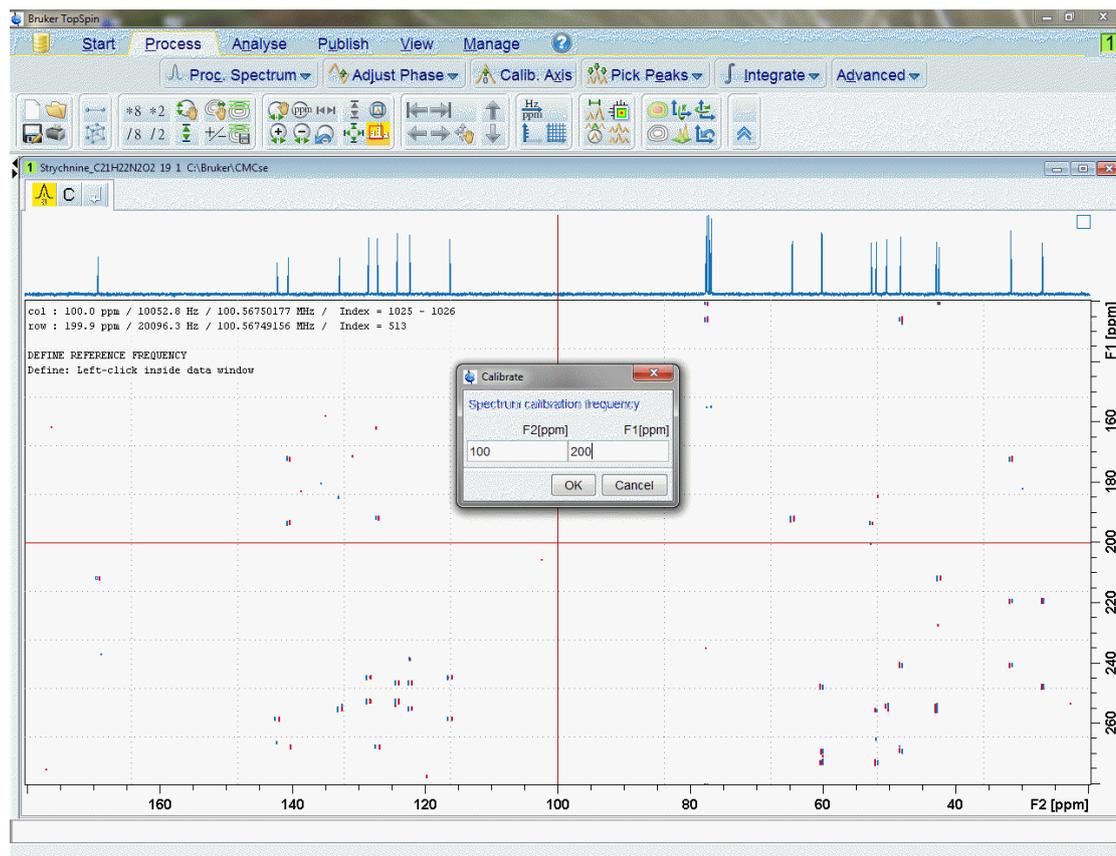


Figure 3.51: The INADEQUATE spectrum must be manually referenced appropriately prior to peak picking in CMC-se.

Similar to the other CMC-se Combined Spectra Display tabs, after the automatic analysis on the other spectra in the project has been completed, the ^{13}C projection of the INADEQUATE spectrum of the C-C tab includes the CMC-se Carbon numbers. Unlike the other windows however, the C-C tab has multiple cursors instead of a single cross-hair. The reason for this is this is a graphical tool to assist the user in interpreting the INADEQUATE spectrum.

Unlike traditional 2D NMR spectra, the indirect dimension of the INADEQUATE spectrum is not simply chemical shift but a sum of chemical shifts. For example in the figure below, the cross peak from C4 at 132 ppm and 254 ppm. The chemical shift in F2 of 132 ppm is that of C4, but the 254 ppm is the combined chemical shift of C4 + its neighboring carbon. And thus $254 \text{ ppm} - 132 \text{ ppm} = 122 \text{ ppm}$ which is the chemical shift of C8. The double cross-hair in the C-C tab of CMC-se does this math for the user.

In the example below, it is easy to see that this is a real correlation because there is a cross-peak at both locations, however when the signal is just above the noise level (as it usually is for INADEQUATE spectra) this is a powerful tool to help determine if a correlation is real or not. (It would be real, if the other cursor lines up with another peak in the 1D- ^{13}C spectrum).

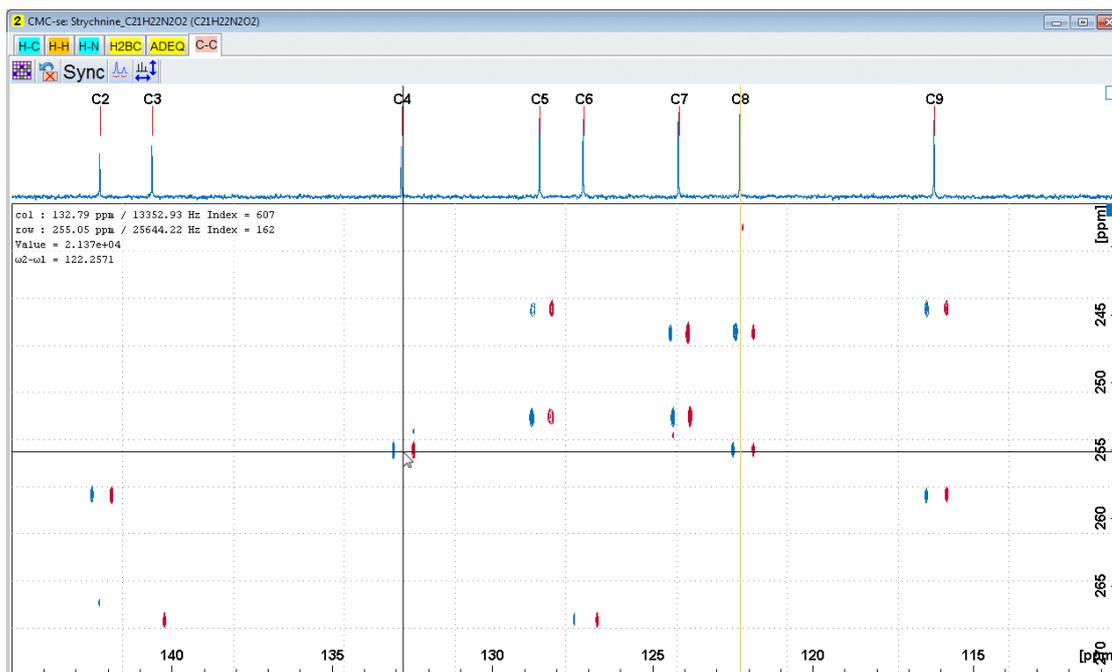


Figure 3.52: In the C-C tab of CMC-se, there is an additional cross-hair that indicates the corresponding carbon chemical shift of a correlation, allowing for quicker visual analysis of the INADEQUATE spectrum.

In the figure above, the additional cross-hair is yellow. However in the figure below, it is red to indicate to the user that this is a folded peak. This is only a graphical representation of the data and calculation, there is no consequence to the user, and the correlations are added to the table in the same manner, whether they are folded or not.

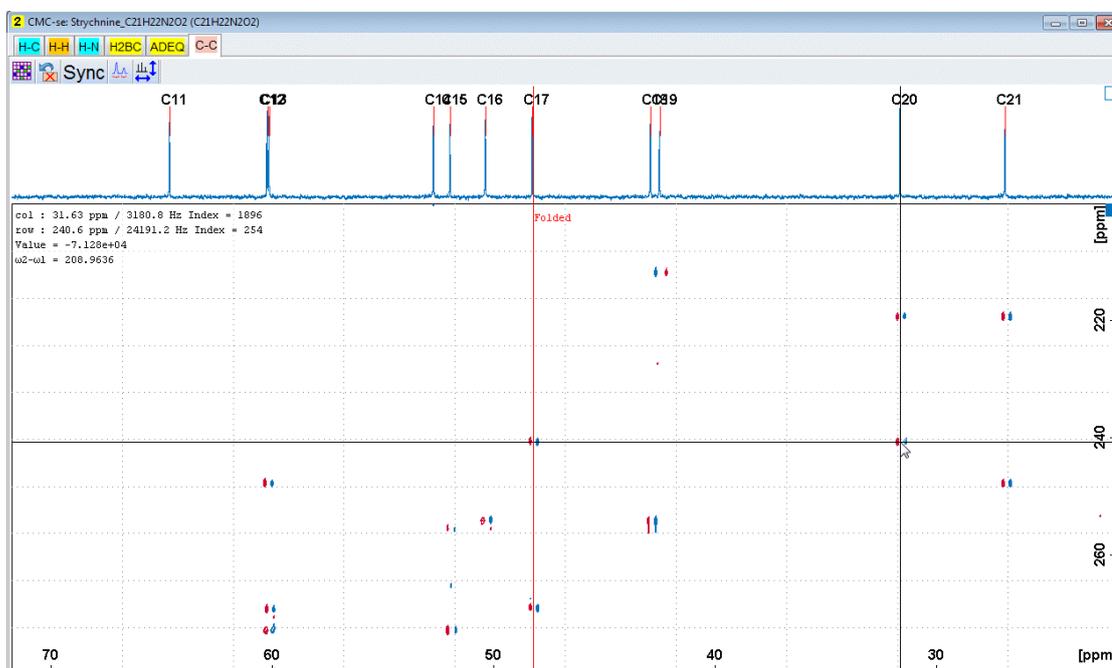


Figure 3.53: When the INADEQUATE spectrum is folded, the CMC-se module will recognize this and adjust the cross-hair accordingly. It still indicates the corresponding ^{13}C chemical shift, but the cursor is red to indicate that it is a folded correlation.

To manually define an INADEQUATE correlation, simply line up the multiple cursors on the corresponding correlations as shown in the figures above, and right click to define the multiplet as shown below. Once the multiplet is defined, it is automatically added to the table, and highlighted on the spectrum as shown for the bond between C2 and C9:



Figure 3.54: Defining the Multiplet

Once the cross-hairs are aligned with the 2 carbons for which an INADEQUATE correlation is desired, simply right click to define the multiplet. It will be added to the table, and also highlighted in the spectrum as shown.

If the Pop Up box similar to that shown in figure below is obtained, then the user should pay close attention to the exact position of the cross hairs. The vertical cross hairs should line up exactly with the 1D carbon projection.

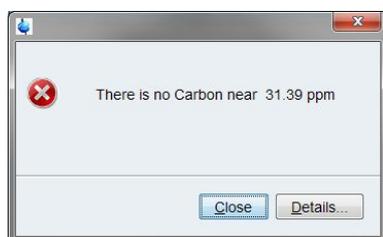


Figure 3.55: Pop up box prompting the user to double check the position of the cursor and vertical cross hairs because the software can't find a carbon to correspond to the current position.

3.7 Additional Modifications to the Correlation Table

There are some cases when picking the correlations through the Combined Spectra Display isn't sufficient or desired. Examples of this could include:

- Deleting all correlations for a given atom in a single click rather than deleting each correlation individually through the Combined Spectra Display (as described in the section [Adding and Removing Multiplets Through the Combined Spectra Display \[▶ 57\]](#)), or deleting the atom from the project entirely.
- Cases of extreme overlap and the software can't assign the correlation to a single proton/carbon or proton/proton cell in the Correlation Table, but a visual inspection by a trained spectroscopist can distinguish the difference.
- The situation where the user wants to enable or disable (rather than permanently changing the peak/multiplet lists in the original data files as described above) correlations for any reason to see what affect it has on the structure generation process.
- Correcting or adding additional information such as Hybridization State, # of attached protons etc.

The following sections discuss how such modifications directly to the Correlation Table are executed.

3.7.1 Deleting Atoms from the Correlation Table

Just as it was possible above to delete a variety of information about an atom through the Combined Spectra Display, it is possible to perform a similar action through the Correlation Table.

Depending on where in the Correlation Table the user right clicks, they will get a slightly different pop up menu as shown in the Figure below. Either of these options *Clear Atom* or *Clear current row* will bring up the pop up window allows the user to decide the extent of the information to be removed from the table.

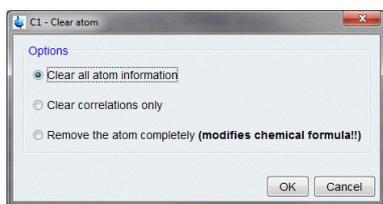


Figure 3.56: Dialogue box asking the user the extent of the information to be removed through this action.

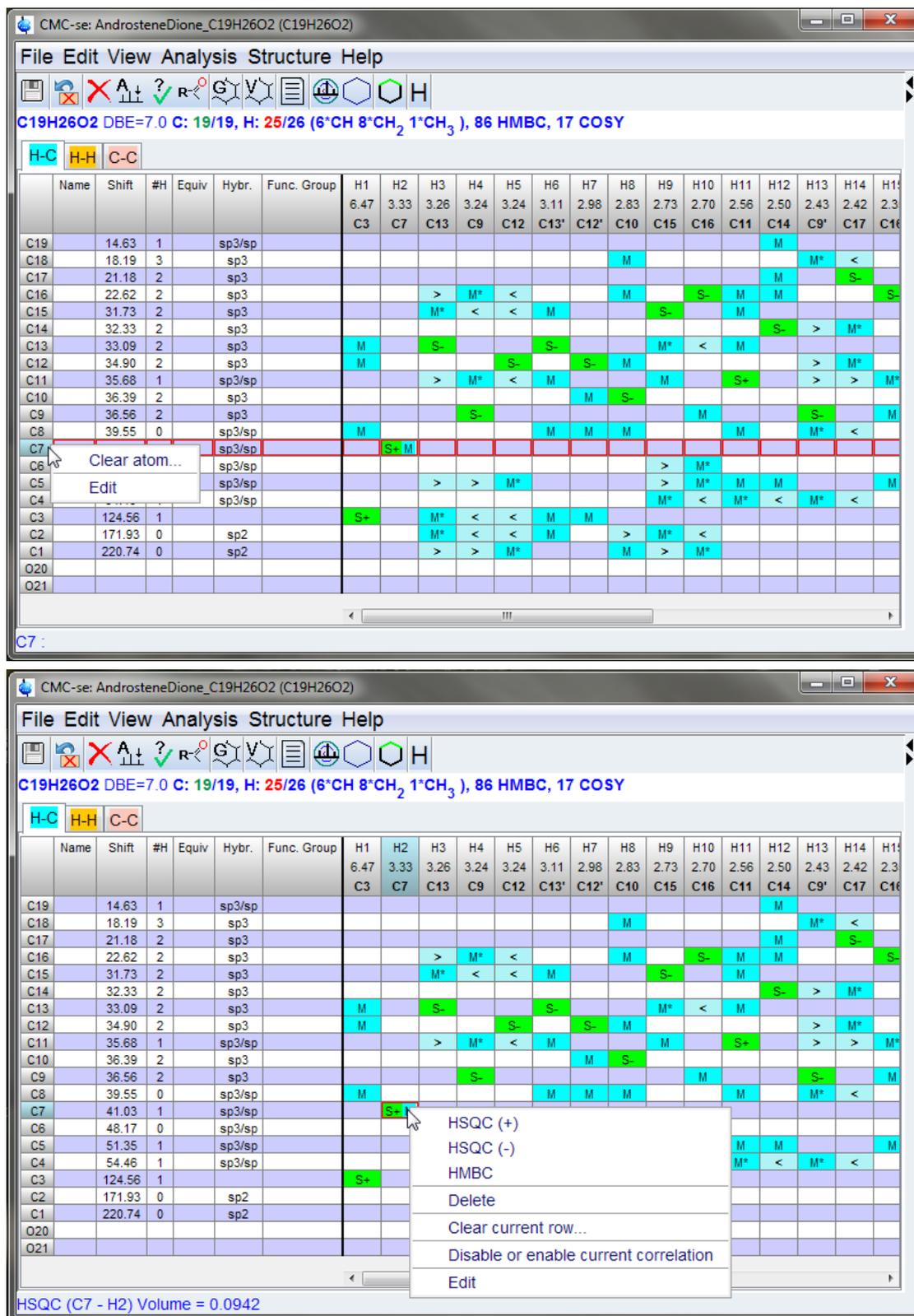


Figure 3.57: Pop Up menus resulting from right clicking in the Correlation Table. Both menus have options for clearing content of the Correlation Table.

3.7.2 Adding/Deleting Correlations through the Table

The link between the cursors in the Combined Spectra Display and the highlighted cells in the Correlation Table (see first figure in section [The Combined Spectra Display Window \[46\]](#)) is a powerful tool to help modifying the Correlation Table. One can move the cursor in the data, and the corresponding cell is highlighted in the table, thus indicating which cell corresponds to a given data signal, and thus which cell to modify.

To manipulate a given correlation (or empty cell) in the table, right-click on the desired cell, and the pop-up menu appears for a HMBC or HSQC correlation:

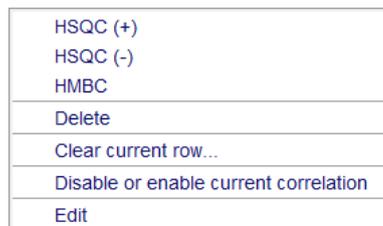


Figure 3.58: Example pop-up menu to manipulate HSQC/HMBC correlations.

The corresponding menu for COSY correlations is illustrated below:

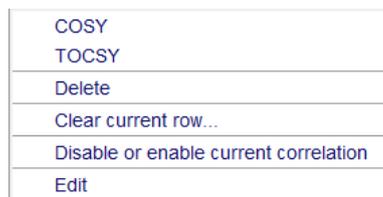


Figure 3.59: Pop-up menu to manipulate COSY correlations.

Clicking on the HSQC/HMBC/COSY option will add a correlation to that cell in the table, regardless of whether or not that peak was previously picked in the data.

This option is particularly useful if for some reason the software isn't able to assign a multiplet in the data to a given pair of atoms.

Additionally, for all correlations it is possible to delete the chosen one. This action is the same as deleting a peak through the Combined Spectra Display, and result in the picked peak icon from being removed from the Correlation Table and the Combined Spectra Display.

3.7.3 Disabling Correlations in the Correlation Table

Removing suspicious correlations is often necessary for a successful structure generation. But given that the suspicious correlation may in fact be real and helpful in evaluating the generated structures, it is often desirable to not delete the correlation entirely.

Thus, an alternative method is to disable/enable the current chosen correlation or all correlations of the same type. Disabling a correlation will remove that correlation from the structure generation process, but still leave it in the table for further evaluation of the structures.

This can be particularly useful for eliminating the weaker correlations in the data from the structure generation process. It is possible, that these weak correlations are long range, and thus will be problematic for the structure generation (section [Chemistry Rules \[95\]](#)), but these correlations are still real, and can help in the final validation process (section [Viewing Correlations on Structures \[110\]](#)).

Correlations can be disabled (or enabled) on an individual basis with the **Disable or enable current correlation** option shown in the pop-up windows shown in the previous section.

Note that while it is possible to disable/enable an HSQC correlation, this will result in the removal of the H-C bond, which can cause other problems and inconsistencies, and thus not recommended.

For a disabled COSY correlation the cell is yellow with a grey C. An enabled one is orange with a black C. For all other correlations (HMBC, ADEQUATE etc) only the color of the letters changed. Disabled correlations have a grey letter, enabled are black:

Assigned C: 10/10. Assigned H: 20/20 (3*CH 4*CH₂ 3*CH₃), 15 HMBC. 6 COSY

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	
							3.30	2.14	1.89	1.58	1.53	1.33	1.17	1.03	0.95	0.90	0.90	0.84	0.83	0.77	-	
							C1	C6	C3	C4	C7	C5	C10	C2	C10'	C7'	C3'	C9	C8	C4'	H15	
C10		20.77	2		sp3								S-		S-							
C9		20.91	3		sp3													S+				
C8		22.08	3		sp3												M		S+		M	
C7		22.90	2		sp3			M			S-					S-					M	
C6		25.44	3		sp3		M	S+										M				
C5		31.52	1		sp3						M	S+								M		
C4		34.42	2		sp3				S-		M									M		S-
C3		44.86	2		sp3				S-									S-			M	
C2		49.83	1		sp3			M		M	M			S+					M			
C1		71.09	1		sp3		S+				M									M		
O11																						

	Name	Shift	Mult	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15			
				3.30	2.14	1.89	1.58	1.53	1.33	1.17	1.03	0.95	0.90	0.90	0.84	0.83	0.77	-			
				C1	C6	C3	C4	C7	C5	C10	C2	C10'	C7'	C3'	C9	C8	C4'	H15			
H15	H15																				
H14	C4'	0.77					C							C							
H13	C8	0.83							C												
H12	C9	0.84			C																
H11	C3'	0.90				C					C										C
H10	C7'	0.90																			
H9	C10'	0.95																			
H8	C2	1.03													C						
H7	C10	1.17																			
H6	C5	1.33					C														C
H5	C7	1.53																			
H4	C4	1.58									C										C
H3	C3	1.89																			
H2	C6	2.14													C						
H1	C1	3.30																			

Figure 3.60: Correlation Table displaying disabled COSY and HMBC correlations

3.7.4 Additional Manipulations to Correlations in Table

The final option in the pop-up menu shown in the previous section is the **Edit** option. Clicking **Edit** opens the *Correlation* window for the current correlation:

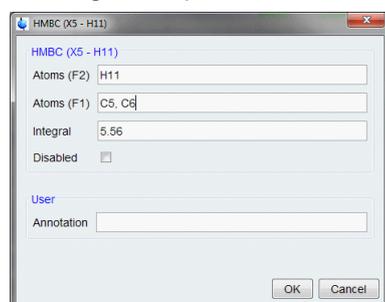


Figure 3.61: Pop-up dialog box to edit correlation properties.

Here all information about the correlation can be seen and just like from the **Properties** window (see *Multiplet Properties* Window in section [Modifying Existing Multiplets \[52\]](#)) from the Combined Spectra Display, there is the option of writing an annotation for that correlation (see section [Bookkeeping Tools within the CMC-se Module \[78\]](#)).

The correlation being edited in the figure above is a fuzzy HMBC correlation (**M*** and **v** in the Correlation Table). A fuzzy HMBC correlation is one that is overlapped such that the software could not determine with confidence which resonance (^{13}C or ^1H) it should be assigned too. The structure generation algorithms will use both possibilities equally. As shown in this figure, this correlation can be assigned to H11 and either C5 or C6.

To be on the safe side, the CMC-se automatic analysis is rather generous when it assigns a fuzzy correlation, and in many cases it is possible for the spectroscopist to resolve the ambiguity. In these cases, it is advised that the fuzzy correlation be removed to reduce the possibilities and thus the time for the structure generation algorithms.

This can either be accomplished by deleting the **v** cell if the **M*** is in the correct place, or removing the incorrect atoms from the Properties box.

3.7.5 Adding Additional Information to the Correlation Table

After the [Automatic Spectra Analysis \[39\]](#) there still might be nuclei with only partial information filled in the Correlation Table (or in some cases incorrect information). As every bit of additional information will help the structure generator to work faster and produce fewer and more reliable structure proposals, this information should be filled in and verified by the user if possible.

The columns *#H*, *Equiv*, *Hybr* can all be edited in a simple manner by right clicking in the cell, and modifying by selecting the desired entry from the resulting pop-up window. More information about these properties is included below.

To modify multiple property entries for a given hetero-atom, the window shown below is advised. This pop-up window is obtained by right clicking and selecting **Edit** from either the *atom number*, *name*, *shift*, *#H*, *Equiv*, *Hybr* or *Func. Group* cells, or as explained in the section [Modifying Existing Multiplets \[52\]](#), by double clicking on an HSQC correlation in the Combined Spectra Display.

Figure 3.62: Atom Properties window for a selected heteroatoms

- If necessary, the *Number of attached H atoms* can be modified from this window, just as by right clicking in the appropriate cell as described above.
- If the automatic analysis did not detect the symmetrical atoms correctly, the *Equivalence* field can be updated. Empty cells indicate no symmetry. Cells containing a number n , indicate that the corresponding C or other hetero atom has $n-1$ symmetrical equivalents. For example, an Isopropyl group would have an Equivalence entry of 3.
- Both generation algorithms allow the *Hybr* entries to be undefined, but will likely take longer, and/or generate more structures if not completed so they should be completed if at all possible.
- Sometimes it is easy to assign an atom to a predefined functional group – while not done automatically by the software, only those functional groups reasonable for an atom with that chemical shift are shown in the pull-down menu. **The BRUKER algorithm for generation uses this more specific information as additional constraints for generation.** Additional functional groups and larger fragments may be assigned to that atom using the Fragment tool (see section [The Fragment Tool \[▶ 81\]](#)). As shown in the figure above, if a functional group is defined for that atom, a fragment displaying this functional group is also drawn in this *Properties* window.
 - In molecules that contain hetero atoms besides carbon, the definition of “hetero-aromatic” is likely preferred over the plain “aromatic”. This hetero-aromatic designation indicates that it is an aromatic ring that **may or may not** contain a hetero atom. The aromatic definition restricts the ring to only containing carbon atoms.
- The **Multiplicity** of the proton attached to the atom for which the **Properties** window was opened can also be defined from this window. When possible, defining the multiplicity **greatly** increases the speed of the generation, and reduces the number of structure candidates, as this provides additional restraints for building the structure.
 - A Methyl or aromatic proton defined as a singlet cannot be adjacent to another CH_n group.

- A Methyl or aromatic proton defined as a doublet, will only have one protonated neighbor.
- An aromatic proton defined as a triplet (in this case it is a pseudo triplet – degenerated doublet of doublets), will have two protonated neighbors.
- UNDEFINED is default, and has provides no constraints on the structure generation.

The proton multiplicity, and other information about a given ^1H in the molecule can also be defined by right clicking on the *Mult* cell, or editing the ^1H atom properties by double clicking on the *Mult* column. The proton properties pop-up window is shown below. Note that shown here is the automatic re-naming of the proton based on the connected carbon:

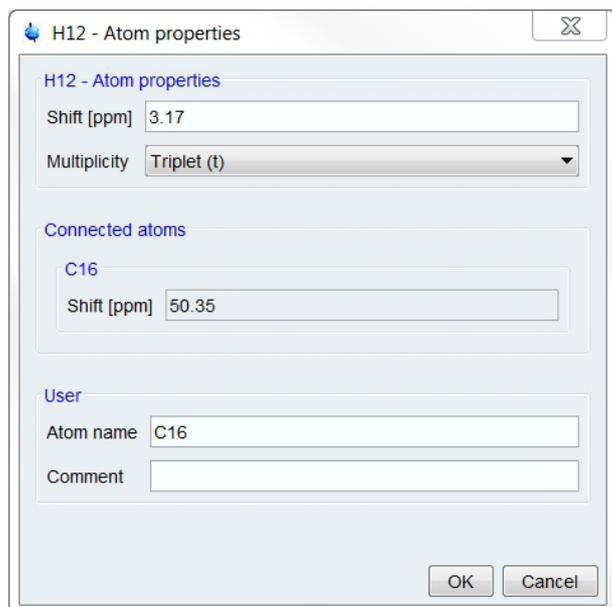


Figure 3.63: Atom Properties window for a ^1H atom

For either hetero-atoms, or protons it is possible to add an **Atom name** or user **Comment** which will be displayed in the Correlation Table, and on the Combined Spectra Display. More about these entries will be described below in the section [Bookkeeping Tools within the CMC-se Module \[78\]](#).

Also note that a good portion of the information contained in the *Atom Properties* window is also shown at the bottom of the *Project* window when the cursor hovers over a correlation, or a given atom column (example shown later in the following section).

3.8 Bookkeeping Tools within the CMC-se Module

As seen in previous sections, there are multiple places within the CMC-se module that the user can add additional names, annotations or comments about various atoms or correlations in the spectra. These tools make CMC-se a valuable tool for working with the data manually even if one of the Structure Generation Algorithms to generate possible structures is not used.

Double Clicking on an HSQC correlation or right clicking on an atom in the Correlation Table (section [Adding Additional Information to the Correlation Table \[76\]](#)) will bring up the *Atom Properties Menu* shown:

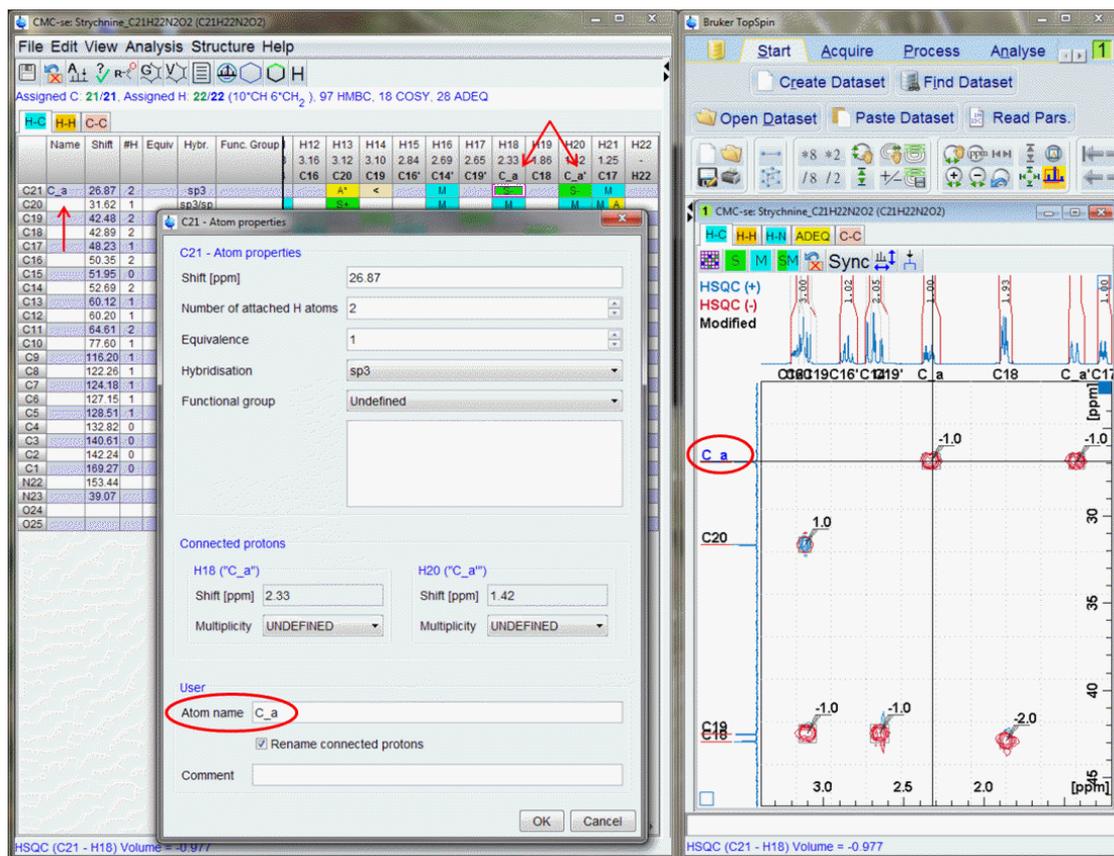


Figure 3.64: Atom Properties pop-up, and resulting changes in Correlation Table, and Combined Spectra Display when an atom is given a name by a user.

As shown above, one of the HSQC peaks corresponding to C21 was double clicked on to bring up the *Atom Properties* window, and the **Atom Name** (red circle in figure) was filled in to correspond to a user defined C_a instead of the CMC-se provide number which is based on chemical shift. Once the **OK** button was pressed, the resulting changes in the Correlation Table and Combined Spectra Display are seen.

- The **Name** column is updated in the Correlation Table (highlighted by red arrows).
 - Because the **Rename connected protons** option was checked, the protons corresponding to this carbon are also automatically renamed accordingly. Because this carbon corresponds to a CH₂ group with non-equivalent protons, they are named C_a, and C_a'.
- The carbon label on the projection of the Combined Spectra Display is also updated to show the user name rather than the CMC-se provided name (highlighted with a blue arrow).

Note that this user provided name is blue instead of the black for a CMC-se defined name.

When a user is unsure of the atom name, the **Comment** field can be useful:

CMC-soe: Strychnine, C₂₁H₂₂N₂O₂ (C₂₁H₂₂N₂O₂)

File Edit View Analysis Structure Help

Assigned C: 21/21, Assigned H: 22/22 (10°CH 6°CH₂), 97 HMBC, 18 COSY, 28 ADEQ

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1 8.07 C9	H2 7.23 C5	H3 7.14 C8	H4 7.07 C7	H5 5.86 C6	H6 4.26 C10	H7 4.11 C11	H8 4.05 C11'	H9 3.91 C13	H10 3.84 C12	H11 3.68 C14	H12 3.16 C16	H13 3.12 C20	H14 3.10 C19	H15 2.84 C16'	H16 2.69 C14'	H17 2.65 C19'	H18 2.33 C_a	H19 1.86 C18	H20 1.42 C_a'	H21 1.25 C17	H22 -	
C21	C_a	26.87	2		sp3									M A					A*	<		M		S-		S-	M		
C20		31.62	1		sp3/sp						M M		M M						S+			M		M		M A			
C19		42.48	2		sp3						M* A*				v				S-		v		S-		M A				
C18		42.89	2		sp3							A A			M*				M*						S-				
C17		48.23	1		sp3/sp						M A					M A			M* A*	<<				M		M S+			
C16		50.35	2		sp3										M S-						S-	M			M A				
C15		51.95	0		sp3/sp					M*	M*	v	M		v	A M A					v			M*	M A				
C14		52.69	2		sp3					Δ	Δ	M*			M*	M			S-			M*	S-		Δ			Δ	
C13		60.12	1		sp3/sp							M*			S+							M*	Δ		v	M*	M*	v	M* A*
C12		60.20	1		sp3/sp										S+							Δ	Δ		M*	Δ	Δ	Δ	
C11		64.61	2		sp3						M A	M	S-	S-															
C10		77.60	1								S+	M*	<															M A	
C9		116.20	1		sp2			S+	M A	M	M																		
C8		122.26	1		sp2			M	M	S+	M																		
C7		124.18	1		sp2			M	M A	A	S+																		
C6		127.15	1		sp2						S+				M* A*	<<													
C5		128.51	1		sp2			A	S+	M	M A																		
C4		132.82	0		sp2			M	M	M A	M																		
C3		140.61	0		sp2										M*	<													
C2		142.24	0		sp2			M A	>	M*	<																		
C1		169.27	0		sp2																								
N22		153.44						M																					
N23		39.07																											
O24																													
O25																													

C2: Think this is C6 based on HMBC to H1 and H3

Figure 3.65: Correlation Table highlighting atom properties when a user Comment is added

In the figure above, a user comment indicating a possible assignment was added to the **Comment** field of the *Atom Properties* pop-up window. Because this is just a comment, which can be somewhat lengthy at times, this text is not shown on the Combined Spectra Display, or in the table, but rather at the bottom of the table. When the cursor hovers over a given atom the text at the bottom of the Project Window (below the Correlation Table) changes to reflect the properties of that atom.

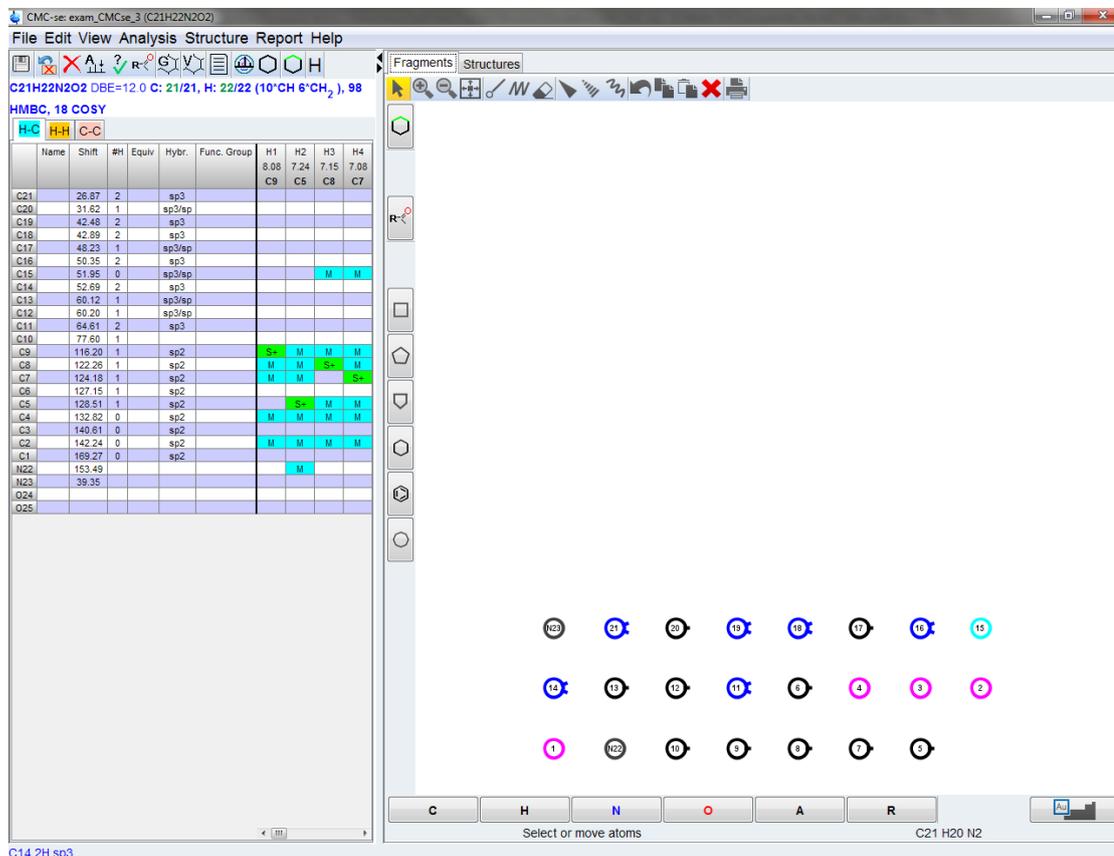
In the figure above, a comment was included in the Comment field of the Atom Properties window. The **OK** button was pressed, and the cursor is now hovering over the row in the table corresponding to C2.

The comment is displayed at the bottom of the table.

Double clicking on a cell in the Correlation Table (or an HMBC in the Combined Spectra Display) brings up a different pop-up window where a user can provide an **Annotation** rather than an **Atom Name**. The results of such an action are shown below:

All additional knowledge about the fragments in the unknown structure can be used in the fragment tool to make the structure generation more explicit.

Because of its great utility in the structure elucidation process, it is automatically opened as one of the tabs in the main CMCse project window, next to the Correlation table. The relative size of this window can be modified with the  buttons.



Name	Shift	#H	Equiv	Hybr.	Func. Group	H1 8.08 C9	H2 7.24 C5	H3 7.15 C8	H4 7.08 C7
C21	26.87	2		sp3					
C20	31.82	1		sp3/sp					
C19	42.48	2		sp3					
C18	42.89	2		sp3					
C17	48.23	1		sp3/sp					
C16	50.35	2		sp3					
C15	51.95	0		sp3/sp				M	M
C14	52.69	2		sp3					
C13	60.12	1		sp3/sp					
C12	60.20	1		sp3/sp					
C11	64.61	2		sp3					
C10	77.59	1		sp2					
C9	116.20	1		sp2		S-	M	M	M
C8	122.26	1		sp2		M	M	S-	M
C7	124.18	1		sp2		M	M		S-
C6	127.15	1		sp2			S-	M	M
C5	128.51	1		sp2			M	M	M
C4	132.92	0		sp2			M	M	M
C3	140.61	0		sp2					
C2	142.24	0		sp2			M	M	M
C1	169.27	0		sp2					
N22	153.49							M	
N23	39.35								
O24									
O25									

Figure 3.67: Fragment Tool within the main CMCse project window

The fragment editor is similar to other chemical structure drawing programs, but here is a brief description of the key buttons:

-  - Used to select and move objects.
-  - Used to control the zoom level.
-  - Draw and erase tools.
-  - Stereo bond definition (Note that at this time CMC-se ignores the stereo chemistry).
-  - Undo.
-  - Simple ring templates.
-  - Add an atom.

- Add a pseudo atom.
- Periodic Table for additional elements. Please refer to [Program Prerequisites \[11 \]](#) for those atoms that are supported in CMC-se.
- Removes the displayed correlations from the window
- Opens fragment selection window to select templates or exclude fragments from the structure generation.
 - This option can also be used to open a pre-existing .mol file into the fragment editor.

Drawing Fragments

In the selection mode it is possible to select bonds and atoms. With holding a bond or atom within the selection it is possible to move it. Right-clicking the mouse button on a bond enables to flip on this bond or flip the bond horizontal or vertical. If a selection is made, it is also possible to rotate the selection around the center starting from the red arrow shown below:

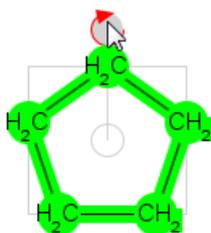


Figure 3.68: Fragment editor in selection mode



Change quickly back to the selection mode with the mouse right-click.

Activate the drawing mode using the button to draw single bonds or the button to draw chains. Drawing normal bonds, either by confirming the suggested bonds (considering the correct bond angle), or with left-clicking and dragging the mouse to an individual position, are illustrated in the following figures. The individual drawing enables the connection of given fragments by dropping the mouse above the desired atom.

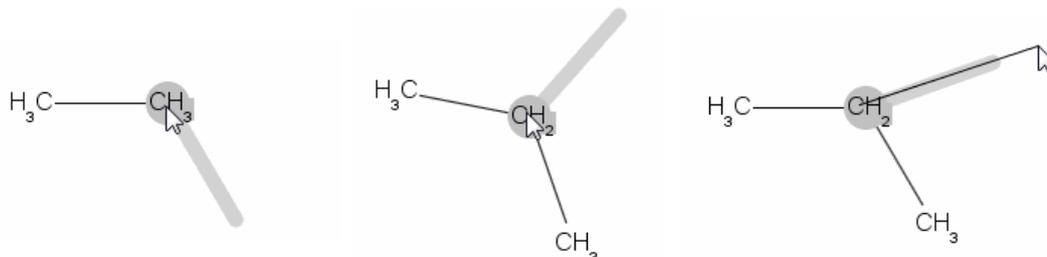


Figure 3.69: Drawing Normal Bonds

To draw chains, left-click the mouse button and drag till the desired length is reached. This is visualized in the following figure:

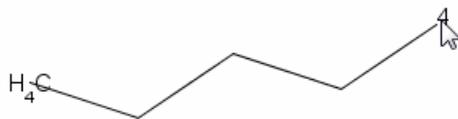


Figure 3.70: Drawing chain bonds

The default drawing atoms are carbon chains. To change single atoms select another atom from the toolbar  or the periodic table . It is also possible to combine this selection with an active drawing mode.

To change the bond type the following buttons are provided . In this mode it is possible to toggle between the different possibilities (e.g. change the direction of the bond or change back to normal single bond) by repeatedly clicking on the bond. While it is possible to draw a bond with a certain stereo chemistry, CMC-se does not interpret or predict 3D structure at this time.

Through the use of templates it is possible to quickly build larger fragments. To do this, drag the selected fragment to the desired binding site, until green circles appear, indicating the successful recognized binding:

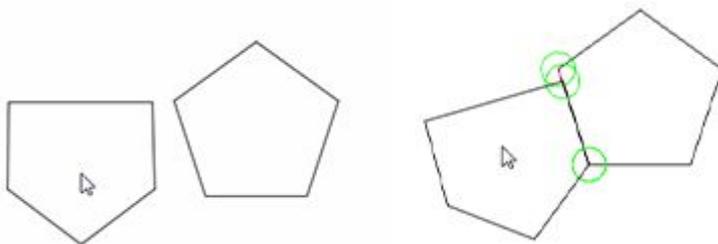


Figure 3.71: Docking Recognition for Binding Sites

For a full description about working with the structure editor see the Manual (edstruc).

Working with Fragments - Using atom information for the structure generation

This window also contains colored circles that correspond to the various carbon and other possible assignable hetero atoms. They are color coded according to hybridization and attached proton count. There are also small bullets on the sides of the circles to visually identify the attached proton count. If the user knows even a single atom assignment for this fragment, they can click and drag that ball to the corresponding position of the fragment. To remove the assignment, left-click on the relevant atom and select **Remove assignment** in the context menu:

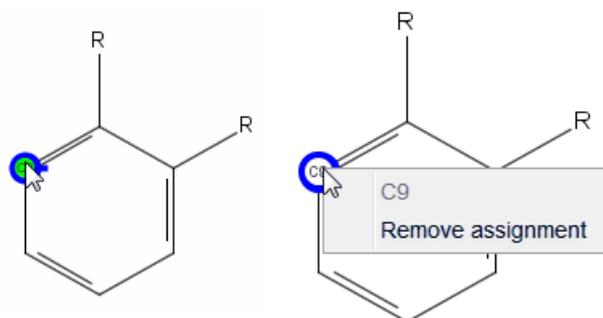


Figure 3.72: Illustration for the Work with Assignments

Once an atom has been assigned to the fragment, the **Func Group** column of the Correlation Table is updated with information regarding the assigned atom. An even more useful feature, is that the ball in the fragment can be clicked on, and the various experimental correlations shown between that ball, and the other corresponding balls; thus facilitating the assignment process in a graphical manner. To remove the correlations from the window display, press the button .

For example, in the figure below, it is clear that C-9 has a COSY (tan line) to C-5, thus it must be the next carbon in the ring:

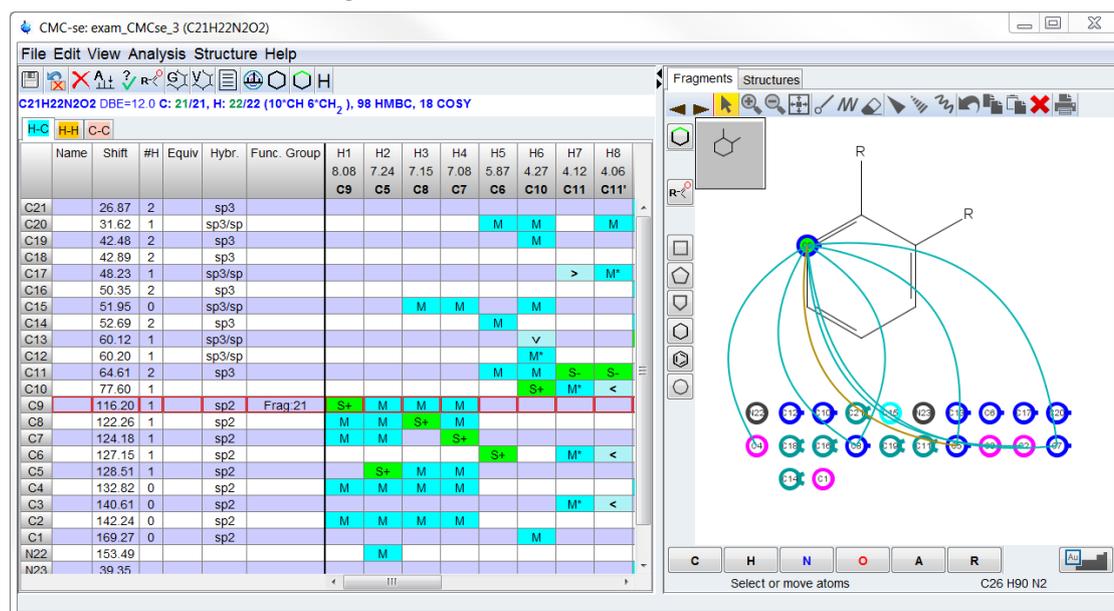


Figure 3.73: Illustration of how the experimental correlations can be used to graphically assign fragments to be used as building blocks for the structure generation.

It is also possible to add fragments in the editor by defining the functional group, as shown in the following figure:

Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2
C21	26.87	2		sp3			
C20	31.62	1		sp3/sp		8.08	7.24
C19	42.48	2		sp3			
C18	42.89	2		sp3			
C17	48.23	1		sp3/sp			
C16	50.35	2		sp3			
C15	51.95	0		sp3/sp			
C14	52.69	2		sp3			
C13	60.12	1		sp3/sp			
C12	60.20	1		sp3/sp			
C11	64.61	2		sp3			
C10	77.60	1		sp3	R2CH-O-		
C9	116.20	1		sp2		S+	M
C8	122.26	1		sp2		M	M
C7	124.18	1		sp2		M	M
C6	127.15	1		sp2			
C5	128.51	1		sp2	Frag 9	S+	
C4	132.82	0		sp2		M	M
C3	140.61	0		sp2			
C2	142.24	0		sp2		M	M
C1	169.27	0		sp2			
N22	153.49						M
N23	39.35						
O24							
O25							

Figure 3.74: Illustration for the adding of fragments through the assignment of functional groups.

While the selected fragments may have overlapping atoms, each atom of the Correlation Table may have only one assignment. Any given carbon atom can have either one assignment on a single fragment, or in a functional group, but not both, and also cannot be assigned in multiple fragments. If the assigned atom truly appears in multiple overlapping fragments, it is necessary to provide a user defined larger fragment.



With the improved Fragment Tool in CMCse 2.3, it is possible to extend and join the existing assigned fragments with the drawing tools mentioned above



There is no limit to the size or count of fragments that could be submitted as user defined fragments. Thus it is feasible that this tool could be used to help manually assign the resonances if a proposed structure is already known.

The non-existing fragments can also be important knowledge for the automatic structure elucidation process. To exclude fragments from the structure generation, use the fragment selection window which is opened with the  button, and then right click on the box with the appropriate fragment.

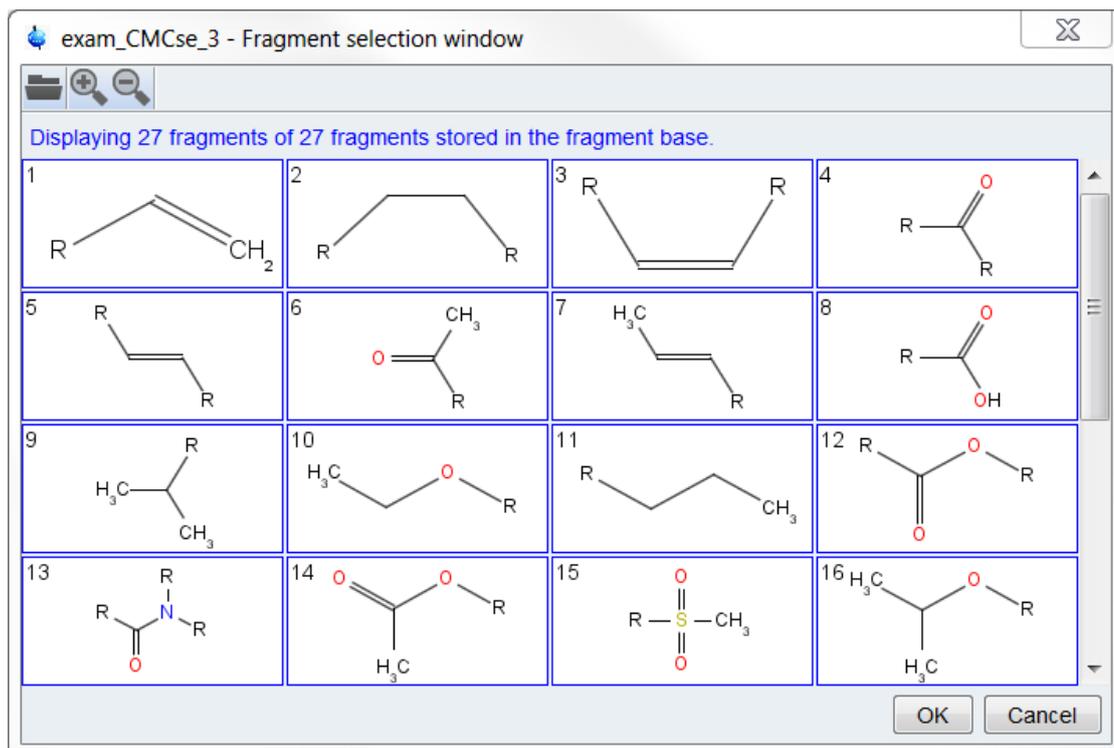


Figure 3.75: Fragment Filter dialog window

The **Zoom** buttons () control how many structures are shown in the display. The **Open** button () enable the possibility of loading additional fragments. Any standard MDL V2000 .mol file is supported.

Double clicking on a particular fragment will toggle between the current zoom of many fragments, and just that particular fragment.

Right-clicking in the box of a fragment a pop-up menu appears which offers the following selections concerning this fragment:

- **Require** → Select this fragment for the fragment editor. If a structure without this fragment is generated, this structure will be removed.
- **Exclude** → If a structure containing this fragment is generated, the structure will be removed.
- **Ignore** → No action is taken for this substructure during structure generation. This is the default assignment for all pre-existing fragments.



Note that by changing the fragment options here, the check-box *Enable substructure filtering option* in the generation options window (Setup of Structure Generation Options) is automatically checked. If it is ever desired to not use these constraints, this box can simply be unchecked, or the fragment(s) can be set back to “ignore”.

All fragment information is project specific; the fragments defined in the *Fragment Tool* window are stored as a model of the project in the file `FragmentEditor.mol`. Additional fragments added in the *Fragment Filter* dialog are stored in the file `Fragment.sdf` in the project directory. Only the default fragments, and those fragments defined for that specific project will be included in this file.

3.10 Setup of Structure Generation Options

Before starting the structure generation some options should be defined. This can be done in the *Structure Generation Options* window (see the following figure). To display the window choose **Structure** → **Generate Structure** from the menu bar of the Correlation Table window or click  button. The various options will be discussed below.

Structure Generation Options ✕

Structure generator: Bruker

Execution Control

Filter results (keep only best ones)

Maximum number of generated structures (0=no limit)

Terminate after this many seconds (0=no limit)

Use multiple processors

Substructures

Enable substructure filtering

Ring rules

Structure does not contain any rings

Maximum ring length (0=no limit)

Forbidden rings lengths(Comma separated e.g. "3,4")

Required rings lengths (Comma separated e.g. "5,6")

Keep epoxydes also if the cyclopropanes are forbidden

Correlations

Use COSY correlations

Use HMBC correlations

Auto-eliminate invalid or long range COSY correlations

Auto-eliminate invalid or long range HMBC correlations

Maximum number of eliminated correlations (COSY+HMBC)

HMBC/COSY autoelimination policy: Optimal

Chemistry rules

Allow sp Carbons

Charged atom states

Use charged atoms

Allow modification of chemical formula due to 1H exchange with solvents

Minimum molecular charge

Maximum positive molecular charge

Special functional groups

Allow peroxides (-O-O-)

Allow triazane (N-N-N)

Allow O-N-O bonds

Advanced

Use Erlenmeyer rule (refuse R₂C[OH]₂)

Use Bredt's rule

Eliminate high energy structures

Use fast symmetry detection (Morgan algorithm)

Generate Structures Cancel

Figure 3.76: Structure Generation Options window highlighting the default generation options.

The individual sections illustrated in the figure above (e.g. Chemistry Rules) may be closed to minimize the dialog size.

3.10.1 Structure Generator

The Structure Elucidation module offers two structure generators, *Bruker* and *LSD* (Uni Reims). The Bruker algorithm was written by Bruker. The LSD (Logics for Structure Determination) structure generator is a public domain tool developed by Jean Marc Nuzillard from University of Reims, France. More information about this method can be found at: www.univ-reims.fr/LSD

The Bruker structure generator differs from LSD 3.4.8 as described below. With the Bruker generation algorithm:

- The *Chemistry rules* are evaluated if the check boxes are enabled.
- Symmetry information is used during the structure generation process, and not just as a filter at the end.
- While both algorithms will use the fragment tool as a filtering tool, **only the Bruker algorithm can use the partially assigned fragments as a building block.**

Choosing a different generation algorithm is incredibly useful in case of doubt. When a second independent algorithm generates the structure of interest, it is additional support that it is indeed the correct structure.

Additionally, in cases of trouble the alternative algorithm may be the trick to obtaining a solution. In cases when there is no symmetry, the LSD generation algorithm is *sometimes* a faster option than the Bruker generator.

3.10.2 Execution Control

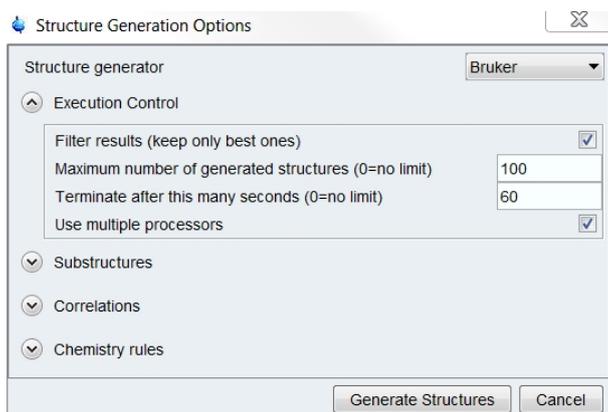


Figure 3.77: Execution Control Options

There are multiple methods to control/limit the amount of time a structure generation

- The **Filter results** option is used to keep the structure generator running until the maximum time elapses or all possible structure proposals are generated. If too many structures are generated, it is possible that the computer will suffer from memory issues, and will crash prior to finishing the process.

By enabling this **Filter results** option, the generated structures are sorted on the fly according a ^{13}C chemical shift prediction. Only the best ones are preserved, the number of structures is defined in the next field **Maximum number of generated structures**. When the filter option is active, the default value is 500, even if this is set to 0.

This is the **recommended** setting, especially for larger, more complicated or ill-defined structures.

- It is also possible to not filter the results according to chemical shift prediction, and only generate a set number of structures.

For this method, set the desired number of structures in the appropriate field, and uncheck the Filter option.

- It is also possible to have a time limit imposed on the generation process. No time limit (0) means that all possible structures will be created. Setting a certain time limit, for example 300s, forces the program to terminate structure generation after 5 minutes, even if not all possible structures were generated.

This limitation can be used along with the filter option.

3.10.3 Substructures / Ring Rules

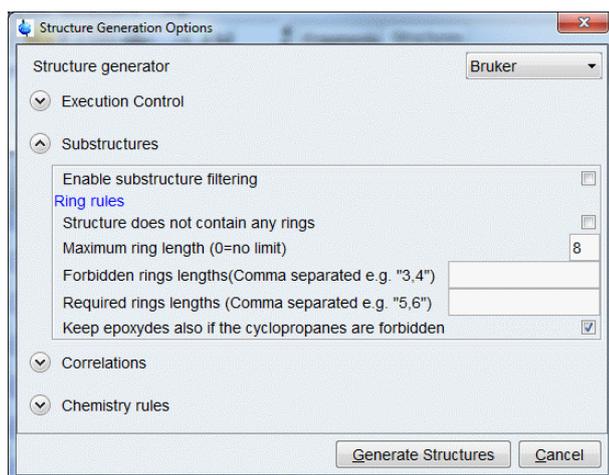


Figure 3.78: Substructure and Ring Rules Options

As described in the previous section, [The Fragment Tool \[81\]](#) is a powerful way of limiting and controlling the structures that are generated. For the fragments to be considered properly (as defined in the fragment filter tool) the **Enable Substructure Filtering** option must be checked. When a fragment is defined as *required* or *excluded* in the Fragment Tool or drawn in the main window, this box is automatically checked. If it is desired to run the generation without these fragment definitions, it is possible to simply uncheck this box.

Another option to sort out generated structures is offered by the *Ring Rules*. If there is an approximate idea of the considered molecule, it might be possible to decide if it contains a ring or not. By default, rings will be allowed, but it is possible to exclude structures with rings. It is also possible to define rules regarding required or forbidden ring lengths, for example, 3,4 would filter out 3 and 4 member rings (figure Structure Generation Options Window above).

As will be seen later in section [Viewing Generated Structures \[101\]](#), in CMC-se Version 2.x it is now possible to filter the structures after generation. By filtering the structures in this manner rather than ahead of time, more structures are ultimately generated, but the user does have the opportunity to at least view them prior to throwing them out, and always get them back if necessary.

If there is a contradiction between the *Fragment tool* and *Ring Rules* (e.g. required 6 member ring in a defined fragment, but ring rule of max ring length 5) the structure generation will fail to generate any structures.

Note that both the *Substructure filtering*, and *Ring Rules* constraints, are filtering options, which means all possible structures are generated, and then if a given structure does not fit the filter criteria, it will be thrown out, and not seen by the user.

3.10.4 Correlations

In this section, all of the non-required experiments that can be used by the structure generator are listed, and the user has the choice to use them or not. Note that only those experiments that are actually in the project and contain at least one correlation will be listed, so if the project does not contain the ADEQUATE it will not be listed in the list of options to use. With these options it is possible to switch on/off the usage of **all** correlations from a particular spectrum. Un-checking this box would be similar to disabling all those correlations through the table (section [Disabling Correlations in the Correlation Table \[▶ 74\]](#)).

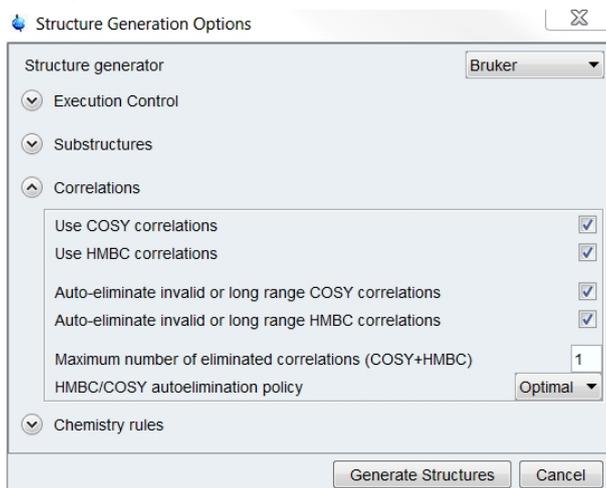


Figure 3.79: Correlation Control Options

In many cases using the default generation options as shown in the figure above, the Structure Generation will likely fail with an error message similar to figure below. This is a representative error of when no structures can be generated which fit within the constraints defined by the algorithms (Section [Structure Generator \[▶ 90\]](#)). As stated in the message, the user is prompted to use the *Auto-eliminate* switch in the generation options. If the user has already activated this switch, it should be examined to make sure that **BOTH the Auto-eliminate invalid long range COSY correlations and Auto-eliminate invalid or long range HMBC correlations** is enabled, and a reasonable number of correlations is set to be auto-eliminated.

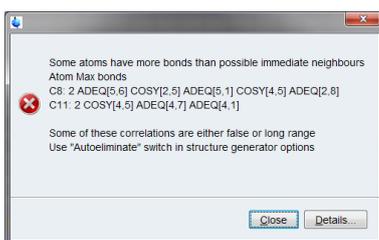


Figure 3.80: Error Message for an Incorrect Bond Count



It is important to note that **BOTH** the number of correlations to eliminate, **AND** the type of correlation to eliminate need to be defined. It is usually recommended to allow the auto elimination of **ALL** types of correlations that are possible

The Auto Elimination Option

The Structure Generation Algorithms assume a fairly strict definition for the correlations, and they are interpreted as follows:

- COSY represents a 3J coupling.
- HMBC represent a $^2J - ^3J$ coupling.
- H2BC, ADEQUATE, INADEQUATE are treated as special case 2J HMBC coupling.
- A TOCSY correlation requires the 2 protons be part of an uninterrupted chain of protonated atoms.

These are the “Standard correlations”, and are such that every cross peak correlates two or three skeletal atoms. Taking longer range correlations into account would reduce their selectivity and exponentially increase the complexity of elucidating the structure.

In some cases, non-standard correlations might be present in the Correlation Table due to either (or both) of the following:

- Long Range Coupling.
- False Correlation Assignment (i.e. real correlation assigned to the wrong atoms).

As neither the number of such “non-standard correlations”, nor their maximum length is known by the software, the user controls the number of possibly eliminated correlations.

Because the intensity of a peak in either an HMBC or COSY spectrum is dependent on absolute value of J , and not necessarily the length of the correlation, it is not always possible to use the data as a guide for which correlations to disable manually (section [Disabling Correlations in the Correlation Table \[74\]](#)).

The *Auto-Elimination* option, as its name suggests eliminates correlations from the generation process in an automatic manner.

Without the Auto-Elimination option checked, structures are generated, and as soon as a correlation is violated (e.g. two protons that have a COSY correlation, are longer than three bonds away), that particular structure is thrown out.

With the Auto-Elimination option checked, the user defined number of correlations can be violated prior to that structure being thrown out.

In the Structure Generation Options window, the user has the option to eliminate either COSY, or HMBC, or both depending on which box(s) are checked, as well as the maximum number of correlations (COSY+HMBC) that are eliminated.



Note that the TOCSY correlations CAN NOT be auto eliminated

The way the generation algorithms handle these invalid correlations has been improved in recent CMC-se versions compared to version 1.0.

In version 1.0, each correlation was handled independently. However as an improvement to the structure generation, they are now treated as more similarly to a bond.

- For example: H_a has an HMBC to C_b and H_b has an HMBC to C_a , while the old generator would have handled this as 2 independent correlations, the improved generator now treats this as a single correlation because it defines a single bond.
- The same is also true for in-equivalent protons of a CH_2 . If they both have a correlation to another atom, they are now treated as a single correlation.

The general process of setting this number of correlations to be eliminated is to start with a small number (3 for example) and increase it until a reasonable number of structures are generated.

Once structures are generated, they should be evaluated by the user (section [User Validation of Generated Structure Proposals](#) [▶ 101]) to choose a top candidate.

It is then advisable to add at least one additional correlation for auto-elimination and generate additional structures. It is quite possible that a better candidate might be generated with this additional flexibility.

- It is often necessary to eliminate more than 10% to get the correct structure.
- Conversely, there are often times, when elimination of only a small number of the correlations results in too much flexibility, and auto-eliminating fewer correlations produces the desired results.

In several cases when there is a large discrepancy in the intensity of correlations, the weaker correlations are likely longer range, but this is not always the case. For this reason, multiple policies for handling the elimination of correlations have been implemented.

Optimal

This is the default mode, as it is a good compromise in speed and accuracy. This mode looks at the intensities of HMBC and COSY correlations, and selects those with the larger volumes and sets them as definitely true, and does not allow them to be eliminated.

- For example: In a particular column, there are HMBC correlations with values of 10, 7, 6, and 0.2. The software would select the correlation with intensity of 10 as definitely 2J or 3J , and the correlation with intensity as 0.2 as most likely long range.

Exact

This mode will eliminate any and all correlations without regard to their intensities.

Fast

This mode is similar to that of optimal, but is more aggressive in defining those correlations which are definitely true.

As will be explained later (section [Viewing Correlations on Structures](#) [▶ 110]) it is possible to display these correlations which have been eliminated on the structures once generated.

The importance of this Auto-Elimination feature will be illustrated in the example data sets in the chapter [Examples](#) [▶ 131].

3.10.5 Chemistry Rules

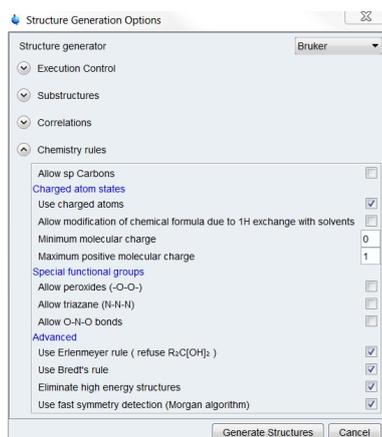


Figure 3.81: Chemistry Rules Options

The *Chemistry Rules* section of the Structure Generation Dialog Box can be expanded by clicking on the text as shown in the figure Structure Generation Options Window above. This section contains typical chemistry rules that are taken as default in the generation process, but can be modified as necessary for the given structure of interest.

Starting in CMC-se 2.1 is the possibility to generate structures containing charged atoms. This leads to large number of generated proposals and longer execution times. You may completely disable the usage of charged atom states (**Use charged atoms** option), or limit the total charge. Setting the maximum charge limit to 0 forces the structure generator to generate neutral molecules, however inner salts and functional groups like nitro group are still generated.

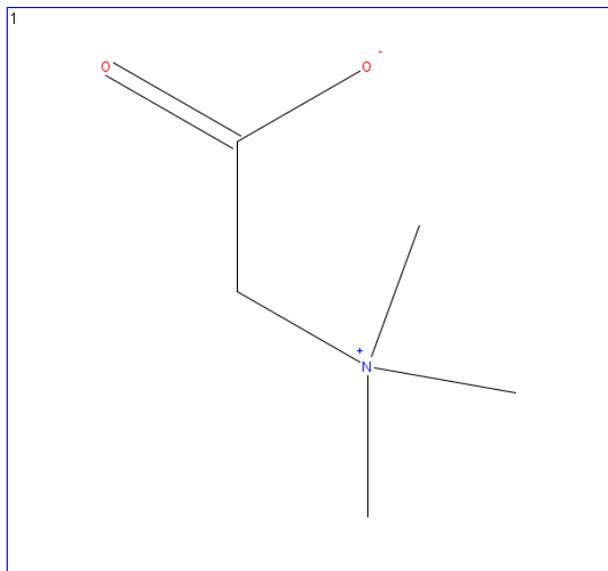


Figure 3.82: Structure of Betaine (trimethylglycine) generated in CMC-se 2.1.

In some cases, the solvent can act as a ¹H donor or acceptor. In these cases, the **Allow modification of chemical formula due to ¹H exchange with solvents** option should be checked. And then any extra/missing protons in the table will be handled by modifying the chemical formula to account for this chemical exchange.

In many cases, the chemical formula is an indication of a charged molecule. In cases like the figure below, where the initial Double Bond Equivalent (DBE) calculation results in a fractional or negative value, it will be highlighted red in the Correlation Table.

The screenshot shows a software window titled "CMC-ser: GibberellicAcid_C19H22O6 (C19H22O6)". The main area displays a correlation table with the following columns: Name, Shift, #H, Equiv, Hybr., Func. Group, H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16. The formula **C19H21O6** and **DBE=9.5** are highlighted in red. Below the table, there are buttons for H-C, H-H, and C-C. On the right, a "Fragments" panel lists: CH3, CH2, CH, C sp3, C sp2, and N.

Figure 3.83: A chemical formula resulting in a fractional or negative Double Bond Equivalent value.

When the chemical formula results in a fractional or negative Double Bond Equivalent value, it will be highlighted red in the correlation table, indicating that the formula is either incorrect, or that the user should allow for a maximum and or minimum molecular charge in the Generation Options.

While this may be due to an incorrect entry of the molecular formula, if it is correct, it indicates to the user that charged atoms, and a **Minimum molecular charge**, and or a **Maximum positive molecular charge** must be allowed/defined in the Generation Options, in addition to **Allow Charged Atoms** box being checked.

3.11 Structure Generation

After all desired options have been set in the Structure Generations Option window, clicking the **Generate Structures** button at the bottom of window will start the generation process. Part of the generation process is automatic on the fly ¹³C chemical shift prediction.

During structure generation, progress is displayed in a new pop-up separate window:

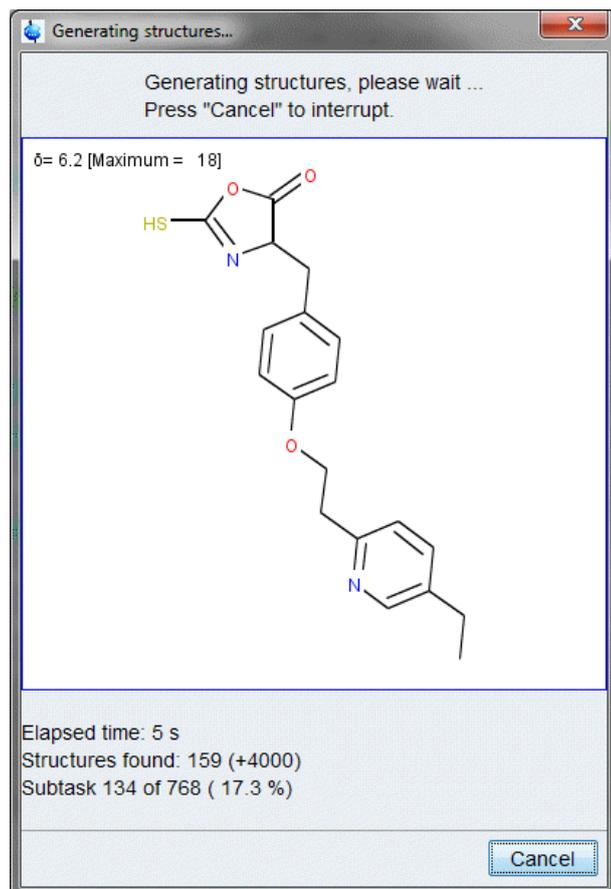


Figure 3.84: Generating structures progress display

Because the ^{13}C chemical shift prediction is running at the same time as the structure generation, the results are filtered based on the results of this prediction. The currently best ranked structure is shown in the window. This provides a rough idea on how successful the generation process is going to be. If no reasonable structures are being shown in this window, it is possible for the user to cancel the generation, and re-set some of the options, and start over.

As shown, this window also provides some information on the status of the generation, in terms of elapsed time, structures found, the status of the current subtask, including % completion.

The 2nd number in parentheses in the *Structures found* indicates the number of structures generated, but waiting to be checked for a variety of aspects, including if it is a duplicate structure, an isomorph of a previous structure, fully testing all the HMBC correlations on the full structure, and any filter elements including ring limits and sub-structures.

If the **Filter structures** option was checked (section [Execution Control \[▶ 90\]](#)), the generation will continue until completion, but the number of “structures found” will not exceed the number of maximum structures defined. Only the best ranked structures will be kept, the remainder will be discarded, thus preserving memory on the computer.

In the bottom right hand corner of the window is a **Cancel** button, offering the possibility to cancel the structure generation if necessary. The structures which have been generated up to that point will be stored in the project.

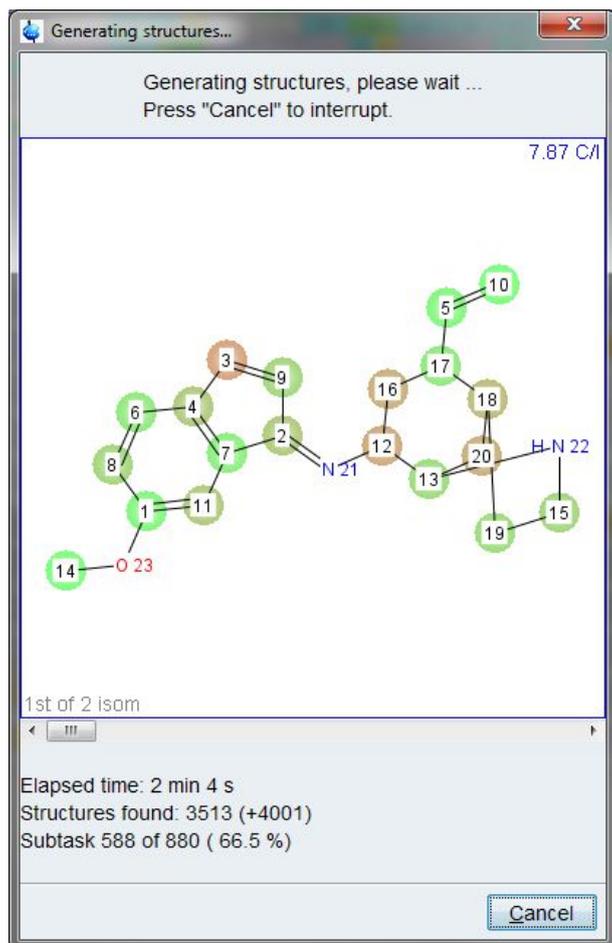


Figure 3.85: Generating Structures Progress Display

When the structure generation is finished the CMC-se Project window is updated as displayed:

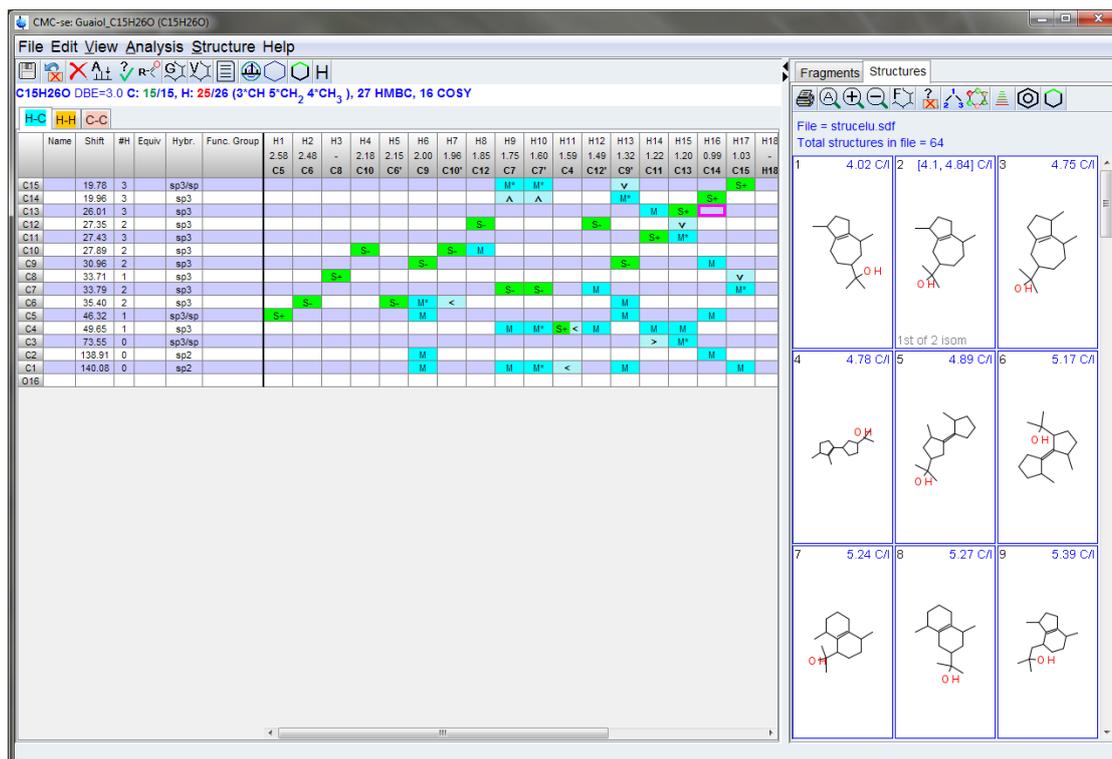


Figure 3.86: Updated Correlation Table with Generated Structures

The generated structures are automatically shown in a panel to the right of the Correlation Table.

The structure generator generates all structure proposals matching the input data and generation constraints (correlations, time, ring rules, fragment definitions, etc.). Some of these generated structures may be isomorphous. This means, they have the same connectivity but differ in atom assignments. The structure viewer groups these isomorphs together and writes the number of isomorphous structures to the left bottom corner (structure #2 in the figure above). Options for viewing these isomorphs and all structures will be discussed in section [Viewing Correlations on Structures \[110\]](#).

3.11.1 Possible Outcomes of Structure Generation

The following results of a structure generation are possible:

No solutions are found

In some cases, no structure was found compatible with the constraints in the input file (i.e. because of contradictory constraints).

- In this case a careful re-inspection of the Correlation Table and the spectra for correctness and the removal of suspicious information is necessary, either manually (sections [Actions Possible through the Combined Spectra Display \[52\]](#) and [Additional Modifications to the Correlation Table \[71\]](#)), or using a larger number of maximum auto-eliminated constraints (section [Correlations \[92\]](#)).
- It is also advisable to double check the Generation Options (section [The Fragment Tool \[81\]](#)), and remove any contradictory information (e.g. 6 member ring in a required fragment, but maximum ring length set to 5).
- It might be necessary to modify the Chemistry Rules of the Generation Options (section [Chemistry Rules \[95\]](#)), to allow for such things as charged atoms, or charged molecules.

In some cases, no solution was found because the correlations were not present to piece the molecule together. It might be possible to reach a solution if additional weaker peaks are picked from the data and added to the Correlation Table.

- This option should be used when necessary, but care must be taken to not use this as a default option, because many times these weaker correlations are actually long range, and thus problematic for the generation process. Whenever additional weaker correlations are added, the “Auto-elimination” option should be set to eliminate possible long range correlations.

In some cases no solution was found and the generation terminated because of the time limit:

- It might be possible to reach a solution if the time limit is increased.
- A re-inspection of the Correlation Table, and Combined Spectra Display are recommended just to make sure there are no mistakes, or additional peaks that could be added.
- Any additional information such as hybridization state, functional groups, multiplicity, as explained in section [Adding Additional Information to the Correlation Table \[76\]](#), should be included if available.
- If possible, adding fragments with known assignments to be used as structural building blocks will greatly increase the chance of reaching solution in a more reasonable time frame.

Even if structures are generated, if the generation finished because of a time-out limit, then it is recommended to perform any/all of the above operations to ensure that the correct structure is generated, and not just those that were generated prior to the time running out.

One or more reasonable solutions are found compatible with the input data

A file <project name>.sdf, containing the structures in MDL file format is automatically created, and structures can be evaluated as described in section [Structure Rating via Chemical Shift Prediction \[107\]](#).

It is recommended that even if reasonable structures were generated, that the user adds an additional degree of flexibility (increase the number of auto-eliminated correlations) and re-run the generation, to see if a better solution is present.

Several thousand solutions are found

In this case it is recommended to define more constraints if possible:

- Reduce the number of auto-eliminated correlations.
- Similar to above, where no solution was generated because of time limit, any and all additional constraints should be added.
 - Resolve fuzzy correlations if possible (section [Additional Manipulations to Correlations in Table \[75\]](#)).
 - Add any additional information such as hybridization, multiplicity, functional group (section [Adding Additional Information to the Correlation Table \[76\]](#)).
 - Define a fragment either as a filter element or a building block (section [The Fragment Tool \[81\]](#)).
- Additional experiments if not already included in the project (^{13}C , COSY, ^1H - ^{15}N HSQC/HMBC, H2BC, INADEQUATE, ADEQUATE, etc.) may also help in these cases.

It is also possible to modify the *Structure Generation Options* in section [Setup of Structure Generation Options \[88\]](#) to reduce the number of structures.

- It is possible to filter the results by one or both methods
 - Chemical Shift Prediction Values, only keep the best ranked structures.
 - Ring Rules, and only allow rings of certain lengths to be generated.
- Chemistry Rules
 - A variety of additional rare functional groups are allowed by default, it may be possible to exclude some scenarios such as “sp Carbons”.

3.12 User Validation of Generated Structure Proposals

After a successful structure generation it is up to the user to decide which of the generated possibilities is most likely the correct structure. The CMC-se module provides a variety of tools for this process. The first of which is the actual viewing of the generated structures.

3.12.1 Viewing Generated Structures

To inspect the generated structures you may use the *Structure* tab next to the Correlation Table. If this window has been closed after the structure generation was finished, it may be re-opened using the following methods:

- The **View Structure** button in the toolbar of the Correlation Table window .
- Choose **Structure**→**Display Structures**→**Show Last Generated Structures** from the menu bar of the Correlation Table window.

These commands display the structures contained in the file <project name>.sdf, which was created automatically during structure generation. The following figure shows a more detailed view of the Structure Panel:

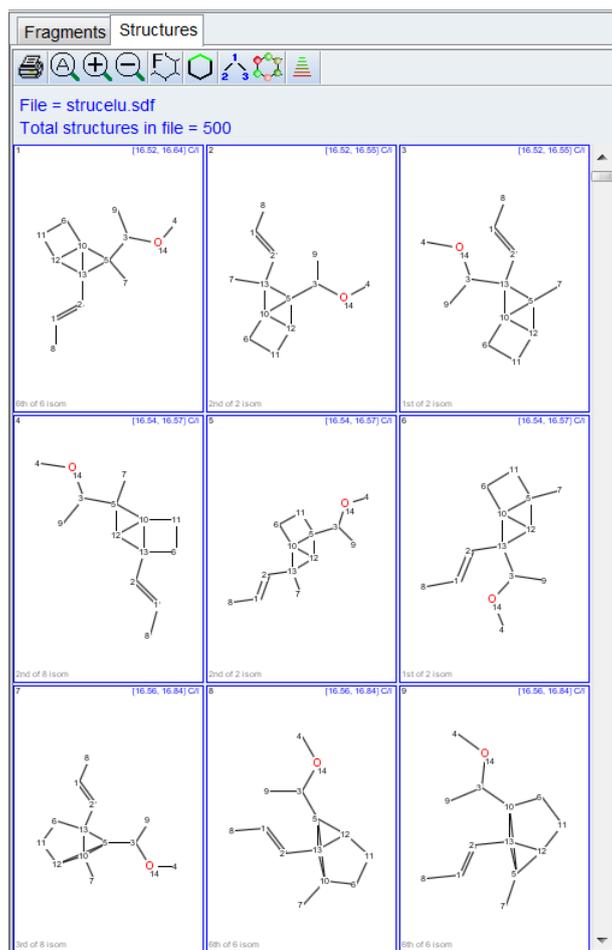


Figure 3.87: Structure Display Window

Note, the ^{13}C prediction ratings are shown in the upper right hand corner of the individual windows. All structures in this window have been sorted by this value with the best ranking structure in the upper left corner.

The structure display panel has its own toolbar and right-click context menu.

- The first button , can be used to print the structures and their chemical shift prediction results. The user has the option of printing all, selected or displayed structures to print.
- The next three buttons with the magnifying glasses  control the number of structures that are displayed in the window.
 - The maximum number of structures displayed in the window is a grid of 4x4. This is default and also obtainable by the  button.
 - The  button will decrease the grid by 1 in each dimension each time it is clicked.
 - The  button will increase the grid by 1 in each dimension each time it is clicked.
 - It is also possible to double left click on a single structure to fill the entire window with this single structure (and corresponding isomorphs which is explained below). A subsequent double click will return to the previous zoom level.

- The  button allows the user to filter the structures very similar to those options that were available during the generation (see the following figure). The benefit of using this option here is that the structures are still generated, and if necessary, the filter can be reset to show all structures again.

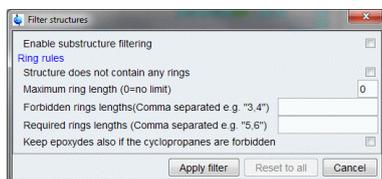


Figure 3.88: The filter structure dialog box that allows user definable elements to be filtered for or against in the generated structure set.

- The  button provides options for correlation highlighting (see section [Viewing Correlations on Structures \[▶ 110\]](#)).
- The next buttons control additional information displayed on the structures:
 -  Will turn on/off the atom name and number defined in the Correlation Table. By default, no atoms are identified in the main window. Turning on the **Atom Number** option will number the carbon atoms based on their numbering in the Correlation Table. If user names (in the *Name* column of the Correlation Table) have been entered, this option can be selected to number the structure accordingly.
 -  Will toggle on/off the color coding based on chemical shift prediction. Once chemical shift predictions have been performed (section [Structure Rating via Chemical Shift Prediction \[▶ 107\]](#)), the individual atoms can be color-coded according to how well they match the predictions. The match is coded using a color gradient between a green sphere (indicates a good match), and a red sphere (indicates a poor match).
 -  Allows one to resort the displayed structures according to different ranking criteria. The user can choose to sort the structures based on ¹³C predictions, or user defined ranking values.
 -  Will close the Structure Panel.

Right clicking on a structure will display the pop-up menu:

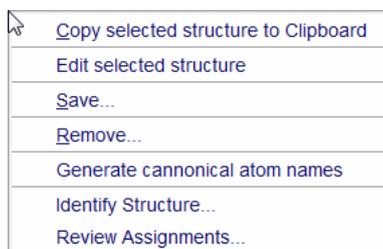


Figure 3.89: Pop-up window for Structure Options

- The first option of **Copy selected structure to Clipboard** will copy that structure to the computer's clip-board temporary memory for pasting in any other application.
- The option **Edit selected structure** offers the possibility to adapt the structure in a separate opening Structure Editor.
- The **Save** option selected offers multiple subsequent options for reporting the final structure possibilities after the user has decided which structure(s) are most likely correct.

- **Selected Structure To Data Set is a final save** type option to be used with the final structure has been determined. This option saves this particular structure to a `.mol` file, and a PDF document summarizing the structure and assigned atoms and correlations to the defined `selu-result.mol` and `selu-result.pdf` files in the same directory path as project directory (default name `strucelu`). This routine also automatically fills in the “Name” column of the Correlation Table with the Canonical Atom Number. The PDF report (as shown in the following figure) will include these numbers.
- **Structure and Report** option allows the user to select individual structure(s) and save the same `.mol` and `.pdf` file as above, but with a user defined name and location. This option is useful when more than one structural candidate is under consideration and reports for multiple structures are desired.

The four pages from the report generated from the above save options. Tables of the chemical shifts, the generated canonical numbers, as well as the original atom numbers from the Correlation Table are shown. The ^1H and ^{13}C spectra are also shown with the peaks labeled with the canonical numbers to correspond to the numbering on the structure that was saved. And the final page is a graphical summary of the correlations that can be explained by the structure on the left, and those that don't fit the structure (given the ^2J - ^3J limits explained previously section [Chemistry Rules \[95\]](#)) on the right:

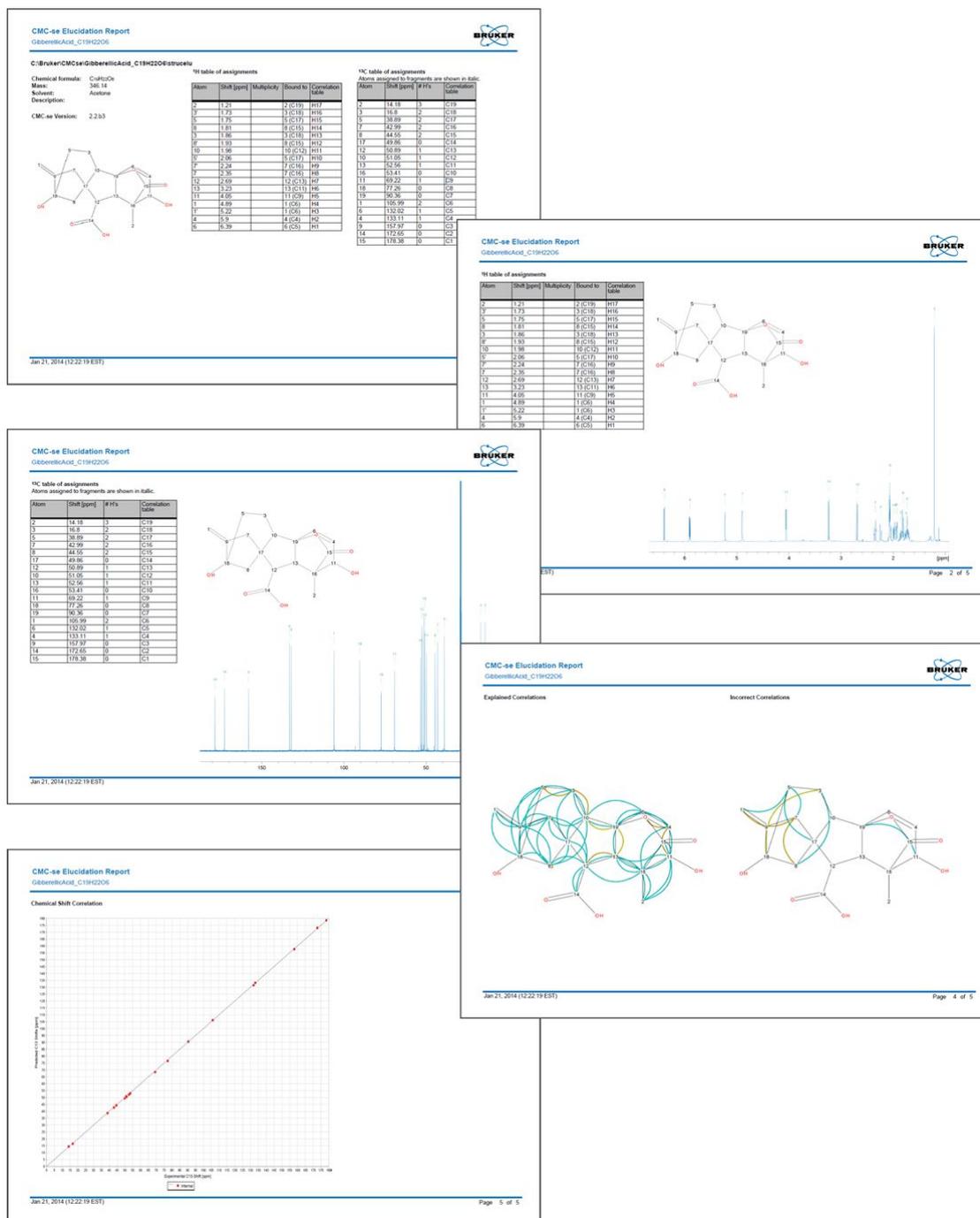


Figure 3.90: The 5 pages from the report generated from the above "save" options.

The report above includes 1) Tables of the chemical shifts and assignments, 2) The 1D ¹H and ¹³C spectra, 3) A summary of the correlations both valid and invalid, as well as 4) A graph of how well the experimental results match the predicted ¹³C chemical shifts.

- The **Remove** option selected offers three subsequent options for removing un-wanted structures from the <project name>.sdf file in the CMC-se project.
 - **Selected Structures**, as the name suggests will remove any selected structure(s). Note it is possible to select multiple structures throughout the .sdf file using the shift key.
 - **Subsequent Structures** option will remove all structures after the current highlighted one.

- *Low-Rated Structures* option allows the user to set a threshold chemical shift prediction value, and remove anything with a higher (worse) value.
- The **Generate canonical atom names** option will automatically fill in the *Name* column of the Correlation Table with the Canonical Atom Number. Note this is different than IUPAC numbering, but it is the IUPAC standardized method of numbering a structure from the InChI code.
- The **Identify Structure** option generates an InChI code for the selected structure, and will then search the server at the University of Vienna (<http://nmrpredict.orc.univie.ac.at>) for this particular structure. If it is already in their server, a separate web page will pop up with more NMR data relating to this compound (See chapter [CSEARCH Structure Identification \[▶ 45\]](#)).
- The **Review Assignment** option makes an assignment check for publications via request to CSEARCH Robot Referee (See chapter [Check Assignment for Publication \[▶ 119\]](#)).

Some of the structures generated will be isomorphic, meaning they have the same connectivity, but differ in atom assignments. As shown in structures #2, 10, 12, 13 of the figure Structure Display Window above, if isomorphic structures are present for a given structure, the number of such will be included in the lower left hand corner of the box. Double clicking on this structure will result in a display as shown in the figure below – where the initial structure is still shown on the top window, and the prediction values highlight the best and the worst values for the different isomorphs. The isomorphs are shown below, automatically displaying the atom numbers on the structures, shedding light on the differences between the different assignments. The connectivity of the structure is the same; the only difference between the two isomorphs is the assignment of carbons #9 and #6 in the 5 member ring:

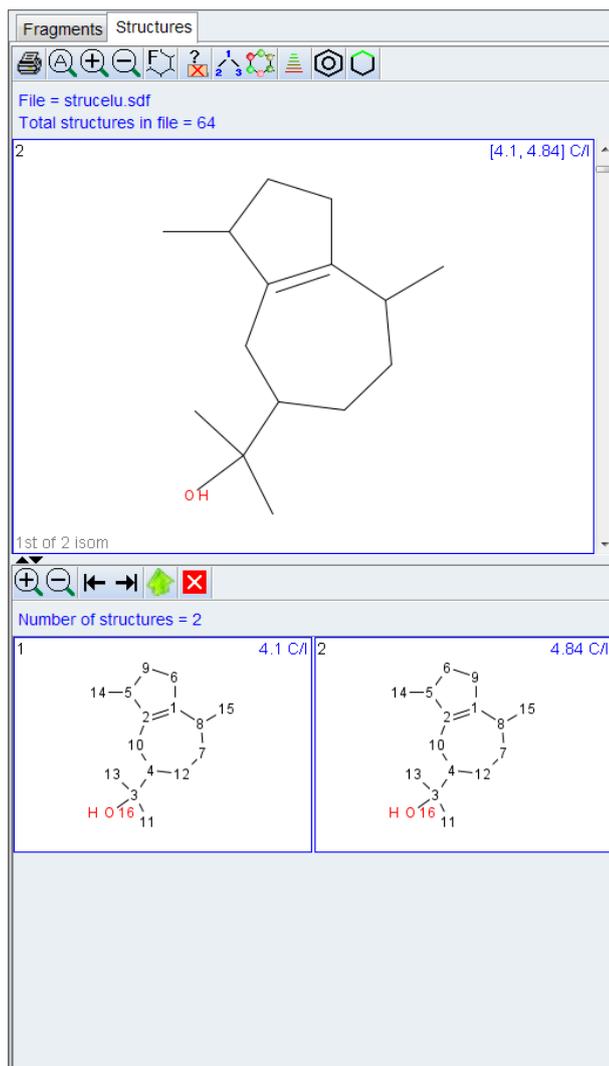


Figure 3.91: Structure panel illustrating multiple isomorphs of a generated structure

In some cases, when the different isomorphs are all valid structures for final consideration, it

is possible to un-group them, and add them all to the full list with the  icon. This will ungroup all isomorphs of that particular structure into individual structures in the main window, where they can be equally evaluated with the remaining structures.

The  button, or double clicking on the main top structure will return the display back to the original display without the individual isomorphs.

3.12.2 Structure Rating via Chemical Shift Prediction

In older versions of CMC-se, it was advisable to run a separate Chemical Shift Prediction on the generated results.

This Chemical Shift Prediction and ranking is now automatically performed during the generation process.

Using an internal prediction routine, the ^{13}C chemical shifts for each generated molecule are predicted, and compared to the experimental values. A rating with how closely they match is automatically provided, and these are the values that are displayed in the upper right corner of the Structure Panel (see figure Structure Display Window.).

After prediction, the mean deviation between the predicted shifts and the experimental shifts will be calculated and displayed with the structure. During this rating process, the structure display is sorted according to the prediction results. The most probable structures, and thus those shown first, will have lower deviation values.

If a calculated structure exists in multiple isomorphous states, the shifts will be predicted for all isomorphs, and the range of deviations is displayed in the main box, and this can be expanded to show the individual ratings below (see for example figure above). The sorting will rank according to the best ranked structure. When the **Display names/numbers** option is selected, the atom numbering/names of the shown structure always corresponds to the best ranked one.

In most (but not all) cases the correct structure will typically have a standard deviation less than 5, and usually, structures with a deviation above 10 can usually be disregarded. In most cases where the deviation is less than 2, this is a strong indication that it is a correct structure.

Using chemical shift prediction is a widely accepted method for the ranking of structure elucidation results, but there is no guarantee that the best predicted (lowest deviation value) structure is the correct one. More critical examination of the structures and various correlations as described in section [Viewing Correlations on Structures \[p 110\]](#) below are necessary in most cases to be truly confident in the validity of the ranking results.

As mentioned earlier when discussing the various icons in this structure display window, it is possible to color-code the generated structures based on ^{13}C chemical shift prediction results. The results of this action are illustrated in the figure below. As shown, the atoms are given a color code according to how well they matched the prediction results. This color coded ranges from green (good match) to red (poor match). If the mouse is hovered over a single atom, there will be a small pop-up indicating how far (in ppm) the predicted value was from the experimental value.

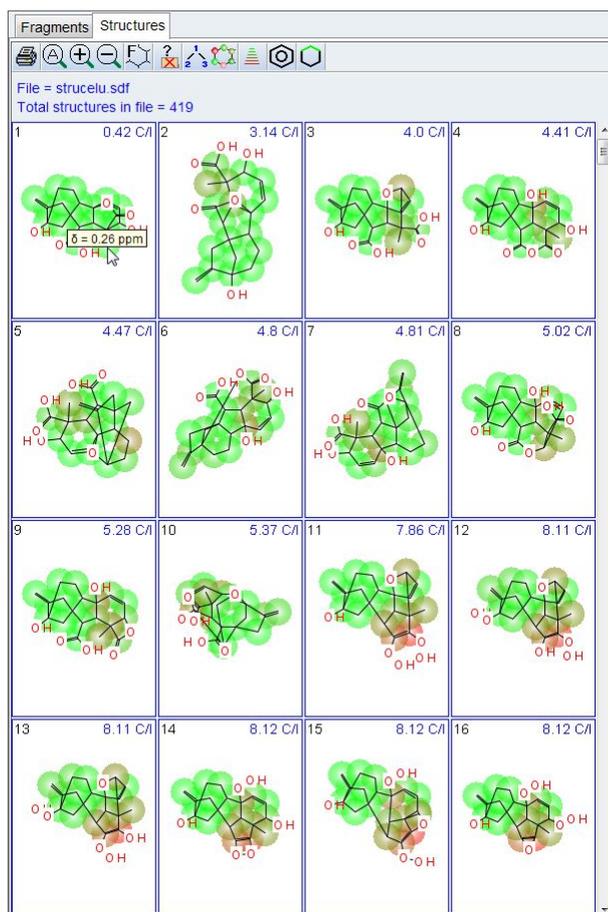


Figure 3.92: The structures have been color coded based on the chemical shift prediction results.

As described in the previous section, right clicking on a structure and choosing **Properties** from the pop-up menu gives a summary of all important information for this particular structure (see the figure below). In the displayed window there is detailed information of the standard deviation and the used method(s) for prediction.

Additionally, there is possibility to set a user specific ranking for the structure and if desired a comment. Both will directly appear in the corresponding Structure Panel.

The user ranking can be helpful when the preferred structure (arrived at via methods described below) and the best predicted structure are not identical. As mentioned above in section [Viewing Generated Structures \[101\]](#), it is possible to sort by any ranking currently available, including this user ranking.

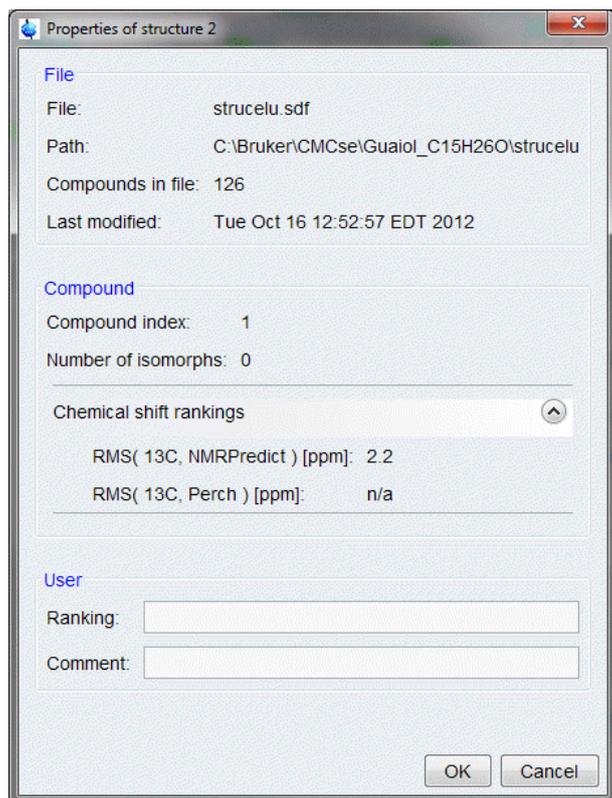


Figure 3.93: Properties of a Structure Selected from Structure Panel

3.12.3 Viewing Correlations on Structures

An important step in validating or refusing any of the generated structure possibilities is visualizing the experimental correlations on the structure itself; not only the correlations that were used in the generation algorithm, but also, those that were eliminated, either automatically, or manually.

Also starting with CMC-se version 2.0 was the ability to use even more types of correlations in this capacity for verification purposes. In addition to those experiments discussed and used previously for generation purposes, the following types of data can also be included in the project, and the correlations used for evaluation purposes in the same manner as the others.

- ^1H - ^1H NOESY / ROESY
- ^1H - ^{13}C HSQC-TOCSY

Viewing Correlations from Selection

Correlations on Structures The atoms of the structures in the structure panel are linked to the corresponding atoms in the Correlation Table. Clicking on an atom in the structure will highlight this atom and correlations present in the table as shown below:

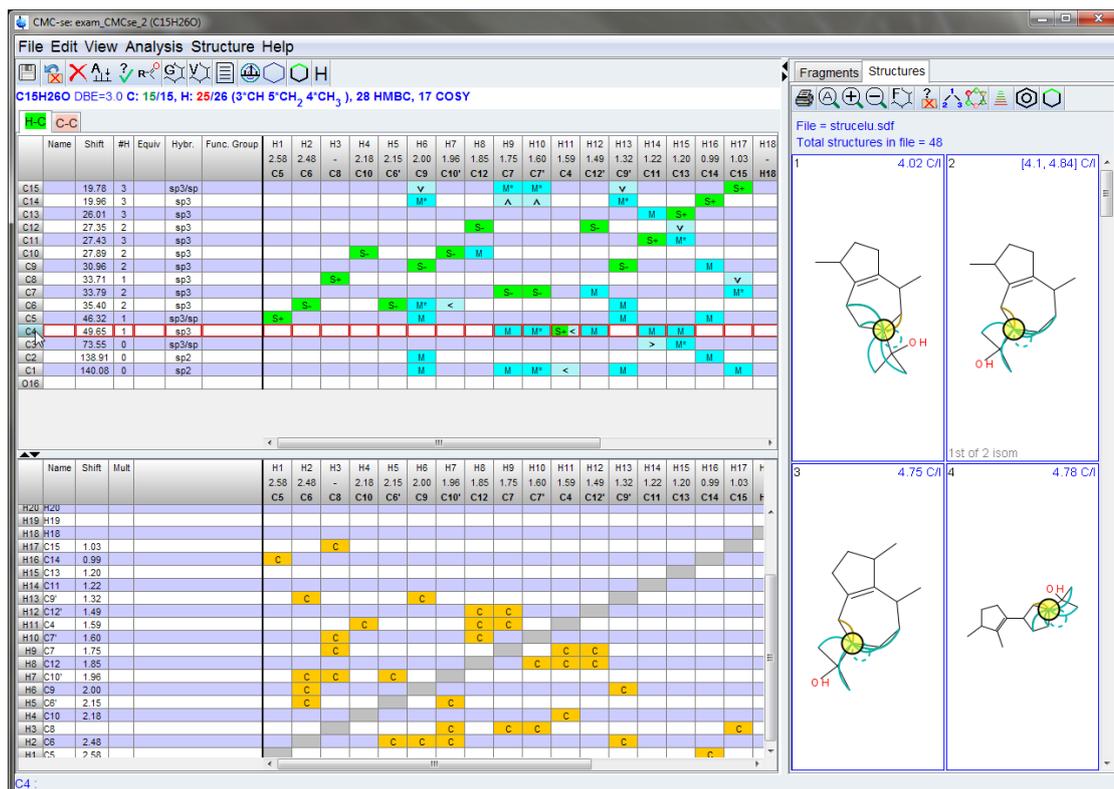


Figure 3.94: Linking of atoms in a structure to the contents of the Correlation Table

The selected carbon atom is highlighted in the Correlation Table and all correlations (HMBC and COSY) are displayed on the structure with curved arrows. The HMBC correlations are teal, while the COSY correlations are brown.

The cells of the table are also tied to the structures. Clicking on individual correlations in the table results in that correlation being highlighted in the structures.

- An HSQC correlation in the table will highlight that carbon atom (and inferred proton) with a colored circle in the structure.
- Clicking on any other correlation (or empty cell) will result in an arrow representing that correlation (present in the table or not) on all the structures. This is illustrated for an HMBC correlation in the figure below. Shown in this figure is an example of displaying a *fuzzy* HMBC correlation. The solid line represents the correlation that was actually clicked on (H9-C14), and the dotted line is the alternative assignment (H9-C15):

Structure Elucidation Basics

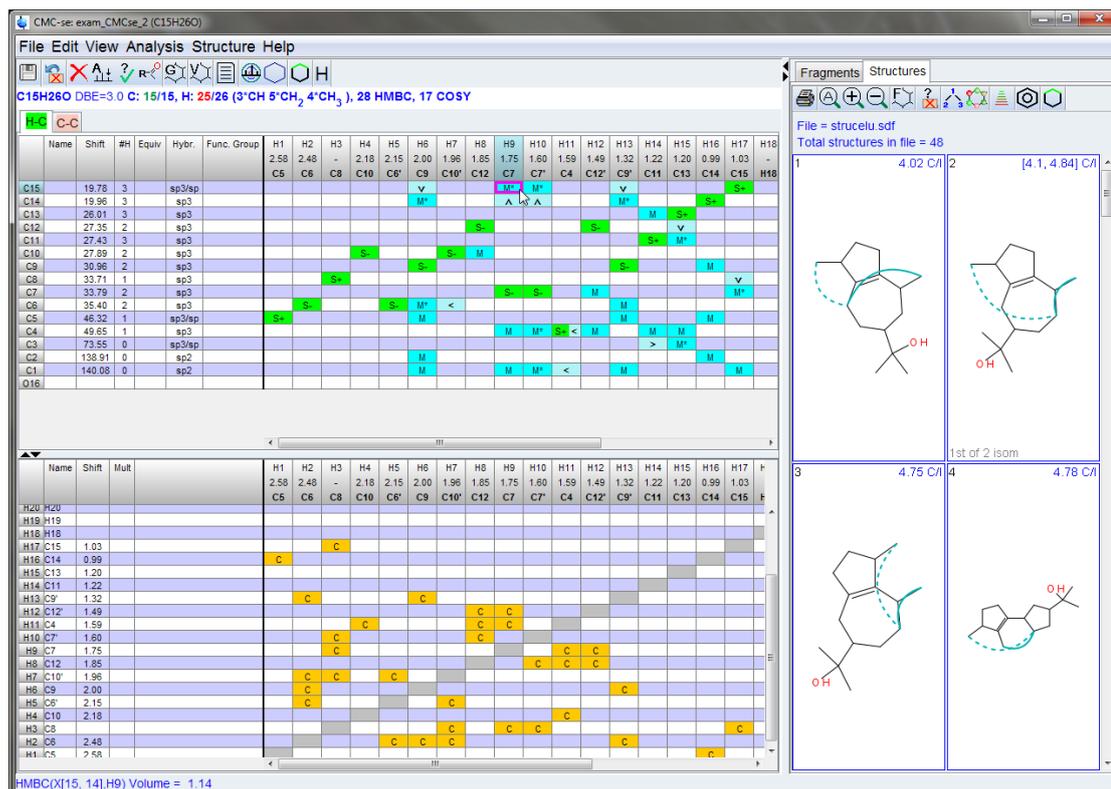


Figure 3.95: Illustration of highlighting a single “fuzzy” HMBC correlation from the table on the structure.

In the figure above, the solid line represents the correlation that was clicked on. The dotted line represents the alternative ¹³C assignment of the correlation.

Correlations in Combined Spectra Display

The data in the combined spectra display is also tied directly to the structures. A correlation does not have to be included in the Correlation Table to be viewed on the structures.

For example, the HSQC-TOCSY shown in the following figure, these correlations are not included in the table, but the user can click on a peak in the spectrum, and have a visual representation of that correlation. In this case the selected HSQC-TOCSY peak gives more confidence that structure #2 is correct despite it having a lower chemical shift prediction result.

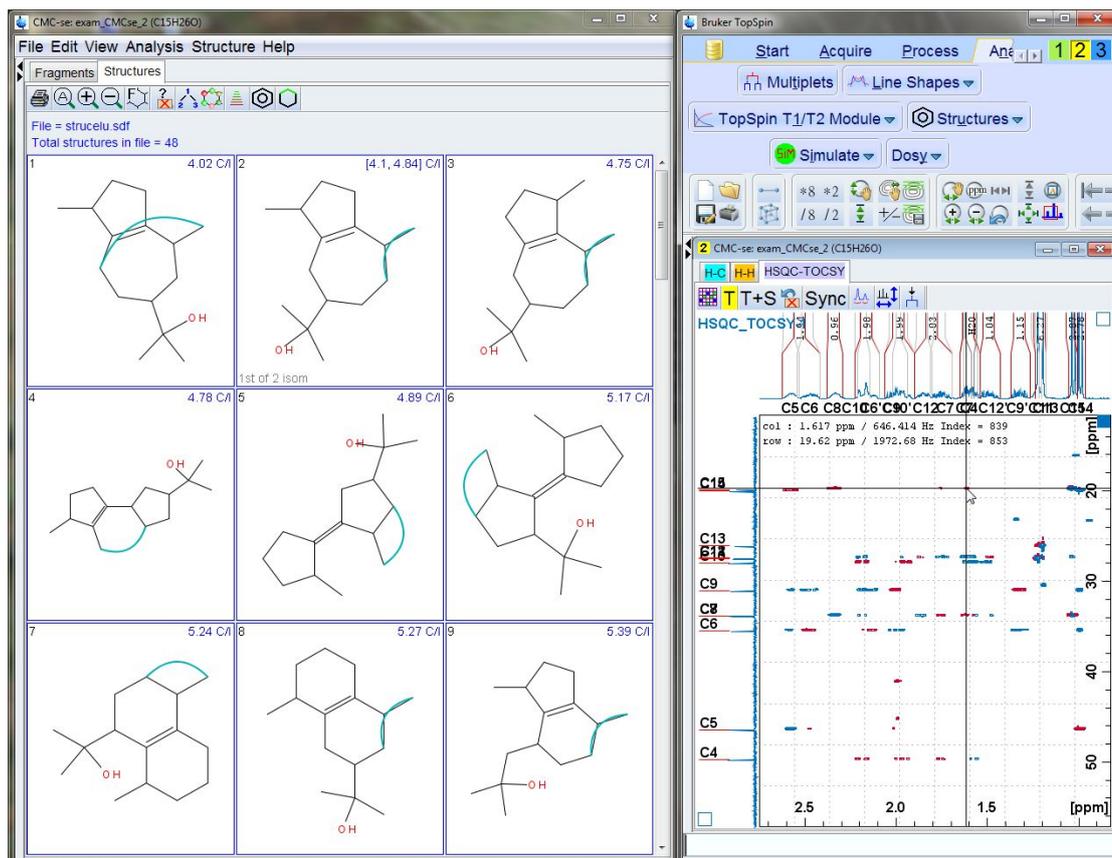


Figure 3.96: Clicking on a peak in the Combined Spectra Display will also result in that correlation being highlighted on the structure candidates for evaluation.

This can be done with any spectra, including those that were used in the generation process. This allows the user to go back to the experimental data and evaluate any weaker correlations that might be present, but likely long range so not included in the Correlation Table due to the problems it would cause with the generation process.

Viewing Correlations topically

For the visual inspection of the result the Correlations can be highlighted separately by using the Highlight Correlations  functionality. The Popup menu lists the different possibilities (see following figure):

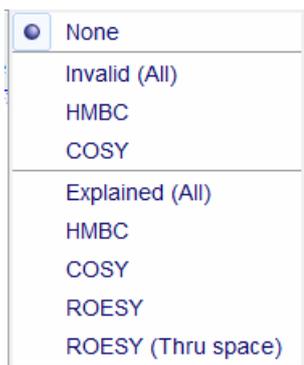


Figure 3.97: Popup Highlight Correlations

Viewing
Automatically
Eliminated
(invalid)
Correlations

An important aspect of validating a structure proposal includes examining those correlations that were eliminated by the software as invalid, and making a judgment call as to whether they are real long range correlations, or not.

Clicking the popup menu item **Invalid (All)** or in the menu **Structure**→**Show invalid correlations**, highlights all correlations which are violated in the selected structure. To filter the displayed correlations choose the respective experimental type (**HMBC** or **COSY**). The correlations are highlighted with pink boxes in the Correlation Table and with arrows on the structure. This is shown in the figure below.

Note that this option does not show the *invalid* correlations for each structure at a time, but rather, the correlations violated in that selected structure are shown on all of the structures for comparison of the *invalid* correlations between structures.

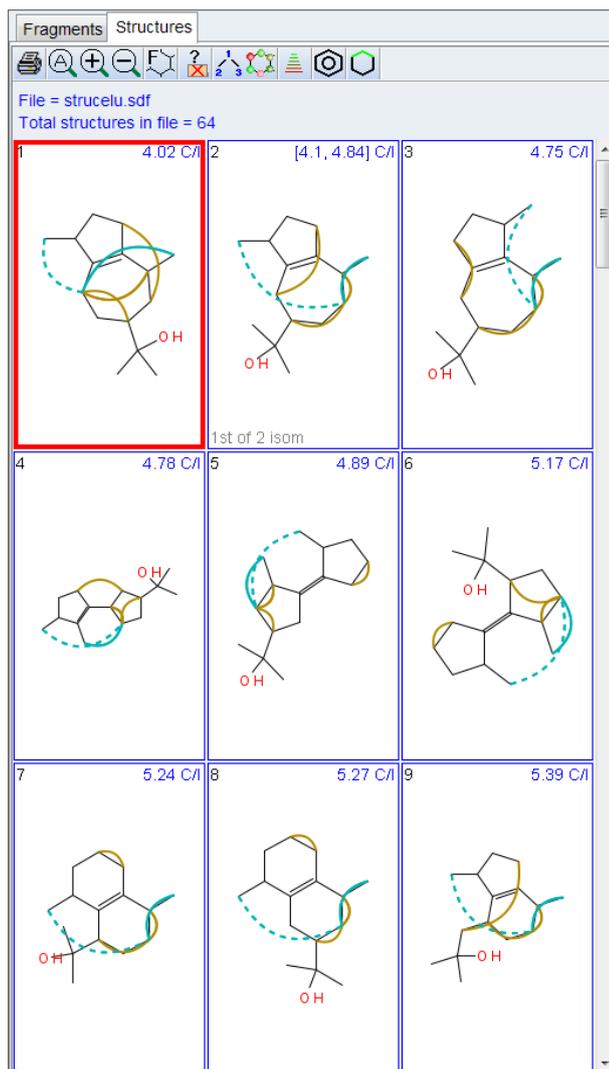


Figure 3.98: Highlighting the invalid correlations found for the first structure on the rest of the structures.

In the figure above, clearly some of these invalid correlations (HMBC and COSY) make more sense on later structures than they do on the selected structure. This provides an additional tool beyond chemical shift prediction to help the user to evaluate which of the proposed structures is correct.

Using the tie between the Correlation Table, and the Combined Spectra Display, the user can now determine if the *invalid correlation* is from a true long range coupling consistent with the structure, or from a miss-assigned atom which would in-validate the structure proposal. As shown in the following figure, the invalid correlations are highlighted on the table, clicking on one of them, will automatically move the cursor to that spot in the spectra, and as seen here, this long range COSY correlation happens to be a comparatively weak signal in the spectrum, so it is definitely possible that this is a long range ($>^3J$) correlation. Additionally, it is a crowded region of the spectra, so it is also possible that it was miss-assigned. Thus the correlation has an explanation, so the structure candidate has not been invalidated.

Structure Elucidation Basics

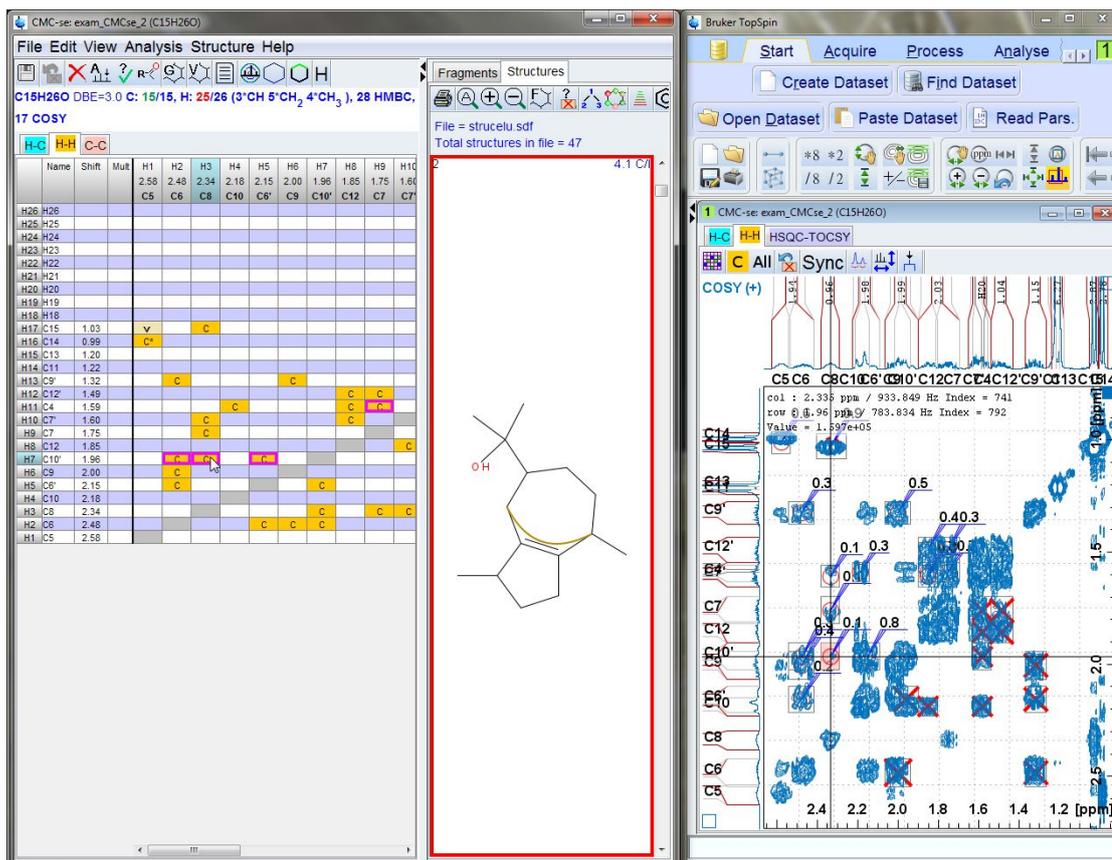


Figure 3.99: Displaying invalid correlations on structure, Correlation Table, and how this information can be evaluated in the Combined Spectra Display with the Correlated Cursor.

Viewing
Atoms
without
Constraints

Also important in the structure validation process can be to visualize those atoms without any constraints. This is accomplished via the following:

The command **Structure**→**Display atoms without any constraints** highlights carbons without any correlations.

As shown in the following figure, this results in those atoms being highlighted in the table, and pink circles on those atoms in the structure:

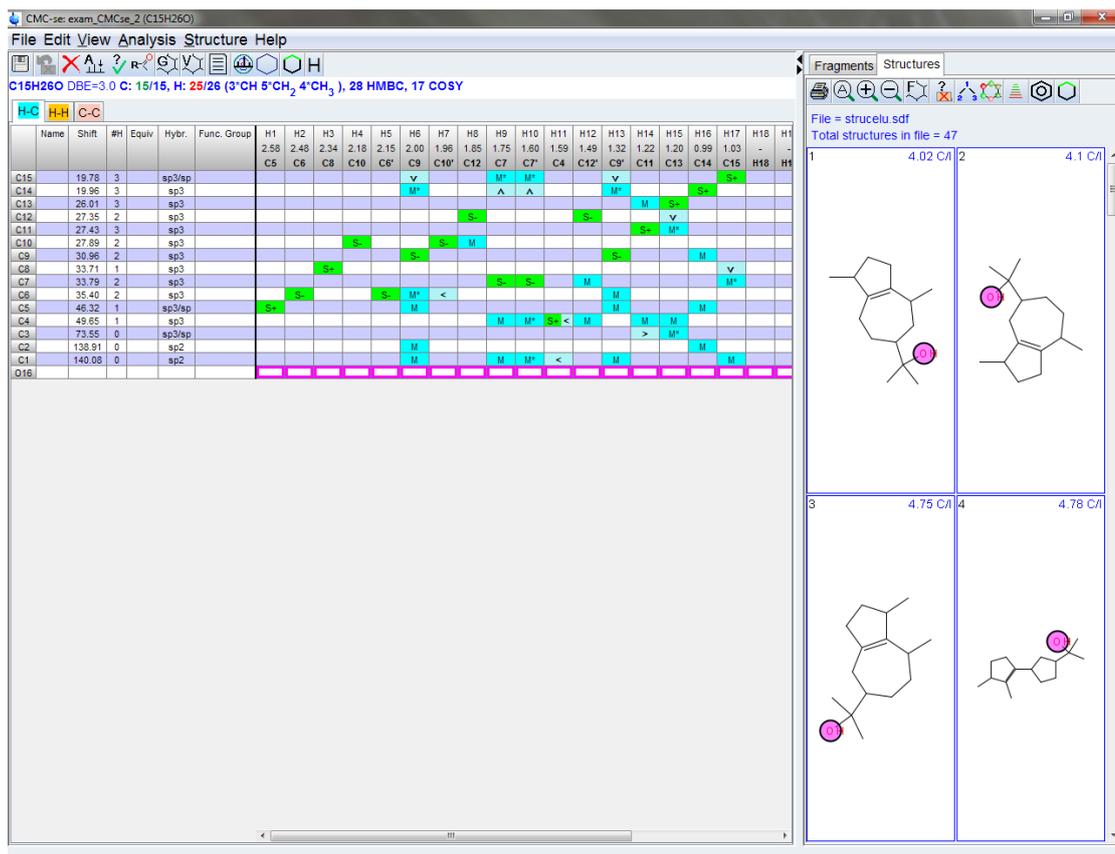


Figure 3.100: Displaying atoms without any constraints

Viewing
Correlatio
ns
between
Protons

Having information about nearby protons from ROESY or NOESY experiments can be an important knowledge for a better understanding about the spatial resolution of a structure. In case there are separate protons drawn in the structure the correlation will be highlighted respectively (to add explicit protons use the option **Edit selected structure** in the context menu of the structure).

Structure Elucidation Basics

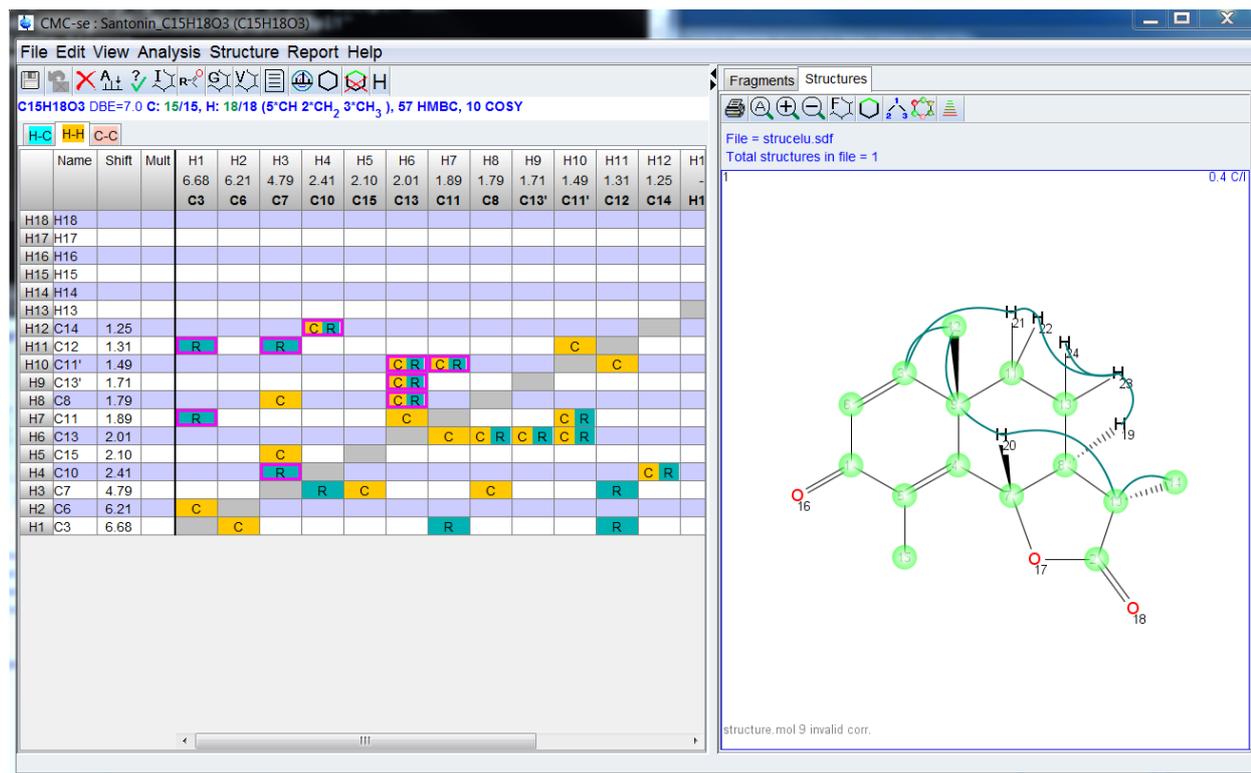


Figure 3.101: Highlight Correlations ROESY

Correlation highlighting between protons can give the appropriate information about the stereochemistry information. This information can be added to the structure via the context menu pressing the right mouse button on the bond (see the following figure).

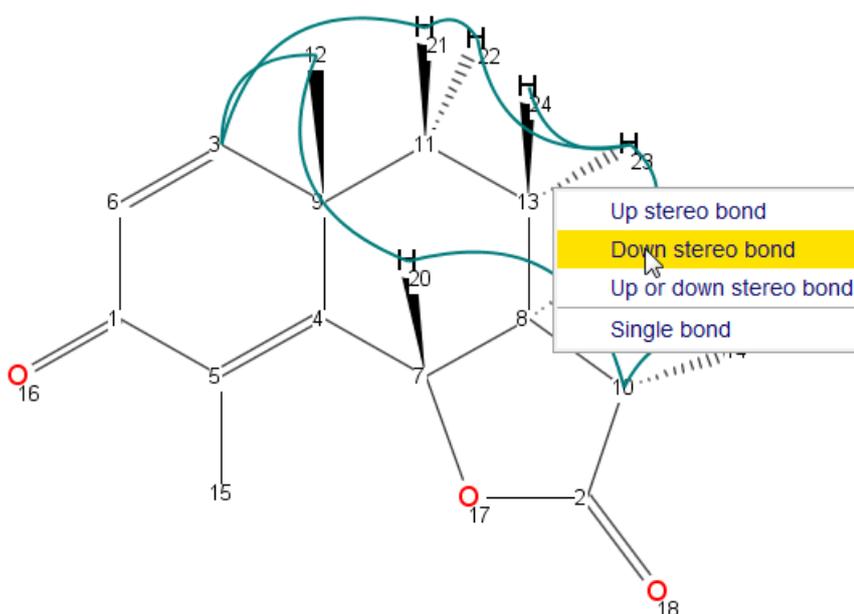
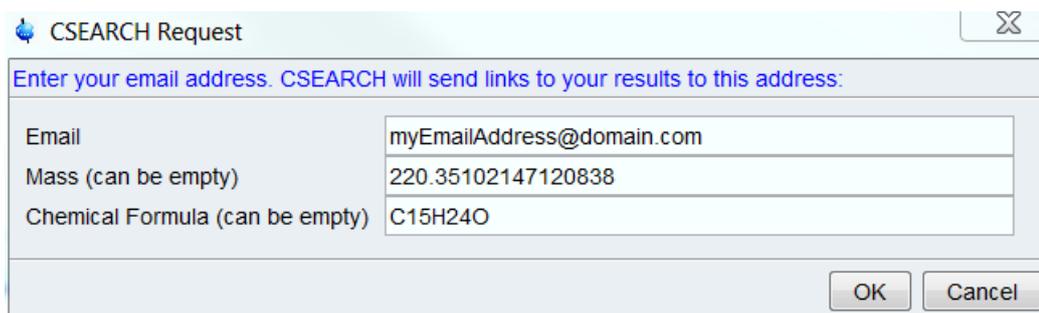


Figure 3.102: Stereochemistry information from proton correlations

3.12.4 Check Assignment for Publication (CSEARCH Robot Referee)

As a last step in the plausibility check of generated structure, the interface to CSEARCH Robot Referee allows an assignment check to get an evaluation of the result ahead a publication and ready for transfer in the manuscript. The result is an email holding URLs to details about the evaluation, the spectrum prediction and the classification of the assigned structure. The structure with its assigned chemical shift values is accessible for importing into Isisdraw/Symyxdraw.

Send the request by choosing the **Review Assignment** entry in the context menu (right clicking the mouse) of the structure. CMC-se automatically generates the request (see the following dialog):



CSEARCH Request	
Enter your email address. CSEARCH will send links to your results to this address:	
Email	myEmailAddress@domain.com
Mass (can be empty)	220.35102147120838
Chemical Formula (can be empty)	C15H24O
OK Cancel	

Figure 3.103: CSEARCH dialog

In a next step the default email program starts with an automatically filled email. Following the instruction in the email text, it is necessary to attach the data file located in the given file path. With the Control V key the filename entry can be pasted in the mail client attachment browser, because CMC-se copied the filename into the clipboard. The email subject is always the program name, ending with a number coding the current date and a time.

The result will be send to the given email address after few minutes. The email includes login data to get access to linked evaluation result.

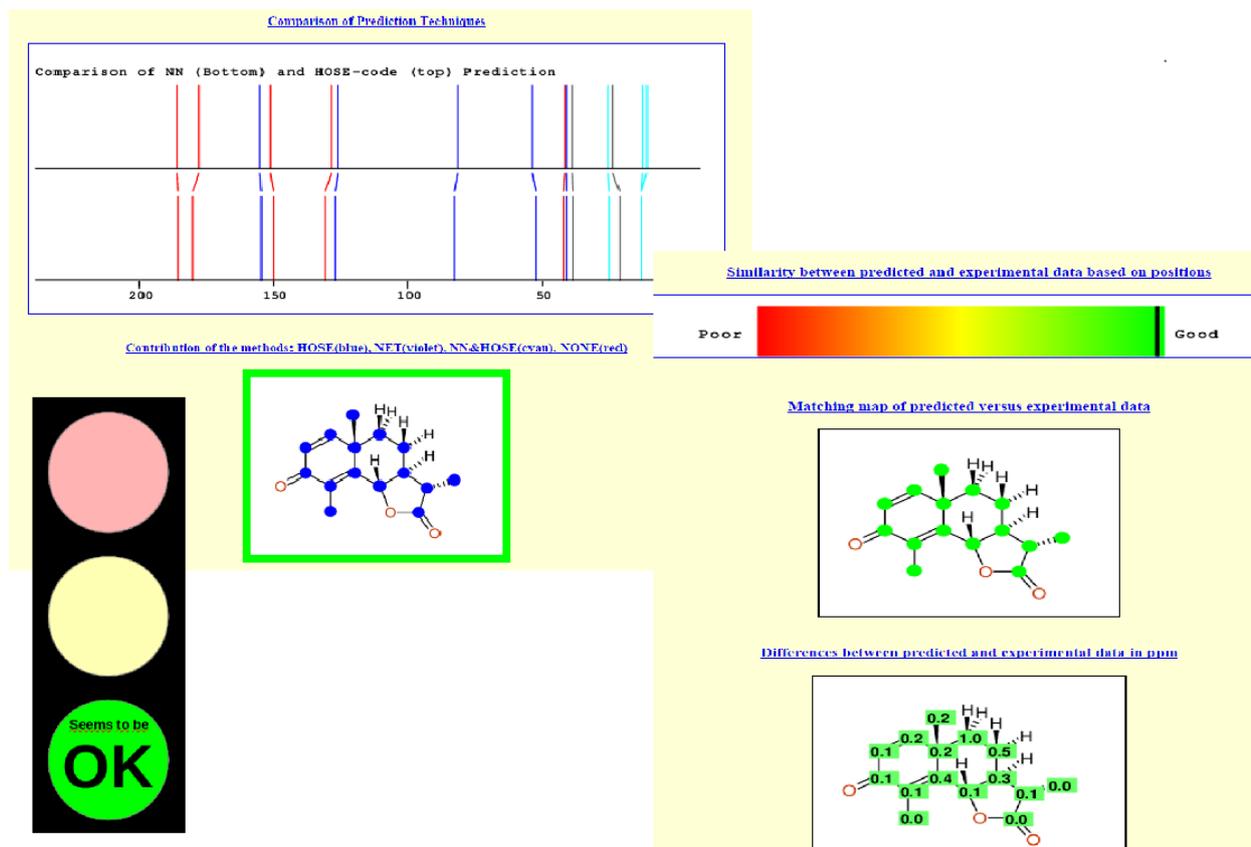


Figure 3.104: CSEARCH Robot Referee Result

The figure above shows an example for the result. For more information about CSEARCH Robot Referee use the following link:

<http://nmrpredict.orc.univie.ac.at/c13robot/robot.php>

3.13 Process Overview

The preceding sections went through the different features of CMC-se in detail. The flow charts in this section describe the intended work-flow of the CMC-se process.

They serve as a general outline and review for the previous content. Additionally they illustrate what actions can be taken upon the result of too many or too few structures generated. While the general process is generally similar, different things should be examined, and changed depending on whether the initial structure generation resulted in too few or too many possible structures.

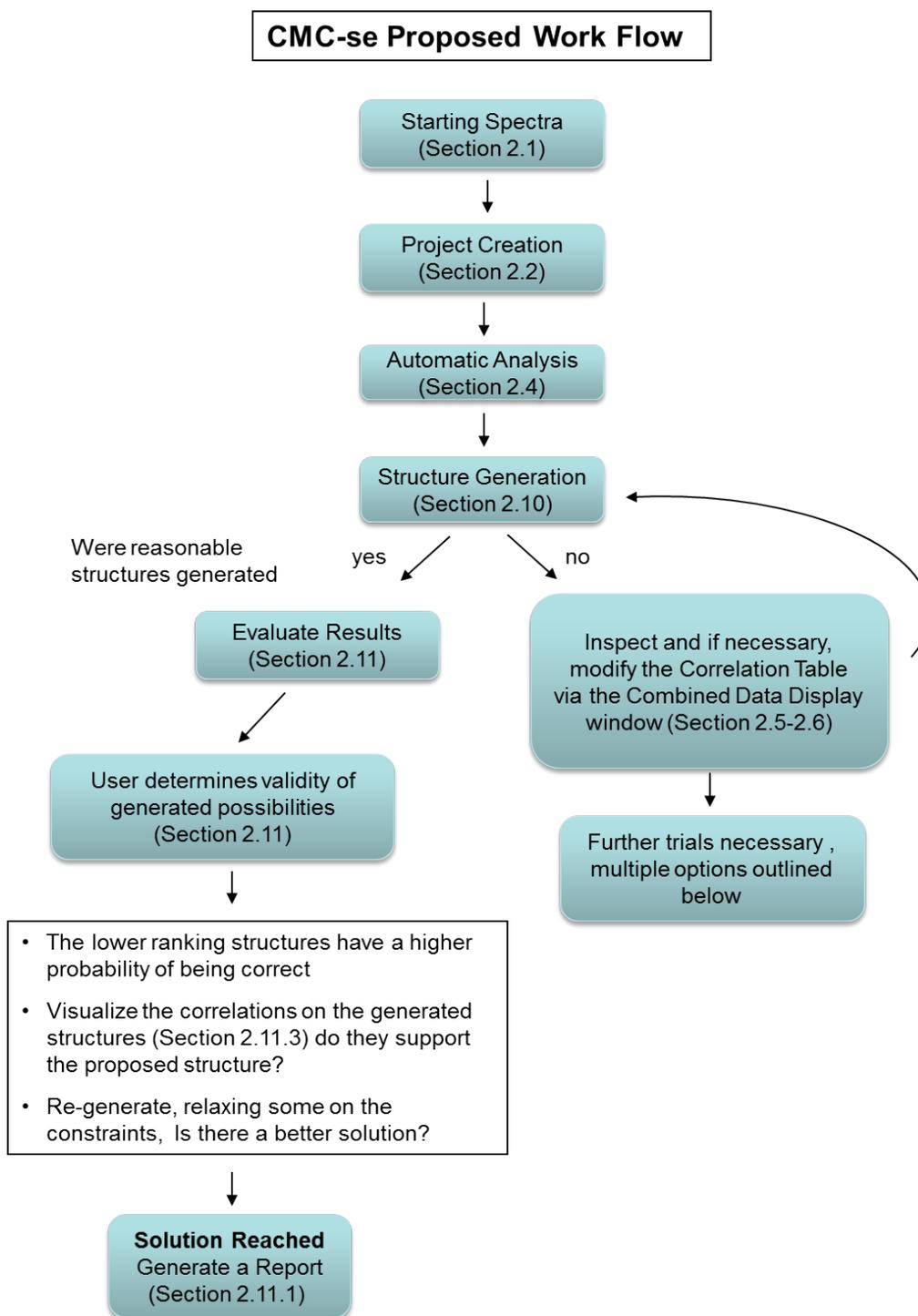


Figure 3.105: Proposed CMC-se Work Flow

As outlined in the following figure, in the case that zero structures are generated, or a few unreasonable structures are generated, the user should **first look for and try to remove any inconsistent constraints**, either by re-defining them in the Correlation Table via the Combined Spectra Display (section [Actions Possible through the Combined Spectra Display \[▶ 52\]](#)), and/or automatically via the “auto-disable” option of the structure generation (section [Chemistry Rules \[▶ 95\]](#)).

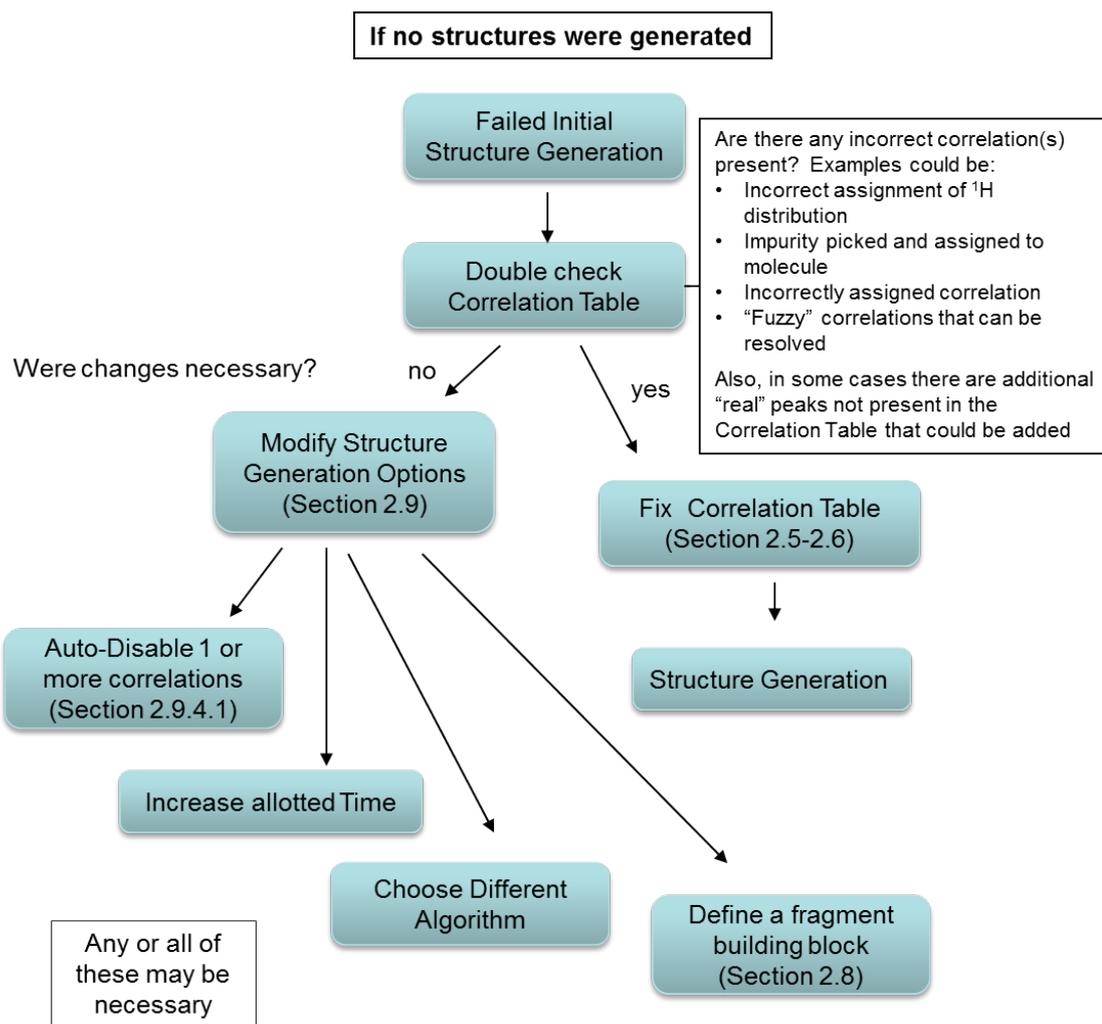


Figure 3.106: Flow Chart for the case when zero or unreasonable structures are generated

Also, because CMC-se uses the correlations to put the molecule together, if a key correlation is missing from the table, the algorithms will not be able to put the pieces together. In cases like this, adding correlations (see sections [Adding and Removing Multiplets Through the Combined Spectra Display \[▶ 57\]](#) and [The Combined Spectra Display Window \[▶ 46\]](#)) can improve the results. However when doing this, the user must be careful to have the Auto-Elimination option set (see section [Chemistry Rules \[▶ 95\]](#)), in the case that these weaker correlations that are added are in fact long range correlations.

Additionally, in the case where the structure generation ran the entire defined time without producing structures, the allowed time can be increased.

Or, in some cases, choosing a different generation algorithm might be beneficial.

Also if possible, providing a fragment with ¹³C assignments to be used as a building block can be incredibly useful.

As outlined in the figure above, in the case that too many structures are generated, the user should **look for any additional constraints which could further refine the generation process.**

This is accomplished either via adding correlations to the Correlation Table (sections [The Combined Spectra Display Window \[▶ 46\]](#) - [Additional Modifications to the Correlation Table \[▶ 71\]](#)), or through the fragment tool (section [The Fragment Tool \[▶ 81\]](#)), or by defining atom parameters such as ¹H multiplicity, or functional group (section [Adding Additional Information to the Correlation Table \[▶ 76\]](#)).

Additionally, if there are supported experiments (^{13}C , ^1H - ^1H COSY, ^1H - ^{15}N HSQC/HMBC, ^1H - ^{19}F HETCOR, ^1H - ^{31}P HMBC, ^1H - ^{13}C H2BC, ADEQUATE, INADEQUATE) that are possibly available but not already included in the project, they could be added for more constraints.

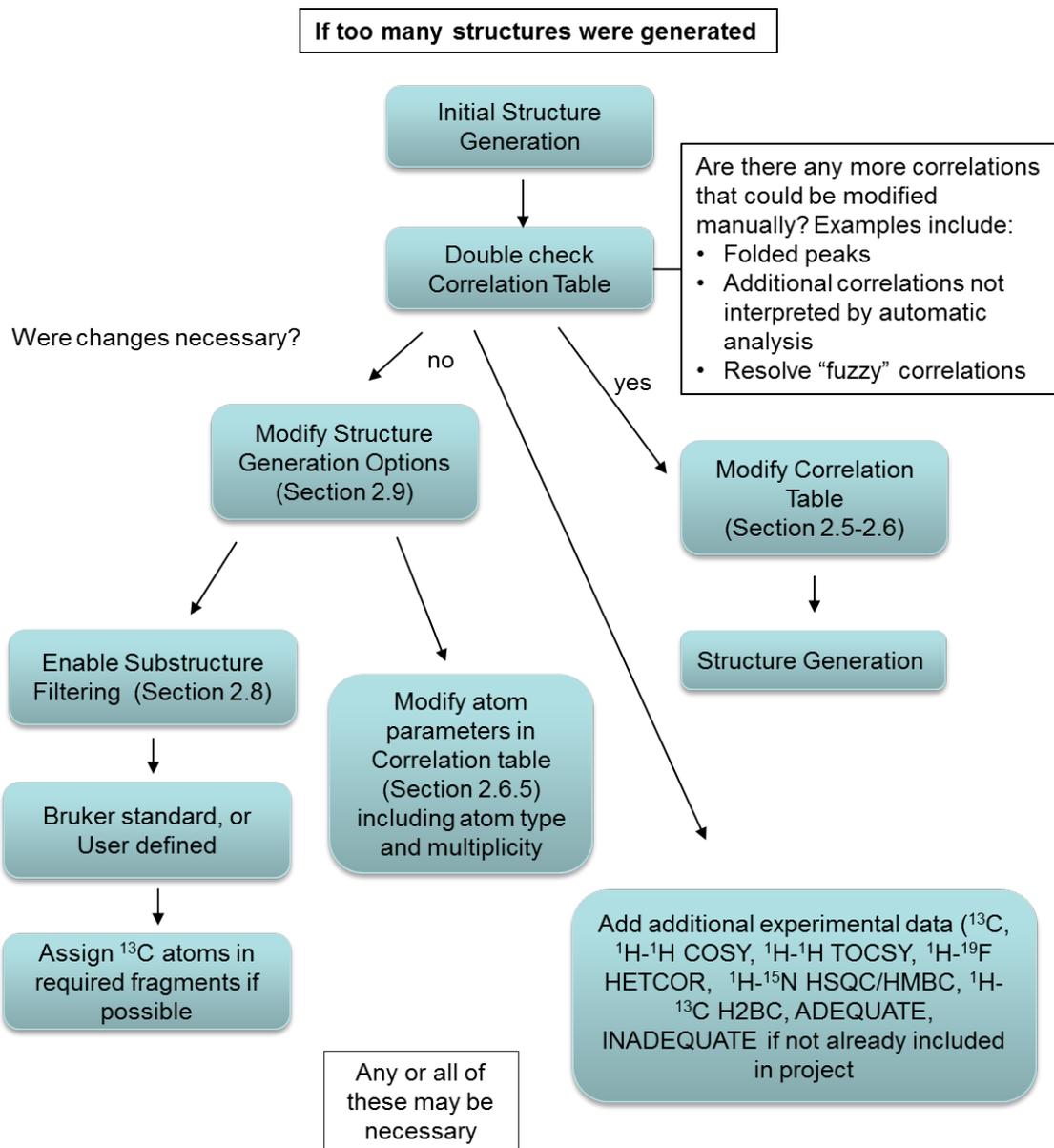


Figure 3.107: Flow Chart for the case when too many structures generated

4 The CMC-se Structure Verification Tool

The graphical tools mentioned above in section [User Validation of Generated Structure Proposals \[▶ 101\]](#) for verifying the correct structure from the pool of candidates generated by the algorithms can now be used on user generated structure candidates.

Starting in CMC-se version 2.0 was the ability to verify the plausibility of user provided candidates based on the information in the Correlation Table.

The procedure for this verification is nearly identical to the process for structure generation.

- Acquire and Process Data. Because this is still part of the CMC-se module, the same requirements for data are true for verification as they were in section [NMR Spectra for a Structure Elucidation \[▶ 15\]](#) for generation purposes.
- Create a project. You may use the automatic project generation - the **cmcse create** command described in section [Creating a New CMC-se Project \[▶ 23\]](#).
- Review Correlation Table. Rather than generating structures with the  button, the  button is used to start the verification process:

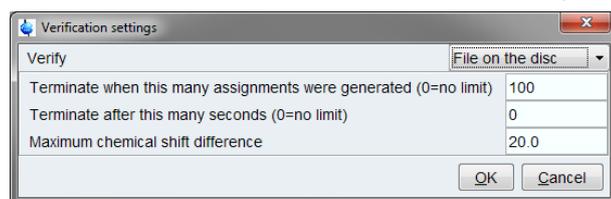


Figure 4.1: Verification Settings

The most important option in the verification settings dialog is the selection of verification mode.

- File on the Disc: Allows verification of multiple structures at one.
- Assigned Fragment: Allows definition of a partially assigned structure.

4.1 Verification of the Structure File on the Disc

The first step in the Verification Process is to select the .mol file(s) that contain the possible structures. As shown in the following figure, it is possible to select multiple possible structures at once:

The CMC-se Structure Verification Tool

The screenshot displays the CMC-se Structure Verification Tool interface. The main window shows a correlation table with columns for Name, Shift, #H, Equiv, Hybr, Func. Group, and chemical shifts (H1-H22, C1-C22). The table contains various correlation symbols (M, S, A, V, >, <, >M*, <M*, S+, S-) indicating relationships between different carbon and hydrogen environments. A dialog box titled "Select verified structure" is open in the foreground, showing a file list with four candidates: "Possible1.mol", "Possible2.mol", "Possible3.mol", and "wrong.mol". The "File name" field contains "mol" "Possible2.mol" "Possibl" and the "Files of type" is set to "*.mol, *.sdf".

Figure 4.2: Selecting multiple structure candidates for the verification to decide which, if any, are consistent with the Correlation Table.

Once the CMC-se verification process is complete, the results are shown to the user similar to as if the structures had been generated by one of the CMC-se generation algorithms. As shown in the figure above the `wrong.mol` was not consistent with the data, and thus no possible assignments could be made on this file. However, the other candidates have the possible assignments, and resulting chemical shift prediction values illustrated.

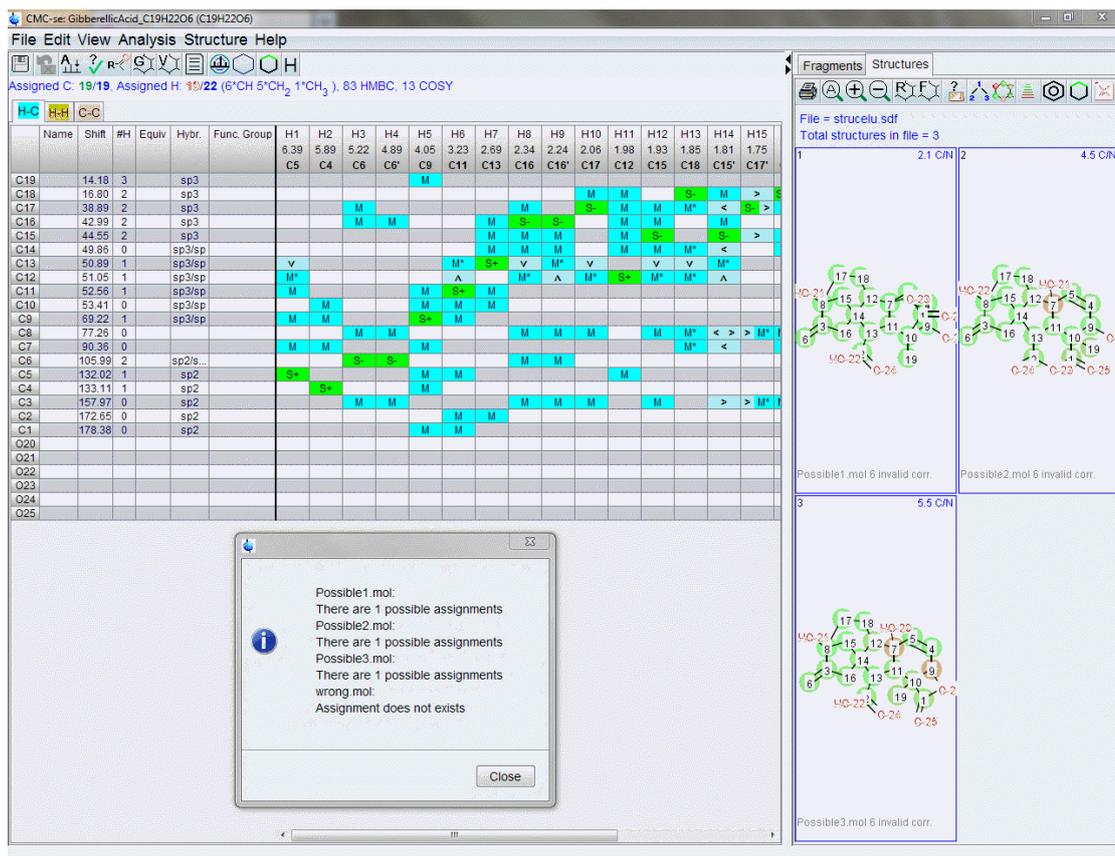


Figure 4.3: The results of the CMC-se verification routine.

The proposed structures which are at all consistent with the Correlation Table are shown, along with the possible assignments and chemical shift prediction results.

4.2 Verification of Partially Assigned Structure

You may use your knowledge and provide a partial assignment for the tested structure. The procedure is the same as using the fragment assignment for the structure generation. First, add the structure to the CMC-se Fragment tool and assign the structure as **required**:

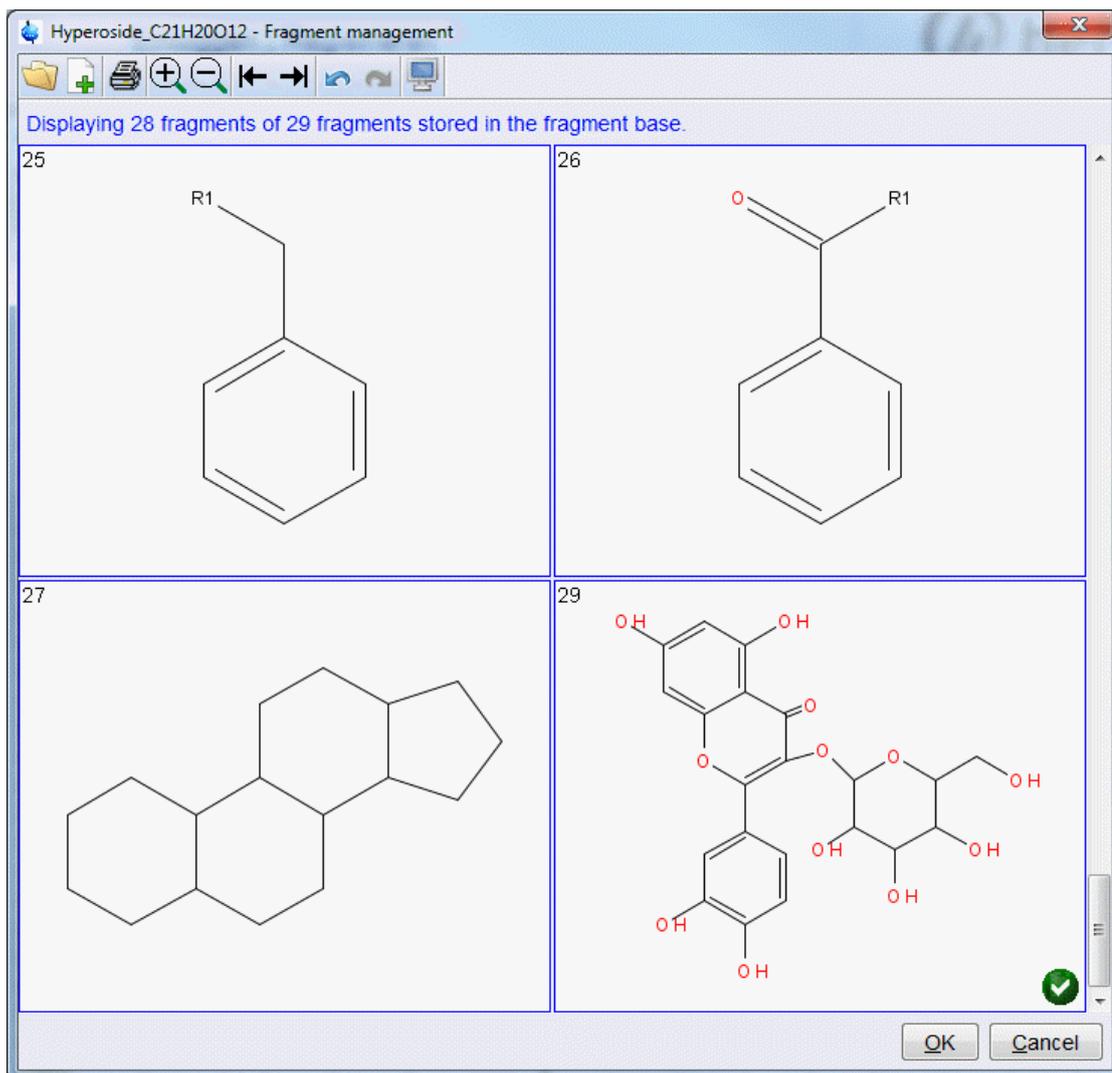


Figure 4.4: Signed structure in Fragment editor for further verification

Switch to the fragment pane in CMC-se window and provide the assignment using the interactive function described in the section [The Fragment Tool](#) [81].

Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	
						H1	C8	C11	C12	C15	C14	C13	H8	H9	H10	C19	C18	C20	
C21	37.77	1		sp3/sp															
C20	58.01	2		sp3	Frag 29										M			S	
C19	65.89	1		sp3/sp	Frag 29										M		S+		M
C18	69.23	1		sp3/sp	Frag 29										M	M		M	S+
C17	71.29	1		sp3/sp	Frag 29										M	M		M	
C16	73.86	1		sp3/sp	Frag 29										M			M	
C15	91.44	1								S+		M							
C14	96.60	1								M		S+		M					
C13	99.65	1			Frag 29					M									
C12	113.24	1		sp2	Frag 29							S+		M					
C11	113.95	1		sp2	Frag 29				M	S+									
C10	119.04	0																	
C9	120.04	1		sp2	Frag 29				S+		M								
C8	131.41	0		sp2															
C7	142.69	0		sp2															
C6	145.37	0		sp2															
C5	154.15	0		sp2															
C4	159.16	0		sp2															
C3	162.04	0		sp2															
C2	219.60	2		sp2															
C1	225.70	0		sp2															
O22																			
O23																			

Figure 4.5: Fragment pane in CMC-se window

Now you may start the verification. Select the **Assigned fragment** as the verification mode

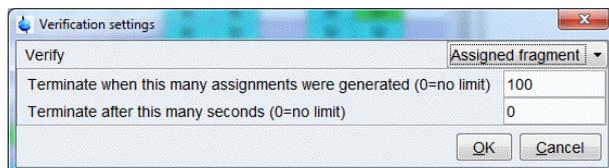


Figure 4.6: Verification settings

The verification starts and shows only assigned structures matching your starting assignments.



Note: This is the way to run a chemical shift prediction routine on a molecule that has been manually generated or assigned so that the user can start to evaluate the correctness of their results.

4.3 Verification Report

A verification report may be generated in the same way as the elucidation report.

Right clicking on a structure will display the pop-up menu:



Figure 4.7: Pop-up window for Structure Options

- The **Save** option selected offers multiple subsequent options for reporting the final structure possibilities.
 - **Selected Structure To Data Set** is a “final save” type option to be used with the final structure has been determined/verified. This option saves this particular structure to a `.mol` file, and a `.pdf` document summarizing the structure and assigned atoms and correlations to the defined `CMCse-result.mol` and `CMCse-result.pdf` files in the same directory path as project directory (default name `strucelu`). This routine also automatically fills in the “Name” column of the Correlation Table with the Canonical Atom Number. The PDF report (as shown the figure below) will include these numbers.
 - **Structure and Report** option allows the user to select individual structure(s) and save the same `.mol` and `.pdf` file as above, but with a user defined name and location. This option is useful when more than one structural candidate is under consideration and reports for multiple structures are desired.

The CMC-se Structure Verification Tool

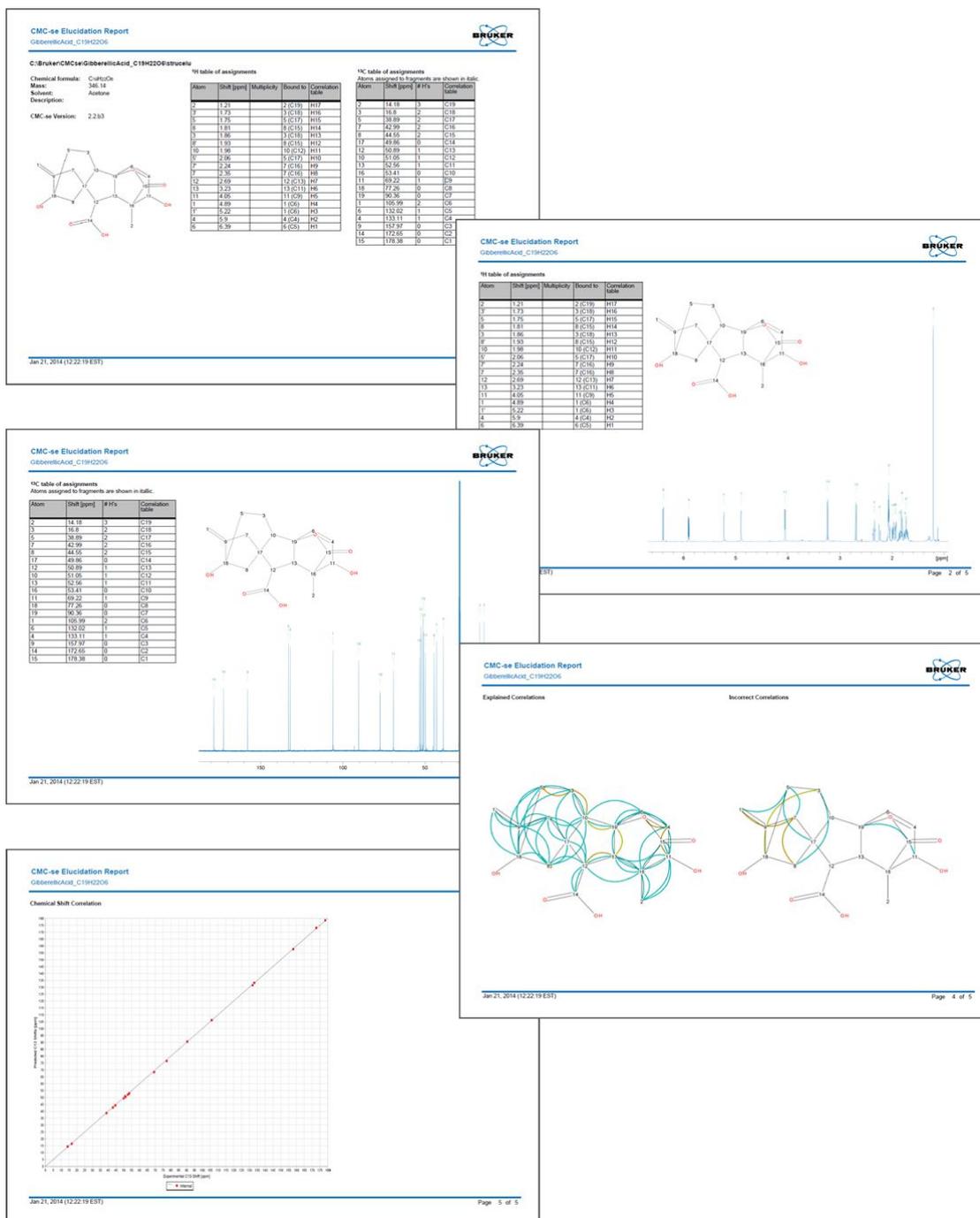


Figure 4.8: The 5 pages from the report generated from the above “save” options.

The report in the figure above includes:

- Tables of the chemical shifts and assignments.
- The 1D ¹H and ¹³C spectra.
- A summary of the correlations both valid and invalid.
- A graph of how well the experimental results match the predicted ¹³C chemical shifts.

5 Examples

Now that all features of CMC-se have been introduced, the best way to get familiar with them is to work with examples. This chapter will present three different molecules which each vary in complexity, and represent the major features of the program discussed in Chapter 2 [Structure Elucidation Basics \[15\]](#).



Going through **ALL** of these examples is **HIGHLY** recommended to give the user an idea of how to work with this program and of the recommended approach to solve the variety of problems that will arise when trying to elucidate unknown structures.

The results shown below are from the current version of the software at the time this manual was written. Minor deviations from these results might be apparent if a newer version of the software is being used. Also exact times and number of structures generated in a given time will vary depending on speed of computer being used. However, the general process is the same, and these still serve as examples for ways of manipulating the software to produce quality results.

All of the data sets used in the following examples are part of the “Example Data” within the TopSpin 3.5 installation. Please refer to the [Program Prerequisites \[11\]](#) for the custom install options. It is necessary to make sure the example data sets are installed along with TopSpin.

5.1 Structure Elucidation for α -Ionone

The first example is that of α -Ionone. Its molecular formula is $C_{13}H_{20}O$, and the sample was in $CDCl_3$.

This is a relatively small compound, with well dispersed NMR spectra. Thus it represents a simple structure elucidation problem which is useful for users first starting to use the CMC-se program.

It also can easily be used as an example molecule for illustrating the power of the new Fragments tool for manually elucidating or assigning a structure.

The following spectra have been acquired (**exam_CMCse_1**) and processed properly (see section [Data Set Processing \[19\]](#)):

- PROTON [1H]
- COSY [1H , 1H]
- HSQC [1H , ^{13}C]
- HMBC [1H , ^{13}C]
- C13 [^{13}C]

Create a new CMC-se project as described in section [Structure Elucidation Project Management \[21\]](#), using all available spectra.

Performing the Automatic Analysis () using all spectra, results in the Correlation Table as shown below:

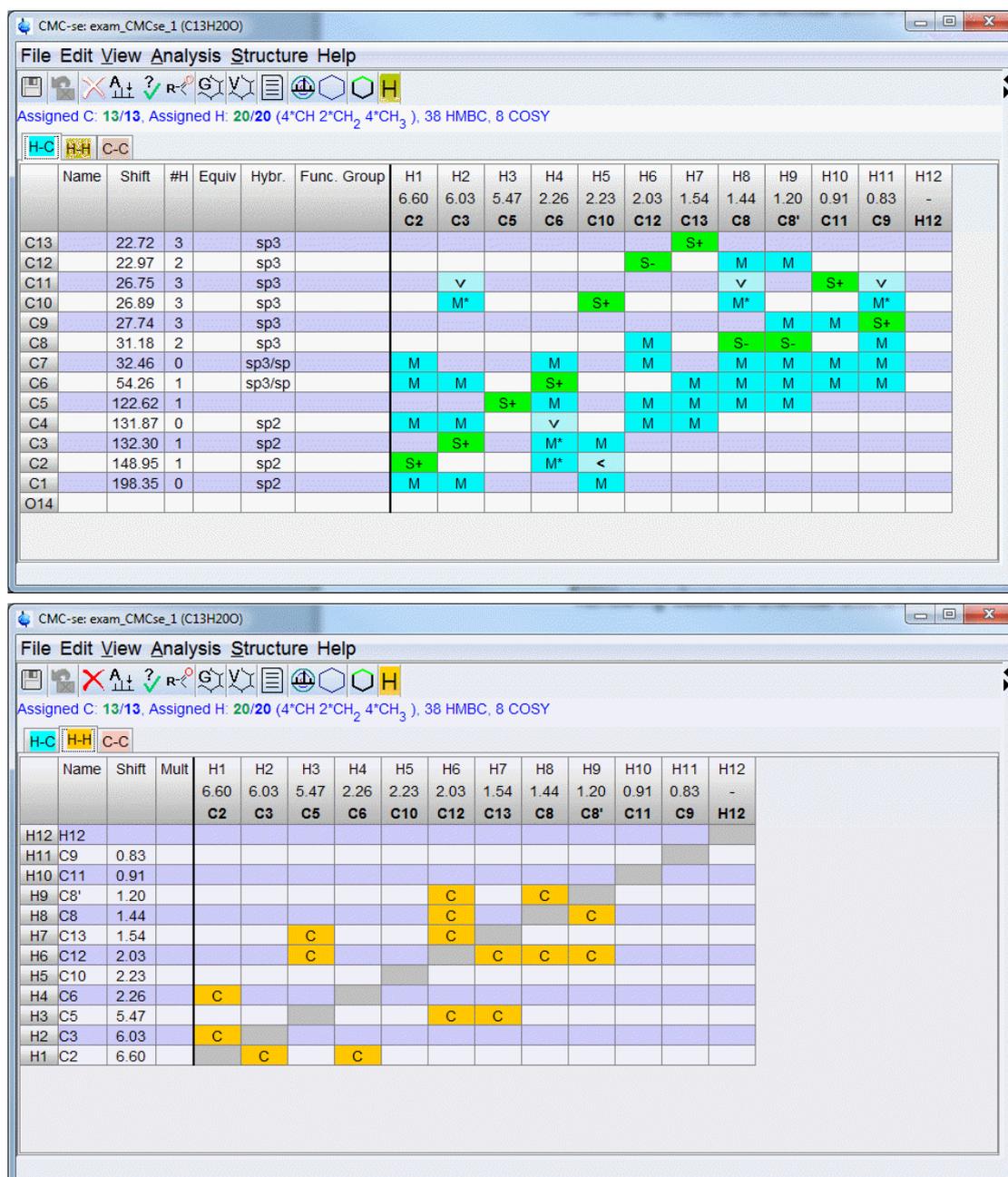


Figure 5.1: Correlation Table after automatic analysis. The top panel is the H-C tab, and the bottom is the H-H tab.

As described in the section [The Correlation Table](#) [33], the Correlation Table has multiple tabs. The first (top) displays the ^1H - ^{13}C connectivity information, including the ^1H and ^{13}C chemical shifts, HSQC, and HMBC correlations.

The second tab (bottom), displays the ^1H - ^1H connectivity information from the analysis of the COSY spectrum. As shown here, the ^1H atoms have been re-named according to the attached carbon atom. This was an option in the project creation **Derive proton names from skeleton atom** which was checked, so the protons have been renamed so the H-C pair can be treated as a single entity rather than multiple atoms. The original ^1H numbering based on chemical shift is also included as a reference if desired.

In addition to the Correlation Table being automatically populated by the analysis, the picked peaks are also indicated on the combined spectra display:

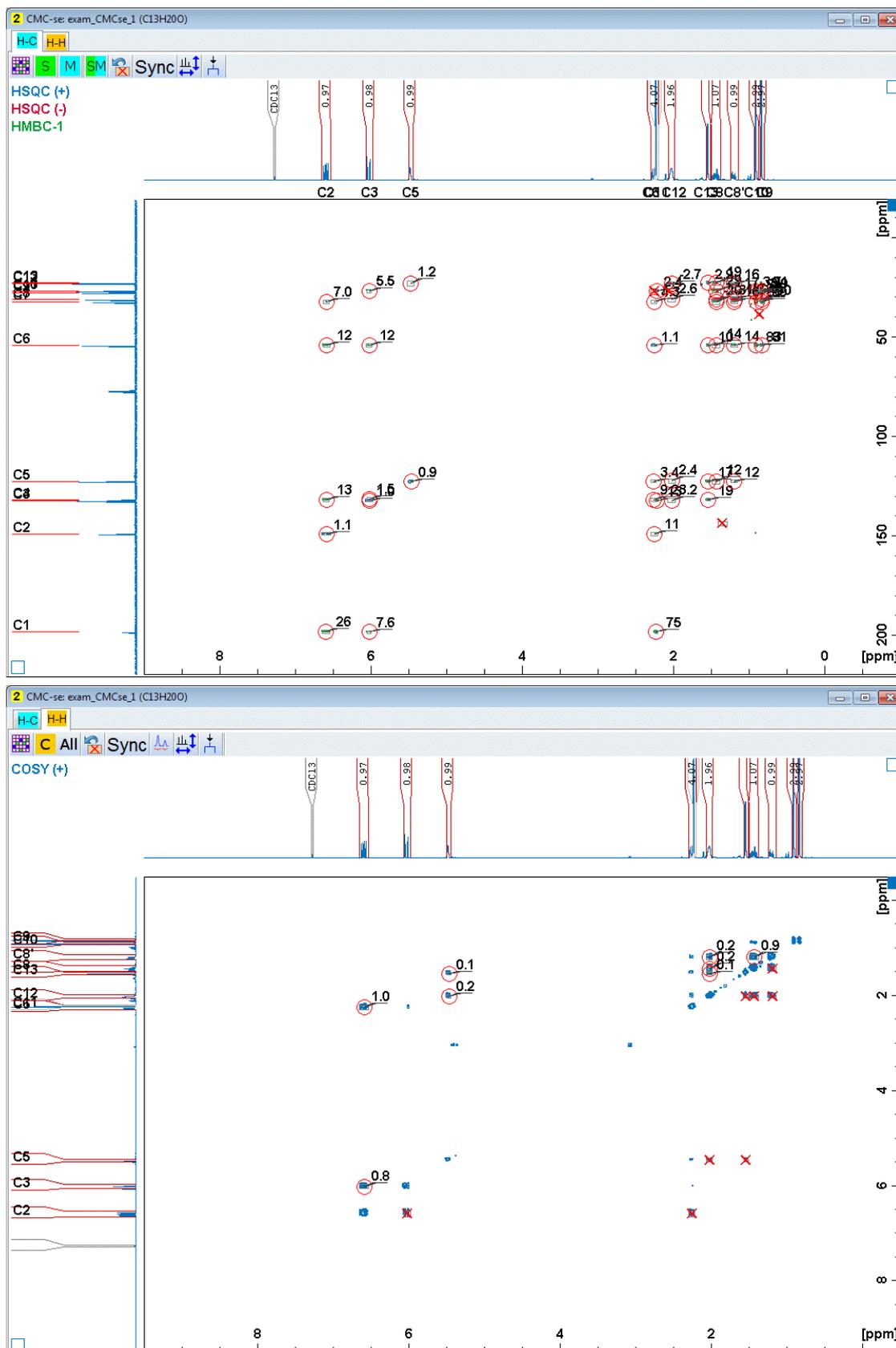


Figure 5.2: The Combined Spectra Display is updated with the peaks that are picked during the automatic analysis.

In the figure above, the first tab (top) is an overlay of the H-C experiments, and the second tab (bottom) is the H-H experiments which are included with the project.

Because this is a simple project with a small molecule and relatively few spectra types, it might be advantageous to switch the viewing layout back to the *old* style that does not have the tabs. As described in above in section [The Combined Spectra Display Window \[46\]](#), this is accomplished through the **Toggle Layout** option in the **View** menu of the Correlation Table. As shown in the following figure, it is also possible to us the **Arrange Frames** option to manipulate both windows to be shown on the same screen.

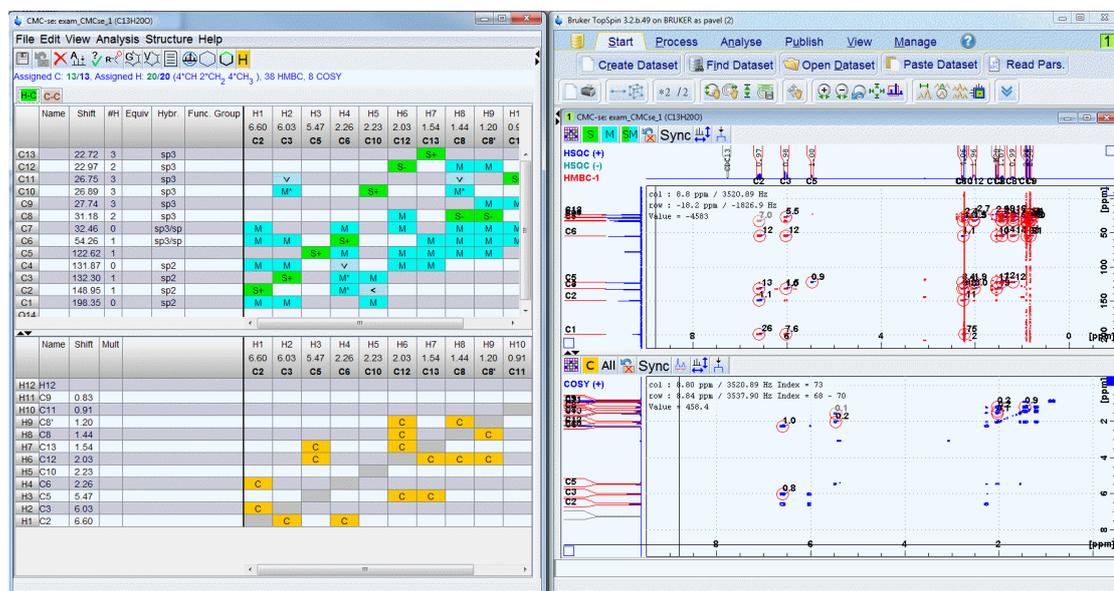


Figure 5.3: An alternative view option which when working with simple spectra like this example allows all data to be examined at the same time (Arrange Frames).

After the automatic analysis is complete, the user should quickly evaluate the results. As seen in the top of the Correlation Table above:

- The DBE is a whole number so it is not likely necessary to deal with charged structures.
- All 13 carbons and 20 protons were assignable by the software.
- The determined proton distribution was 4 CH groups, 2 CH₂ groups, and 4 CH₃ groups.
- It is also shown that 38 HMBC correlations and 8 COSY correlations were assignable by the analysis.

Clicking on the automatic consistency check button () results in the following pop-up window:

• CMC-se Consistency Report

Project path = C:\Bruker\CMCse\exam_CMCse_1\strucelu

Chemical formula : C13H20O

Mass : 192.297

DBE : 4

H/CNO= 1.4

Inconsistency found

Some atoms have more bonds than possible immediate neighbours

Atom Max bonds

C12: 2 COSY[3,6] COSY[6,9] COSY[6,8] COSY[6,7]

C13: 1 COSY[3,7] COSY[6,7]

Some of these correlations are either false or long range

Use "Autoeliminate" switch in structure generator options

Consistency check of proton distribution and 1D integrals

Interval [ppm]	nH	1D integral	HSQC volume (Both normalized to # of protons)
[6.651, 6.533] :	1	0.969	1.114 C2
[6.065, 5.977] :	1	0.985	1.029 C3
[5.499, 5.440] :	1	0.989	0.856 C5
[2.296, 2.199] :	4	4.070	3.516 C6+C11+
[2.062, 1.984] :	2	1.960	2.732 C12
[1.564, 1.515] :	3	2.990	2.904 C13 Ignored, contains potential solvent peak
[1.496, 1.379] :	1	1.072	1.301 C8
[1.242, 1.144] :	1	0.992	1.312 C8
[0.930, 0.881] :	3	2.988	3.705 C10
[0.852, 0.803] :	3	2.975	3.926 C9

Number of explained H: 17

Ratio 1D_Integral/nH 16.985 +/- 0.514

Correlation coefficient 0.9994

Nov 13, 2012 (12:01:35 PM)

Page 1 of 2

Figure 5.4: CMC-se Consistency Report.

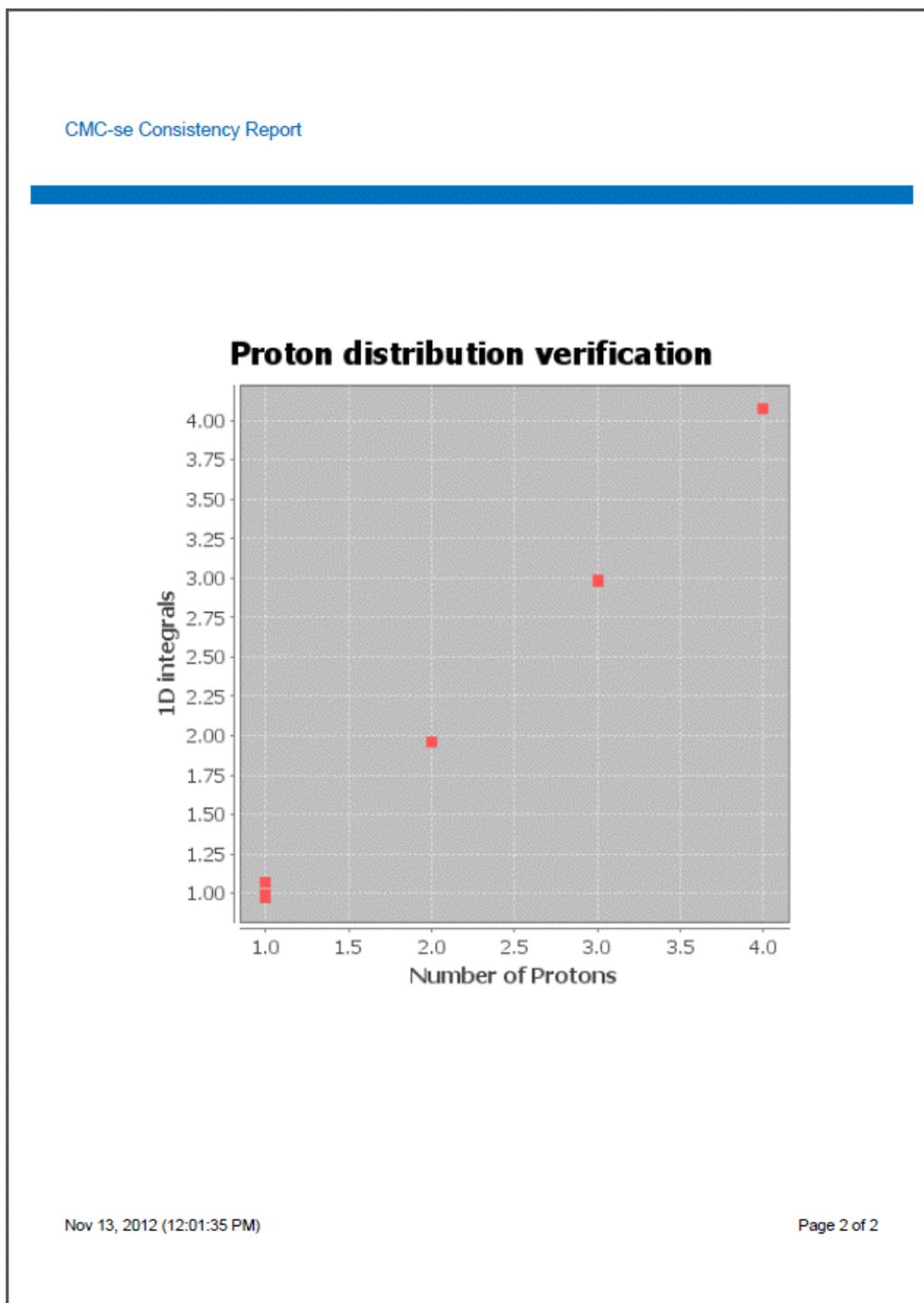


Figure 5.5: Results of Consistency Check

The results of the consistency check (section [Typical Problems with Automatic Multiplet Analysis](#) [42]) are encouraging.

The second section below the general information containing information regarding the chemical formula, *Inconsistency Found* is a section that indicates potential inconsistencies were found. In this case it is that some atoms have more bonds than possible immediate

neighbors. This is an indication that there are long-range correlations present in the data, and as the message suggest, the **Auto eliminate COSY** option should be checked in the structure generation options (section [Setup of Structure Generation Options \[88\]](#)).

The lower section of the table is an indication that the proton distribution was properly assigned by the software. The table is a comparison of the integral volumes from the 1D-¹H spectrum and the HSQC spectrum. After the table, on the second page there is a graphical representation of this information. The correlation coefficient of this comparison for this data is 0.9994; and is high enough to indicate there was not a problem assigning the number of protons responsible for each signal. Also indicated in the text was that 17 explained protons were in this calculation, because as indicated in the table, the software recognized that the Methyl group between 1.564 and 1.515 ppm was too close to the resonance frequency of water in CDCl₃, so that integral region was ignored during the calculation to improve the accuracy.

This consistency check and a quick inspection of the Combined Spectra Display window do not reveal any glaringly obvious problems, thus the recommended action in this case to proceed with the structure generation process.

Clicking on the  button brings up the default Structure Generation Options as shown below:

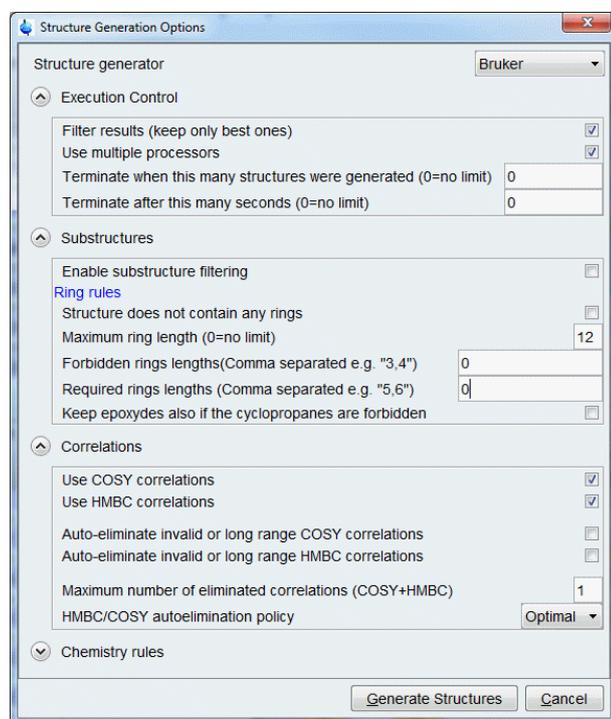


Figure 5.6: Default Structure Generation Options

Using these default options the error message shown below is the result:

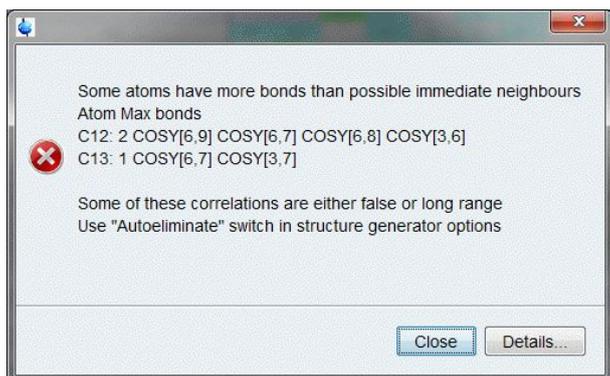


Figure 5.7: Error message resulting from structure generation with default options

This error message is the same information that was contained in the Consistency Check above.

As illustrated in the next figure, the COSY correlations in question are all reasonably strong signals, and nothing is obviously miss-picked by the software:

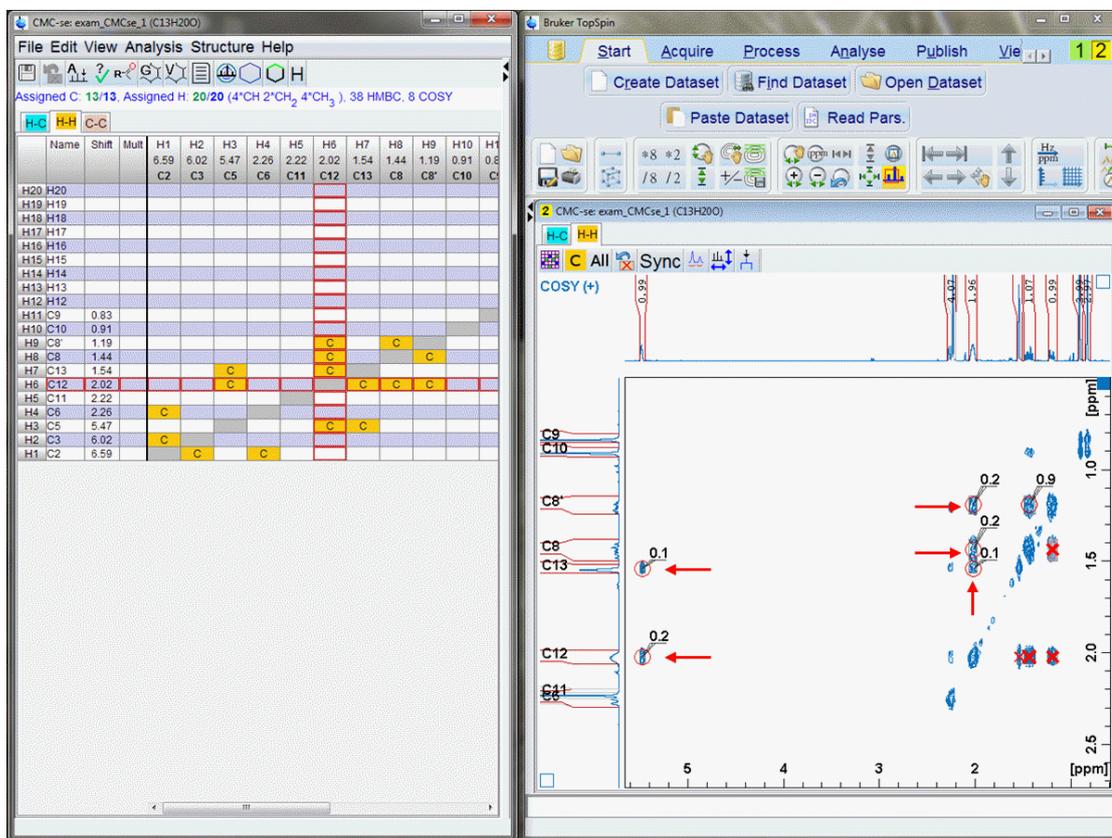


Figure 5.8: Checking the quality of the "Inconsistent" COSY correlations

The protons that correspond to carbons 12 and 13 are those at 2.02 ppm and 1.54 ppm respectively (C12/H6 is highlighted in the Correlation Table) and are the signals to be examined.

- There are 4 COSY cells filled in with correlations for the proton at 2.02 (C12) at 5.47 ppm, 1.54 ppm, 1.44 ppm and 1.19 ppm.
- There are 2 COSY cells filled in with correlations for the proton at 1.54 ppm (C13), at 5.47 ppm, and 2.02 ppm.

In the figure above, the peaks in the Combined Spectra Display for these particular correlations are highlighted by red arrows. All of these signals are of reasonable intensity and reasonably picked, so it is not possible to manually define any of these 5 correlations as incorrectly picked/assigned or long range.

Thus, following the software's suggestion of utilizing the *Auto Elimination* option is clearly the best choice.

To use the *Auto Elimination* option (see [Correlations \[92\]](#) for details), it is necessary to go back to the structure generation options, and select BOTH the **Auto-eliminate invalid or long range COSY correlations** and **Auto-eliminate invalid or long range HMBC correlations** options.

Because there are at least 2 atoms that were identified as inconsistent, the initial starting point for the number of correlations to be eliminated should at least be 2.

When 3 correlations are auto eliminated, a single structure is generated. As suggested in the section [Process Overview \[120\]](#), even though a structure was generated with these options, it is advisable to increase the flexibility by at least one correlation to see if additional (and possibly better) structures are generated. For illustration purposes in this example, the number of eliminated correlations was set to 5 so that more than one structure will be generated and can be evaluated.

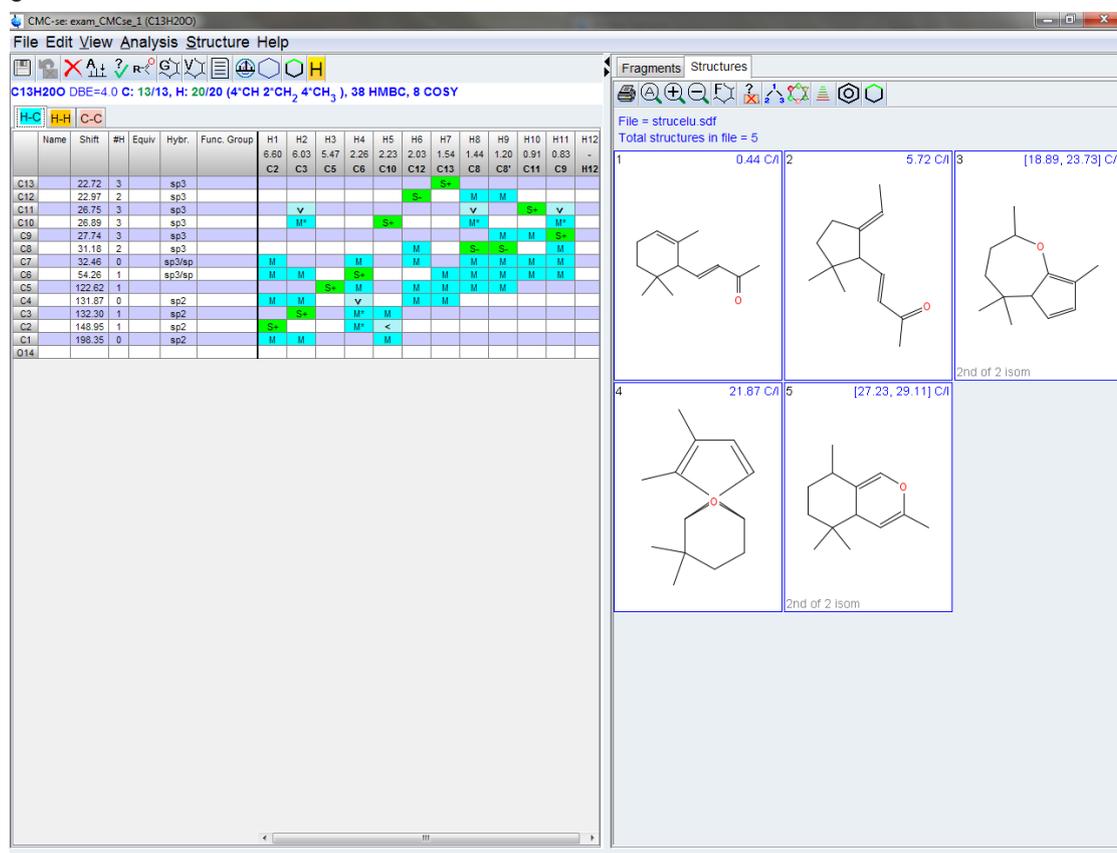


Figure 5.9: Structures generated when the auto-elimination options are activated.

Looking closer at the structures and their ^{13}C chemical shift prediction results in the following figure, it is obvious that the first 2 structures are clearly better candidates than any of the others.

As explained in section [Structure Rating via Chemical Shift Prediction \[107\]](#), a prediction standard deviation value less than 5 is typically a valid possibility, and usually those structures with a deviation above 10 can be disregarded. Thus, attention will focus on structure #1 and #2 for the remainder of the example, because the deviation of 10 for candidate #2 cannot be discounted without good reason.

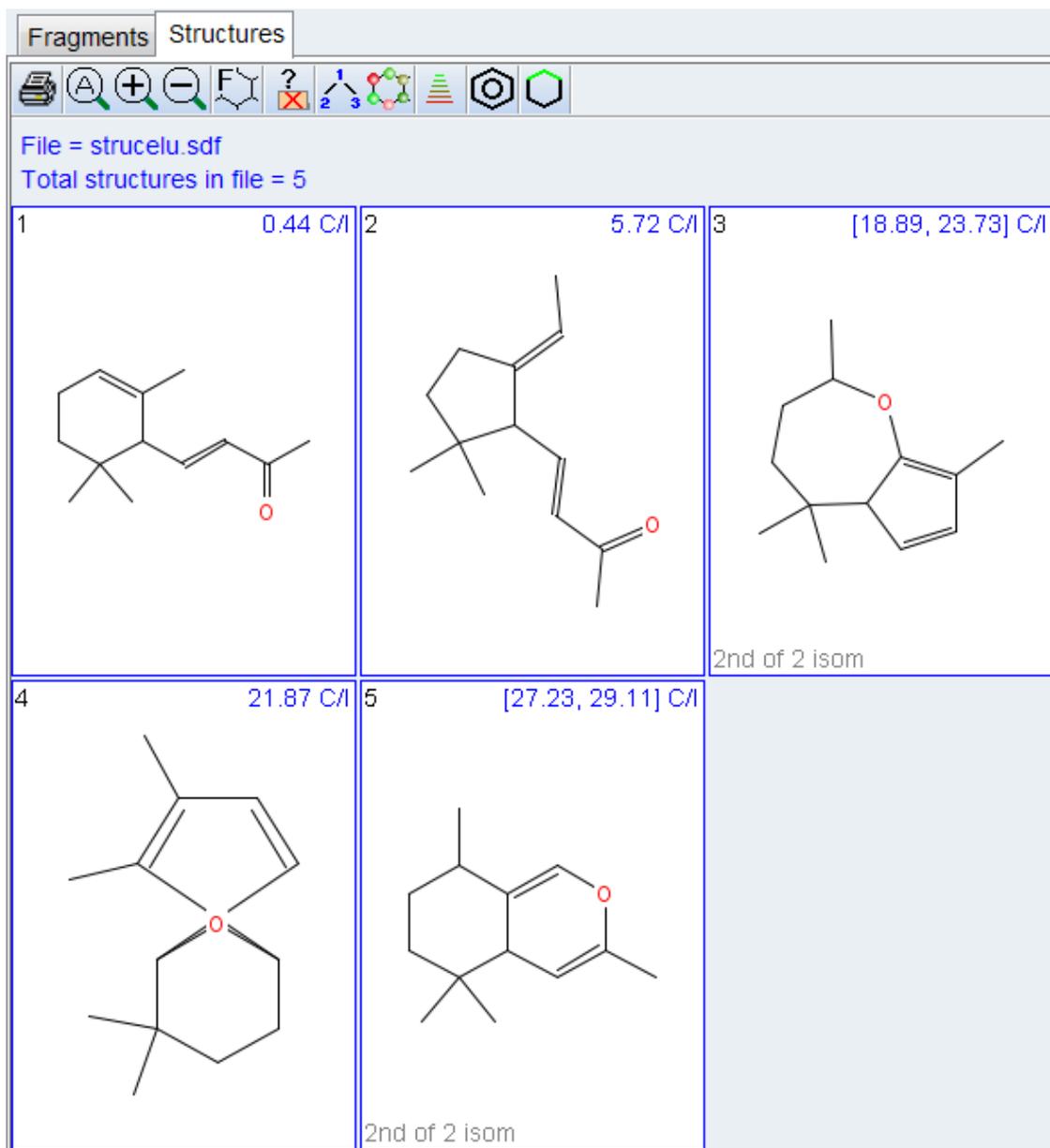


Figure 5.10: Closer look at the Generated and Ranked Structures

As described in section [Structure Rating via Chemical Shift Prediction](#) [107], these chemical shift prediction results can also be viewed graphically on the structures with the  button, as shown in the figure below. This gives further confidence that structures #3-6 can be removed from consideration because there are many atoms with poor matches to the predicted values (red circles), and for similar reasons, that it is likely that structure #2 is not correct because there are multiple atoms with less than ideal matches (brown circles).

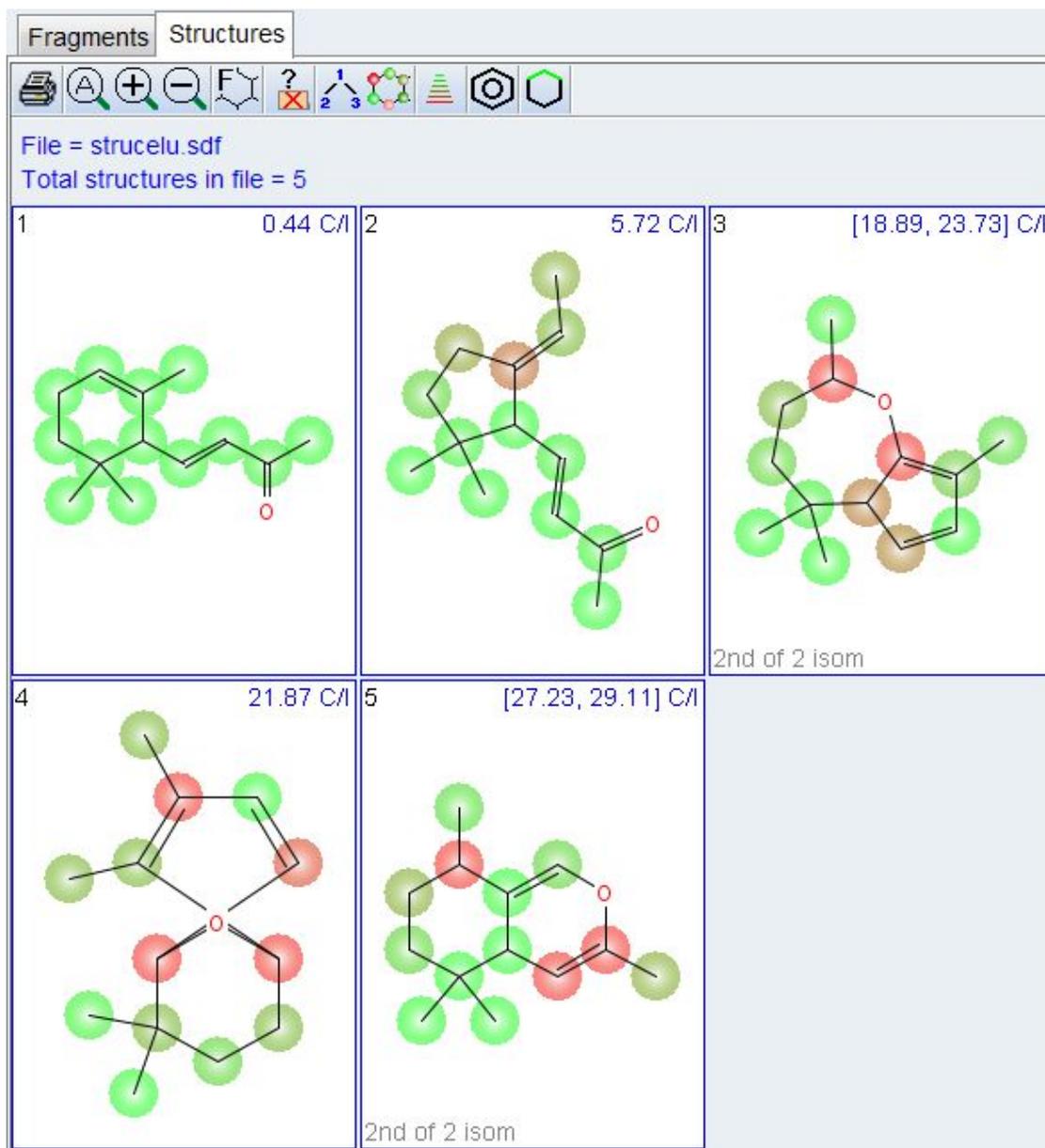


Figure 5.11: Graphical representation of the chemical shift prediction results

As described in [Viewing Generated Structures \[101\]](#), the  buttons can be used to adjust the number of structures displayed in the window.

It may be recalled, that the **Auto Elimination** option had to be selected in order to generate this structure. So a good second step in critically evaluating whether this structure is correct or not is to display which of the correlations the software eliminated during the generation process (section [Viewing Correlations on Structures \[110\]](#)), and the results of this are shown in the figure below. Because structure #1 was generated with only eliminating 2 correlations and #2 was generated when the flexibility was increased to eliminate 4 correlations, it would be a good starting point to see which additional correlations were eliminated with this one.

Clicking on this structure and then the  button will highlight those correlations which were auto eliminated for this particular structure on all structures, and also highlight the corresponding cells in the Correlation Table.

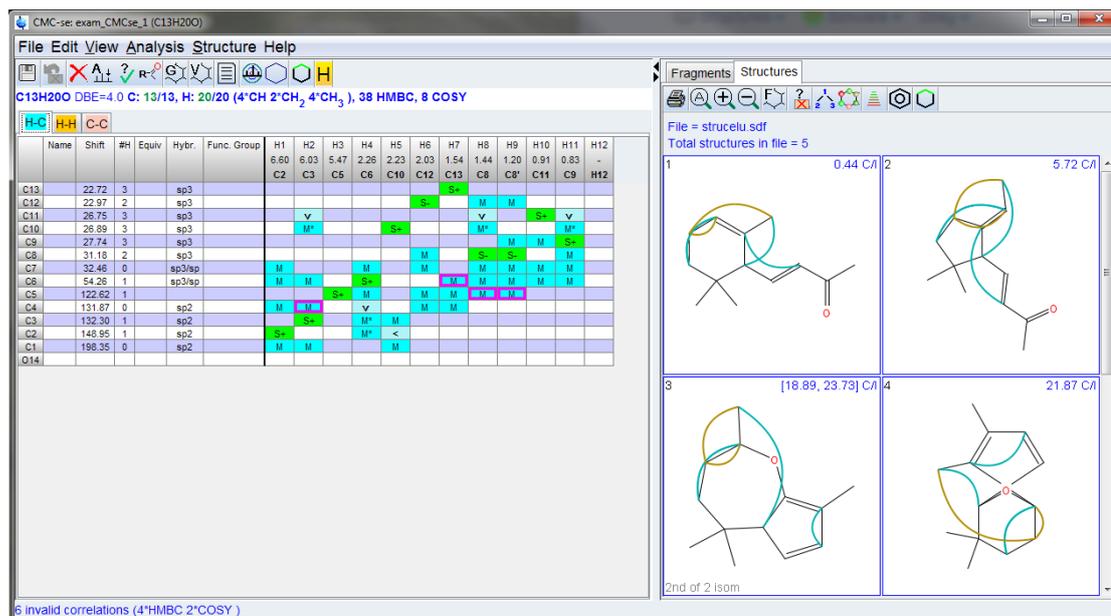


Figure 5.12: Top ranked structures highlighting the correlations that were automatically eliminated for structure #2 on all structures and in the table.

As shown in the figure above, the eliminated correlations are highlighted by arrows on the structure, and with pink frames around the correlations in the Correlation Table. The HMBC correlations which were invalid for structure #2 are valid in structure #1; giving increased confidence that the toped ranked structure is the correct result.

As described in [Viewing Generated Structures \[101\]](#), now that the user has evaluated the structure possibilities, and determined which structure is the correct structure, a report can be generated by right clicking on the desired structure, selecting the **save** option, and then the subsequent **Structure and report** option.



Remember that this is only the atom connectivity of the structure, and further analysis needs to be done to determine the geometry. While this example is drawn in the trans conformation around the sp² carbons, this is only one possible orientation, no data has been analyzed to support or refute this conformation.

This option will produce a .pdf file that contains:

- The structure which has automatically been numbered according to the canonical numbering scheme
- Assigned 1D ¹H, and ¹³C spectra
- Tables with the chemical shift values.
- A summary of all (both the valid and invalid) correlations used in the generation
- A graphical representation of how well the experimental ¹³C chemical shifts matched the predicted values.

α-Ionon is a small molecule with well separated NMR signals. So it is a good example for a relatively simple structure elucidation with only a few problems which can be directly solved. The program did not have any problems analyzing the spectra and only a few COSY correlations were problematic for the generation.

This simple structure elucidation is a good example for someone first working with the program, but other structures are likely to cause more problems. Such problems are examined in the other examples.

5.1.1 Manual Structure Elucidation of α -lonon using the Fragment Tool

Because α -lonon is a simple molecule, it can also serve as an example on how CMC-se can be used as a **Structure Workbook**, for manually working with the 2-dimensional data of small molecules for structure elucidation or assignment purposes.

Just as above, the data have been acquired, processed, and a CMC-se project has been created. After the automatic analysis, the Correlation Table is filled in as shown in the following figure. Notice how the divider has been moved to the left, so the Fragments window is also shown. This will be the primary working space for the manual elucidation of this structure.

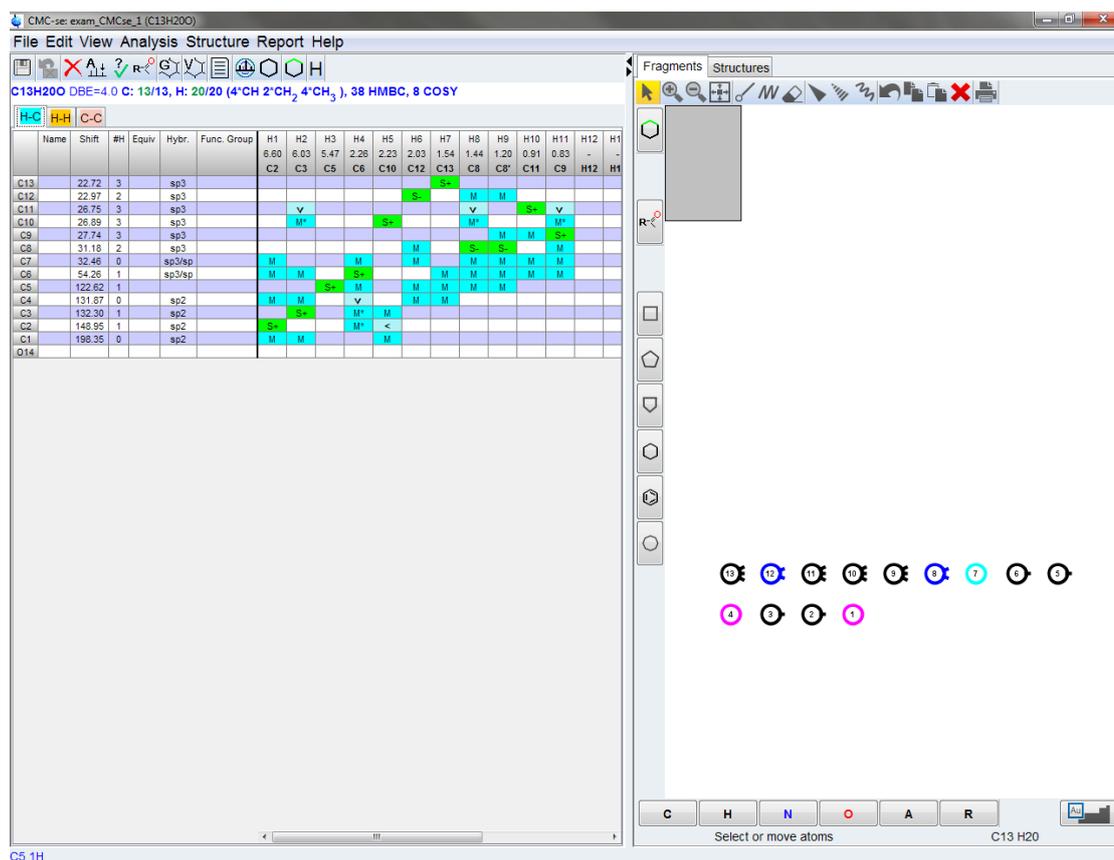


Figure 5.13: Correlation Table with the Fragments Window Open for Manual Elucidation

Prior to starting a manual elucidation, it is highly recommended to check the Correlation Table with the data for consistency and to see if it can be simplified any. A quick examination indicates that all of the required atoms were found (13 of 13 Carbons, and 20 of 20 protons) and there don't appear to be any issues with peak picking. One particular area that should be examined are the "Fuzzy Correlations". The computer algorithms are very efficient at trying multiple possibilities for a single correlation, but when the user is attempting this manually, it can greatly help the process if these possibilities are reduced.

In this case, there are multiple "Fuzzy correlations" that can be removed. For example the correlation shown in the figure below was assigned in the table as belonging to H2(C3) and either C10, or C11. Clicking on the **M*** cell in the table will highlight that particular correlation in the Combined Spectra Display as shown in the figure below.

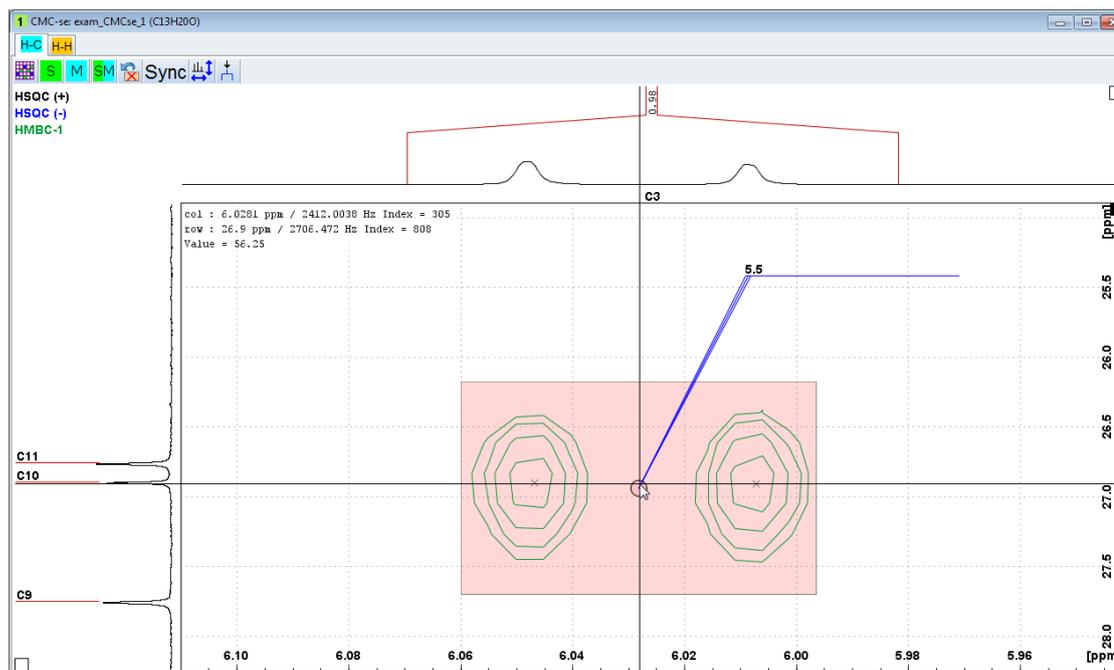


Figure 5.14: Using the Combined Spectra Display to clarify Fuzzy Correlations when Possible

Using the view as in the above figure, it is clear that the cursor when hovered over the center of the 2D peak is in better alignment with C10 than C11. Therefore, it is possible for the user to safely say that this correlation belongs to C10, and the C11 can be removed either by deleting the cell from the correlation table, or right clicking, on the peak in the spectrum and modifying the properties box to not include the C11 as shown in the figure below.

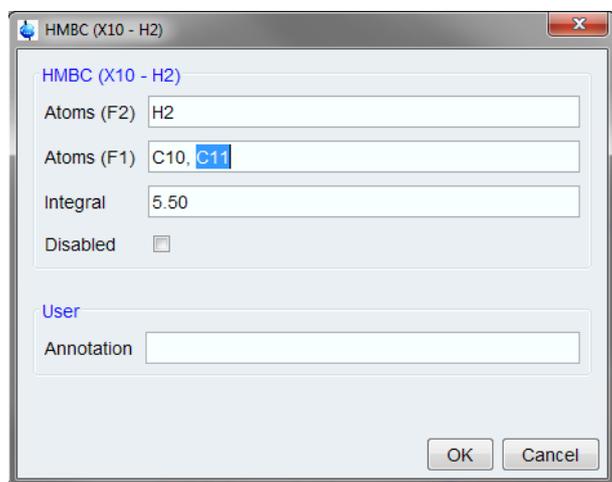


Figure 5.15: Modifying the properties of a correlation to remove the alternate atom assignments

This similar process can be repeated for the remaining “Fuzzy correlations” in this example data set.

Now that the correlation table is set, the user can start to define fragments and begin to put them together based on the experimental data.

One easy starting point in this molecule is the C1 carbon. Based on chemical shift, it is known that this carbon has to be a Ketone. This can be defined in the **Atom Properties** dialogue box that can be opened by double clicking on the **C1** cell in the correlation table. This is shown in the figure below.

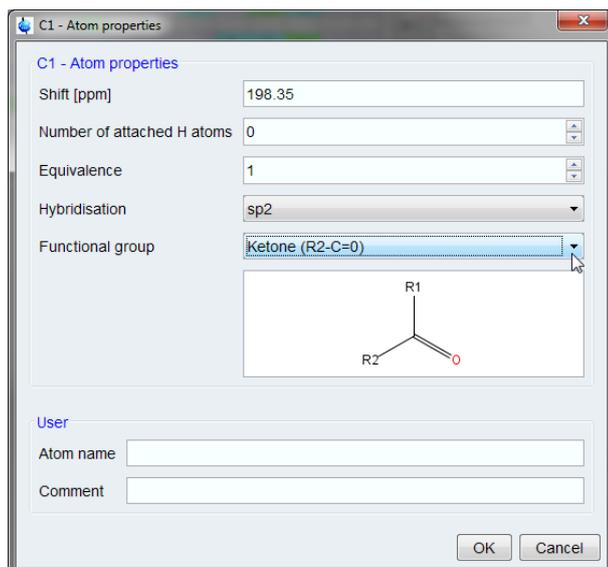


Figure 5.16: Defining C1 as a ketone through the Atom Properties dialogue box.



Only those functional groups which are consistent with the chemical shift of an atom are shown in the dialogue box. Therefore, this can be used as a guide to help the novice user determine possible functional group identification when necessary.

Once a fragment is defined, it will automatically appear in the Fragments window with the corresponding atom assigned accordingly. This is shown for C1 in the figure below.

Examples

CMC-se: exam_CMCse.1 (C13H20O)

File Edit View Analysis Structure Report Help

C13H20O DBE=4 0 C: 13/13, H: 20/20 (4*CH 2*CH₂ 4*CH₃), 38 HMBC, 8 COSY

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13
							C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
C13		22.72	3		sp3														
C12		22.97	2		sp3														
C11		26.75	3		sp3														
C10		26.89	3		sp3														
C9		27.74	3		sp3														
C8		31.18	2		sp3														
C7		32.46	0		sp3/sp														
C6		54.26	1		sp3/sp														
C5		122.62	1		sp2														
C4		131.87	0		sp2														
C3		132.30	1		sp2														
C2		148.95	1		sp2														
C1		198.35	0		sp2	R2-C=O													
O14																			

C4 sp2

Fragments Structures

13 12 11 10 9 8 7 6 5
4 3 2

C H N O A R

Select or move atoms C13 H20 O

Figure 5.17: Fragments window automatically updated with the newly defined fragment

Clicking on any atom in the Fragments window will highlight the experimental constraints to that atom. This is shown for the C1 carbon in the figure below. One of the HMBC correlations is to C10, a CH₃ group.

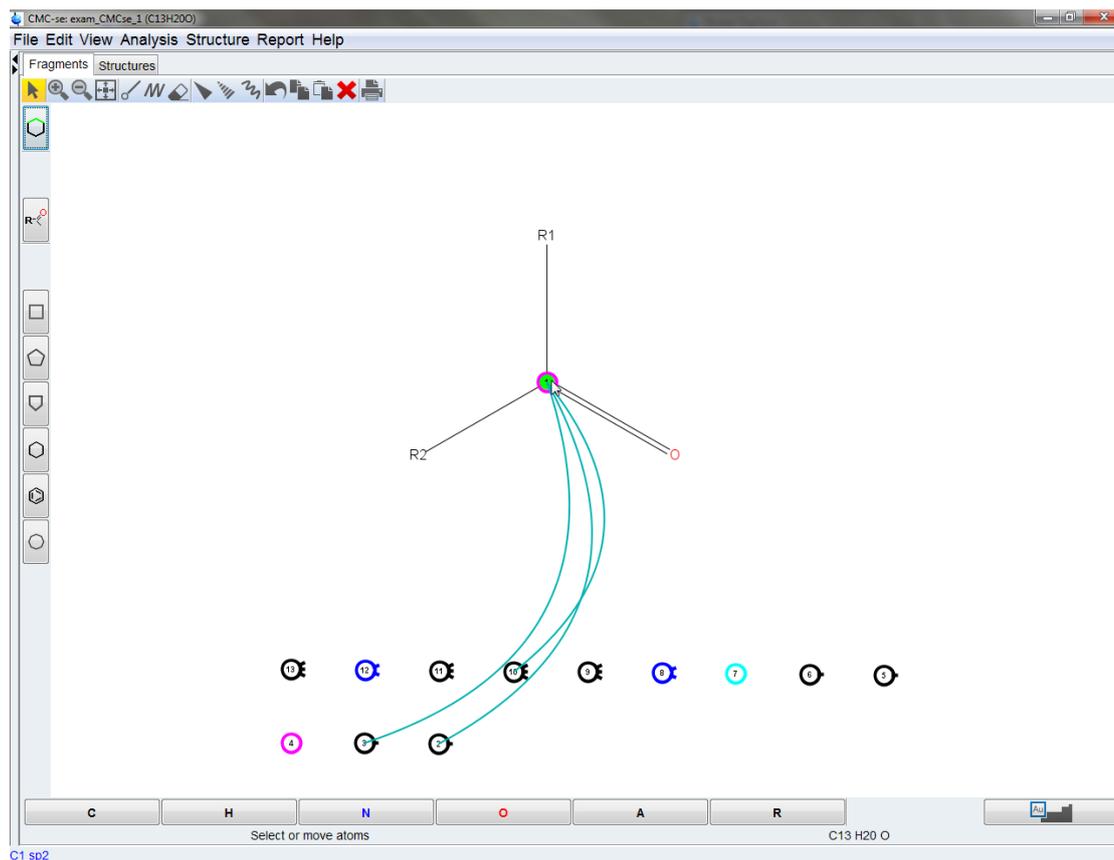


Figure 5.18: Clicking on an atom will highlight the experimental constraints for that atom

With the somewhat downfield chemical shift of the ^1H and ^{13}C of this CH_3 , it is likely one of the R groups bonded to the Carbonyl of C1. This C10 ball can then be assigned to the fragment by clicking and dragging it over one of the R groups as shown in the figure below.

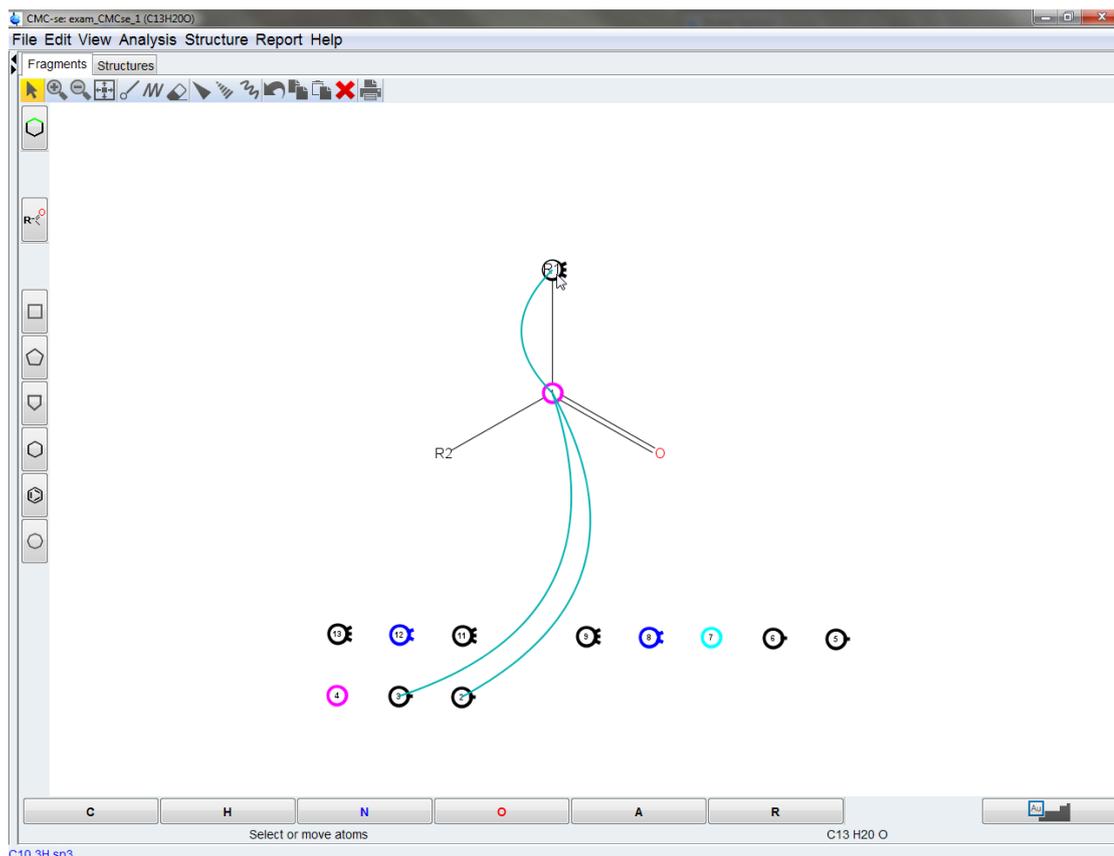


Figure 5.19: Clicking and dragging another atom to an existing functional group to make additional assignments.

The other HMBC correlations to C1 are those to C2 and C3. From a quick inspection of the data as shown in the figure below, it can be determined that both C2 and C3 are sp² hybridized carbons (based on chemical shift) and are connected to each other (based on the COSY correlation). So this is a -C=C- group that is connected to C1.

Looking at the ¹H coupling patterns of the two protons, the proton for the C1 CH, is a doublet, where the ¹H for C2 is a doublet of doublets. Therefore, C3 has to be the carbon directly next to the C1 carbonyl.

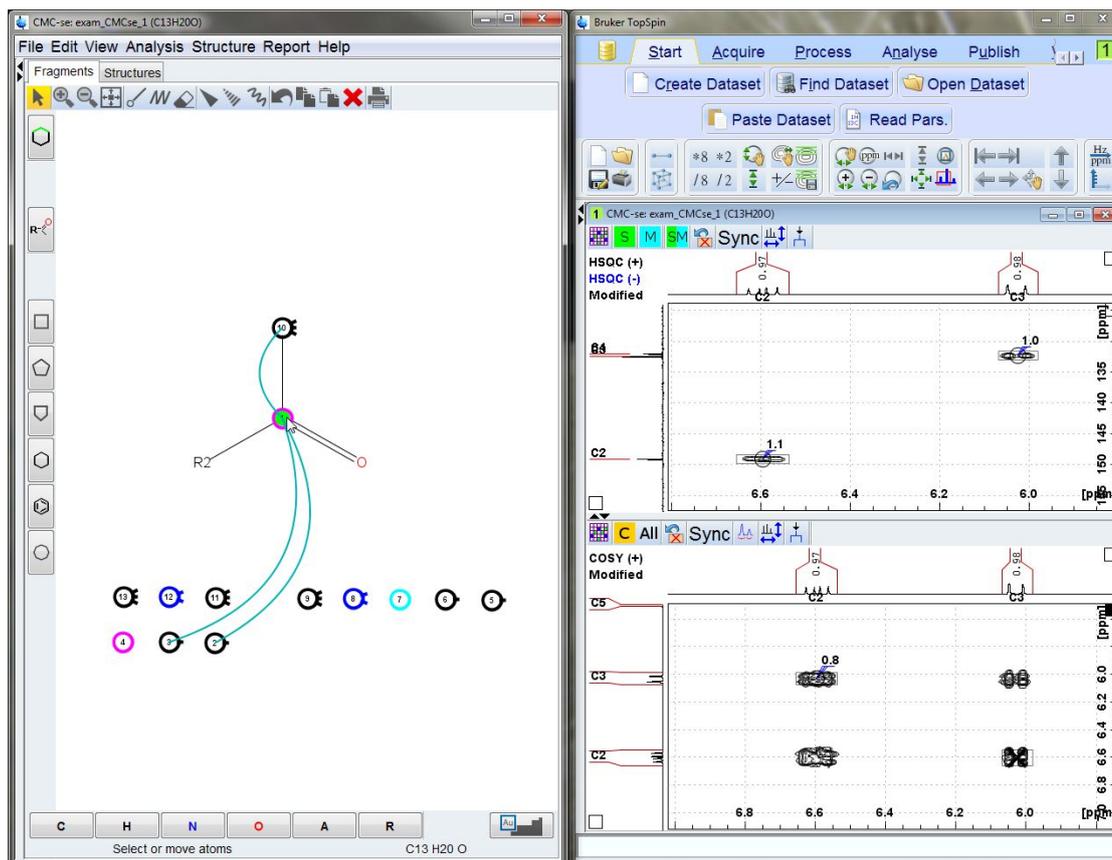


Figure 5.20: Using the NMR data to help extend the current functional group into a larger fragment

From here, C3 can be drug to replace the current R2 group, and then the drawing tools can be used as described in [The Fragment Tool \[81\]](#) to extend the current functional group into a larger fragment as shown in the figure below.

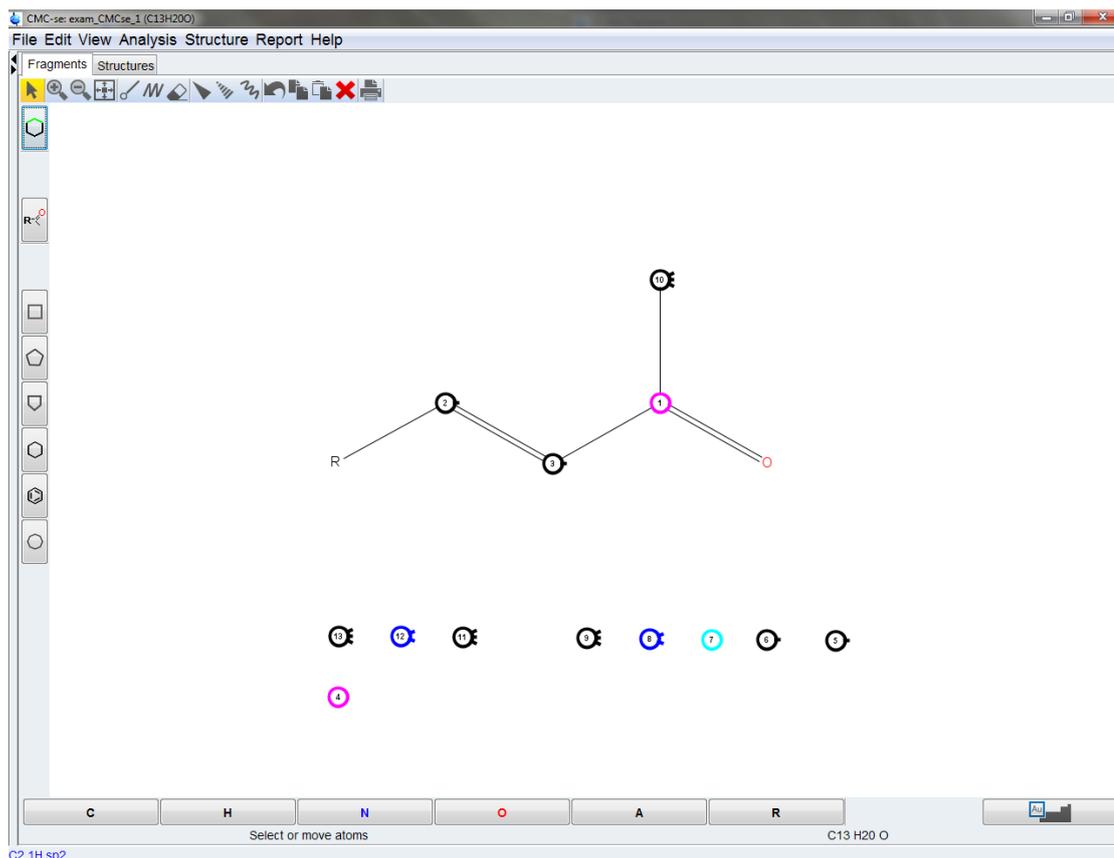


Figure 5.21: The smaller functional group has been extended to a larger fragment

As shown in the figure below, the fragment can then be extended even further because clicking on C2 reveals a COSY correlation to C6, which can then replace the current R group.

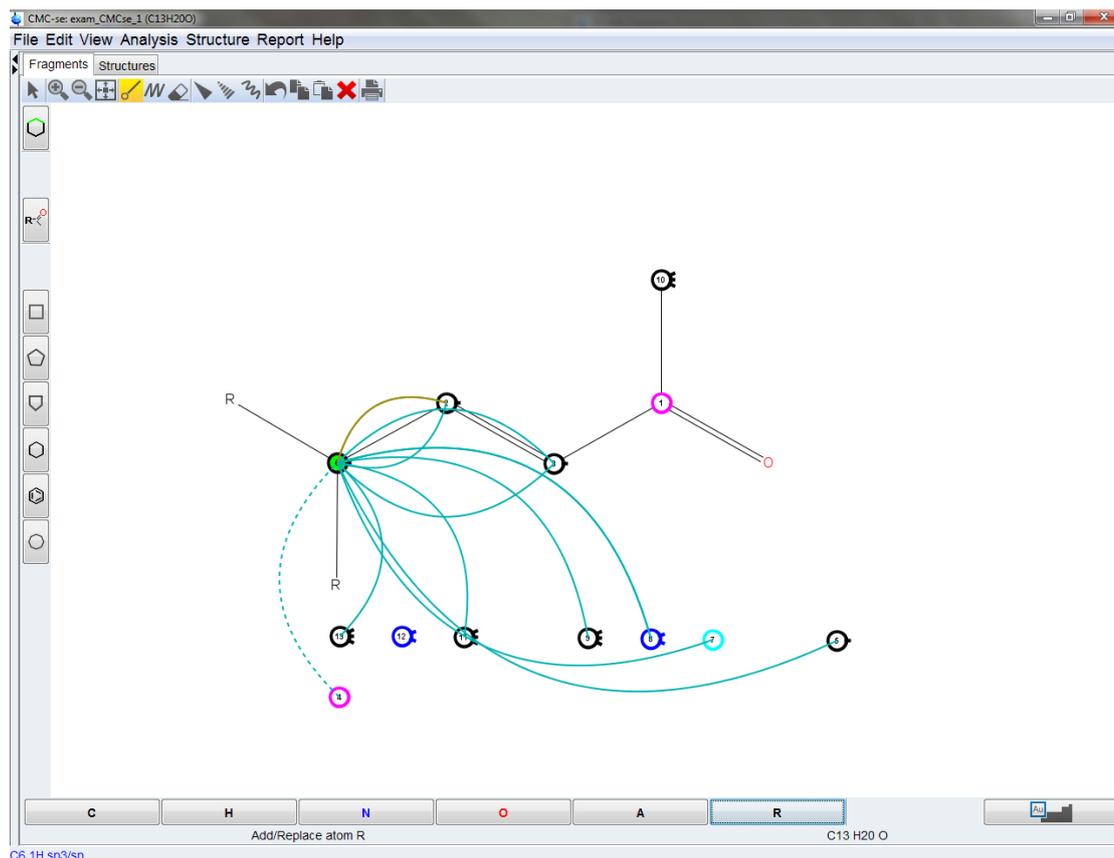


Figure 5.22: Identifying the next atom in the fragment by the COSY correlations

After this atom is assigned to the fragment, the user has hit a point where going further with this fragment is not possible at this time. As shown in the figure below, there are too many HMBC correlations to C6 to identify which 2 additional atoms are also bonded to C6 beyond the ^1H , and C2, so this fragment can be left as is for the time being.

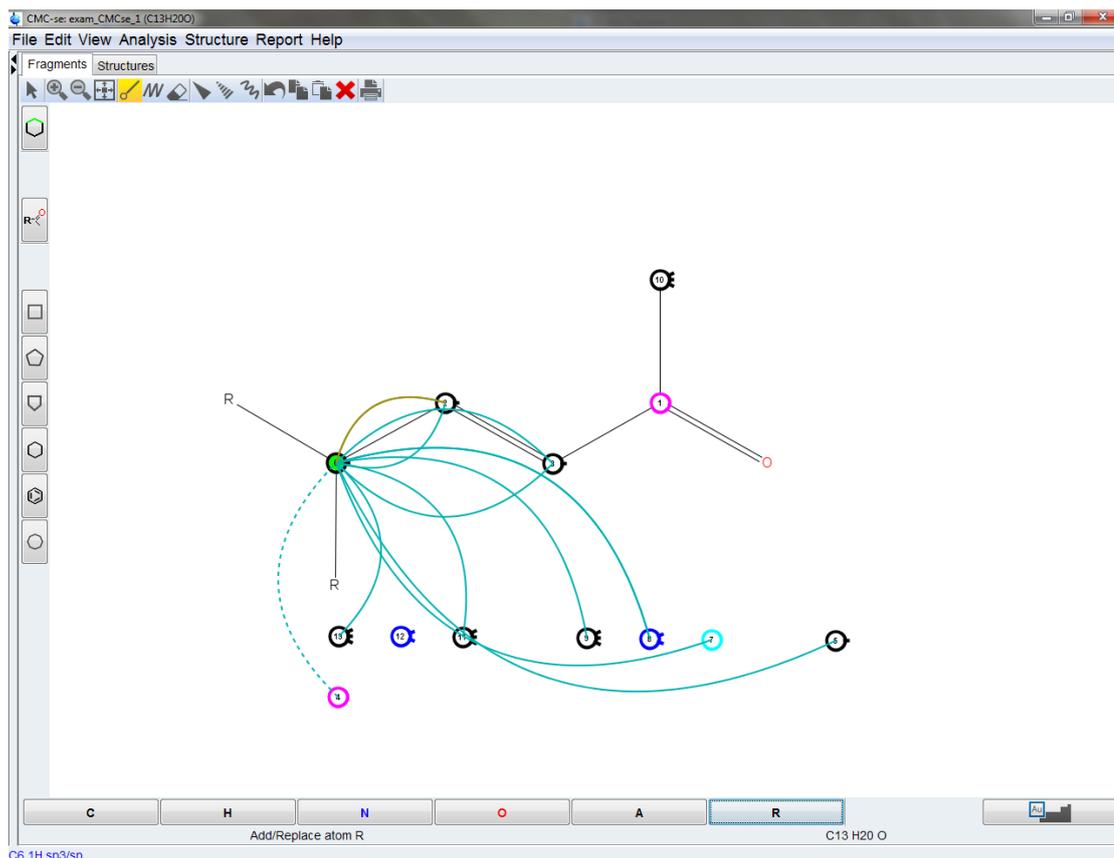


Figure 5.23: No longer clear which atom is next due to too many correlations to C6

The next step after this stopping point on the above fragment is to go back to the Correlation Table and determine if there are any other fragments or functional groups that could be started.

The molecular formula for this molecule ($C_{13}H_{20}O$) indicates there are 4 double bond equivalents (DBE). This information is highlighted in the upper left corner of the Correlation Table as shown in the figure below. There are currently 2 double bonds in the above fragment, so 2 more are necessary. A ring with a double bond is a likely candidate since there are only 2 possible sp^2 hybridized carbons left (C4 and possibly C5).

The screenshot displays the CMC-se exam software interface. The main window is titled "CMC-se exam: CMCse_1 (C13H20O)". The menu bar includes "File", "Edit", "View", "Analysis", "Structure", "Report", and "Help". The toolbar contains various icons for file operations and analysis. The main area is divided into two panes: a "Correlation Table" on the left and a "Fragments" window on the right.

The "Correlation Table" shows the following data:

Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
						6.60	6.03	5.47	2.26	2.23	2.03	1.54	1.44	1.20	0.83	0.83	-
						C2	C3	C5	C6	C10	C12	C13	C8	C8'	C11	C9	H12
C13	22.72	3		sp3							S+	S+	M	M			
C12	22.97	2		sp3									M	M			
C11	26.75	3		sp3											S+	M	
C10	26.89	3		sp3	Frag		M			S+							
C9	27.74	3		sp3											M	M	S+
C8	31.13	2		sp3									S+	S+			
C7	32.46	0		sp3/sp		M		M		M			M	M	M	M	
C6	54.26	1		sp3/sp	Frag	M	M	S+	M				M	M	M	M	
C5	122.62	1						S+	M				M	M			
C4	131.87	0		sp2		M	M		v				M	M			
C3	132.30	1		sp2	Frag		S+	M		M							
C2	148.95	1		sp2	Frag	S+			M								
C1	198.35	0		sp2	R2-C=O	M	M										

The "Fragments" window shows a chemical structure diagram with atoms numbered 1 through 13. The structure is a branched alkane chain with a carbonyl group. The atoms are labeled as follows: 1 (C=O), 2 (C-C), 3 (C-C), 4 (C-C), 5 (C-C), 6 (C-C), 7 (C-C), 8 (C-C), 9 (C-C), 10 (C-C), 11 (C-C), 12 (C-C), 13 (C-C). The diagram also shows a "Fragments" window with a list of atoms and their corresponding chemical symbols (C, H, N, O, A, R).

Figure 5.24: Using the Correlation Table to determine additional fragments that can be defined

Because there are 6 carbons left that could go in a ring (5 carbons not assigned in fragment, not including the 3 CH₃ groups, and the C₆ which still has potential binding partners), a 6 member ring is a valid assumption. The templates in the Fragments window can be used to easily add the 6 member ring to the workspace as shown in the figure below.

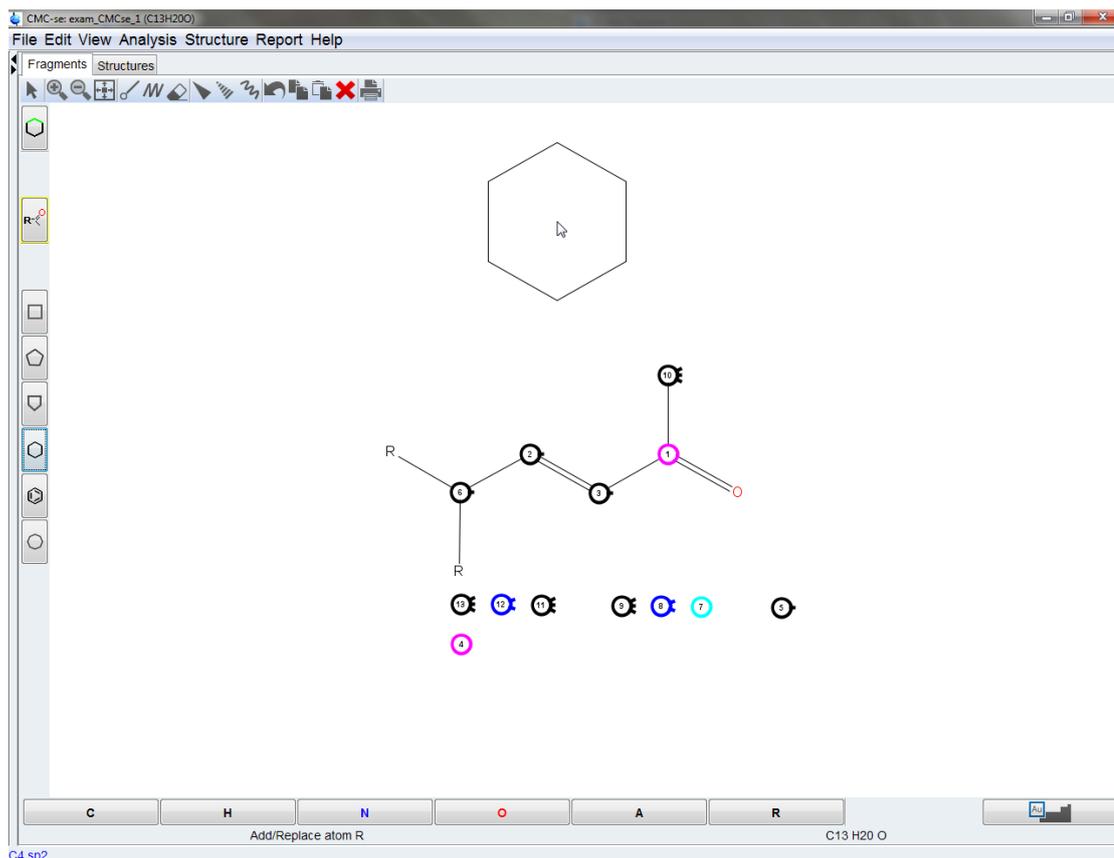


Figure 5.25: Using the templates to add a 6 member ring to the Fragments tool workspace

Because one carbon on this ring has to be C6, and 2 of them the other R groups, the ring should be selected by clicking and dragging with the  tool, and then clicking and dragging it to align with the current fragment as shown in the figure below.

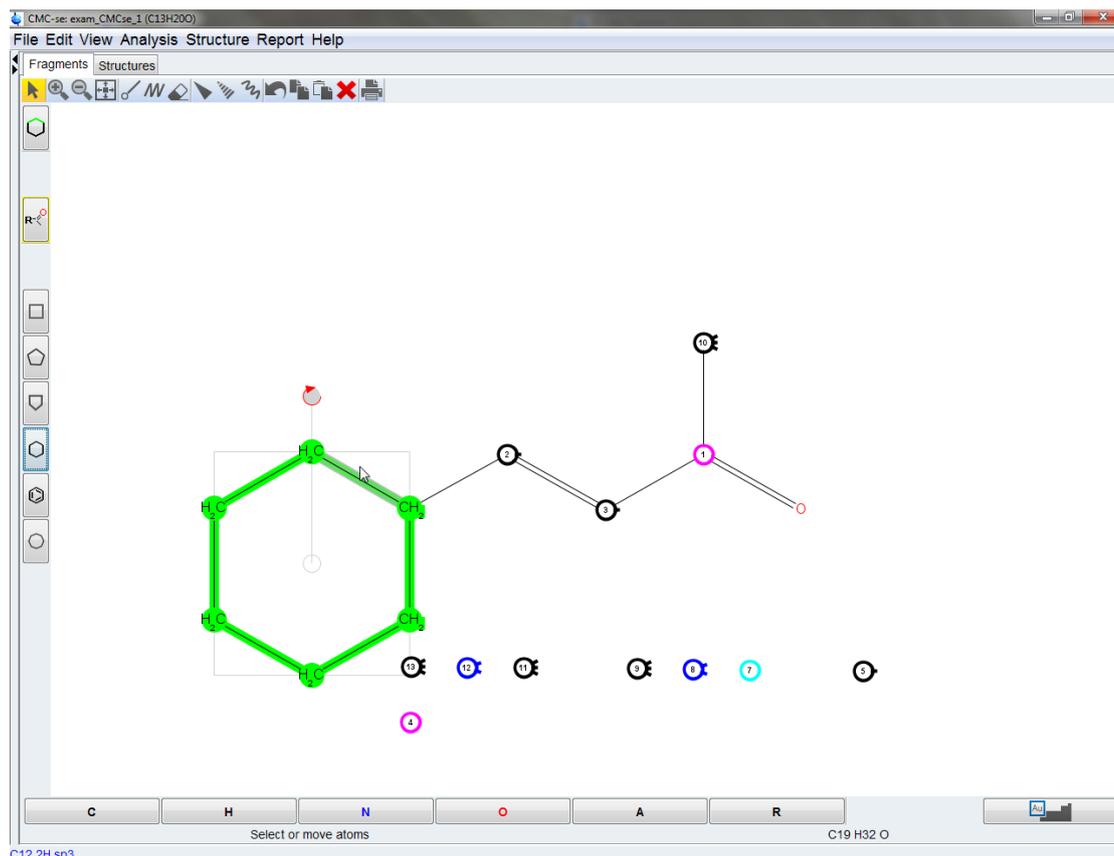


Figure 5.26: Selecting and merging two fragments into one in the Fragments tool

It was seen previously that there are too many correlations from C6 to use it to assign any additional atoms on the ring, however if one clicks on C2 as shown in the figure below, it is clear that there are 2 4° carbons (C4 and C7) neighboring C6 in the ring.

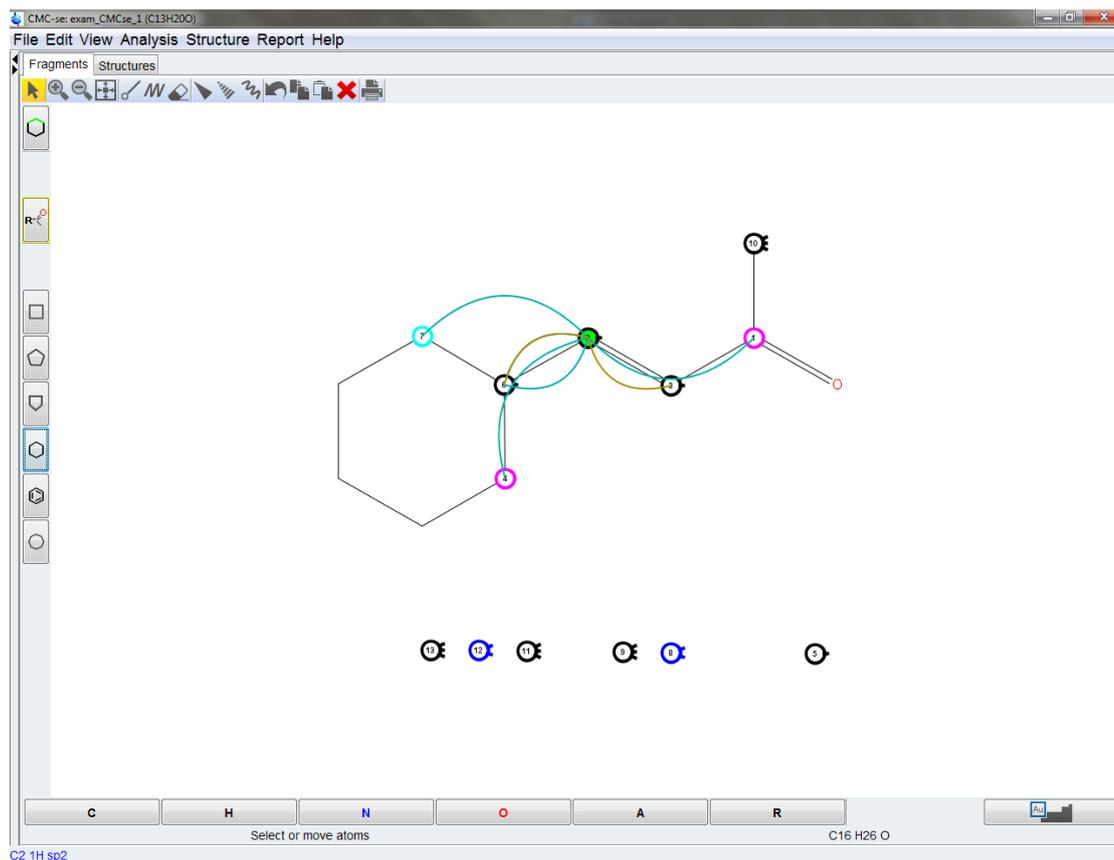


Figure 5.28: Two additional carbons assigned to the ring

Now if the user goes back to see the correlations to C6, it is seen in the figure below that there is an HMBC to C5.

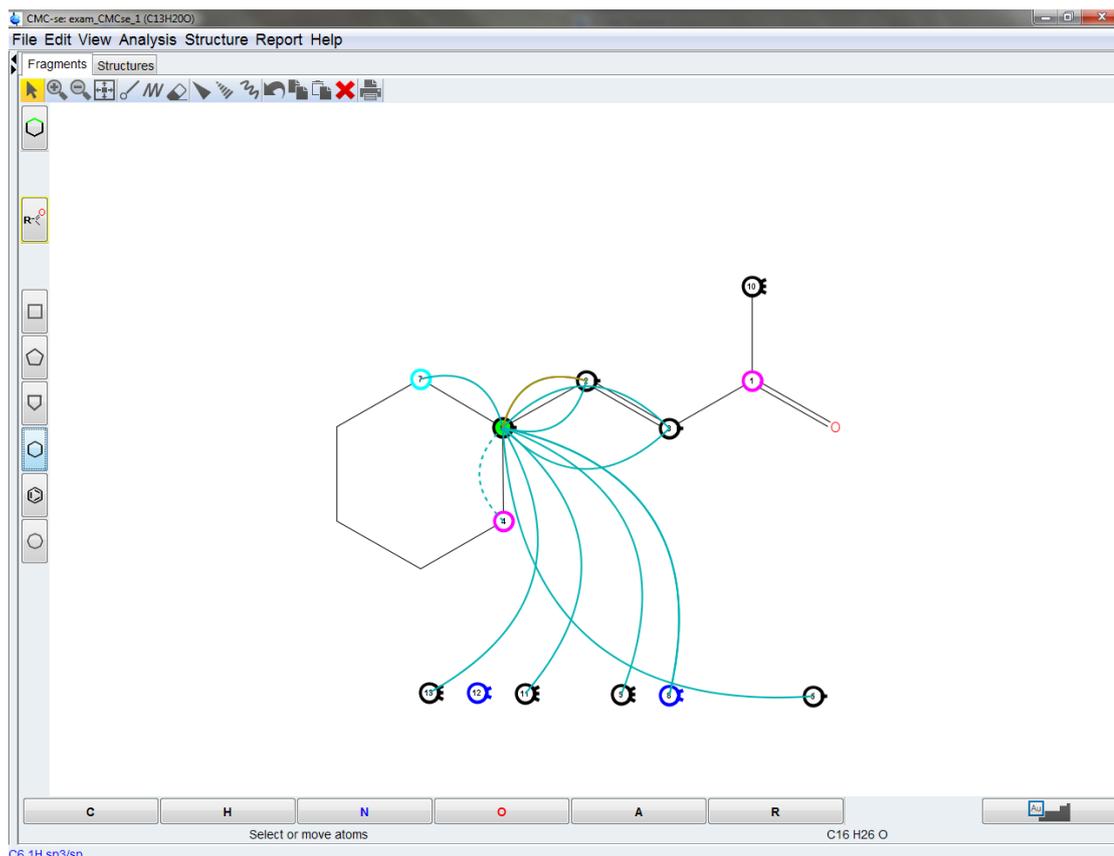


Figure 5.29: HMBC correlations identify C5 as one of the next members of the ring

It was noted above that C5 was potentially a sp^2 hybridized carbon, so this must be the carbon that is next to C4 in the ring, and a double bond can be defined between them as shown in the figure below.

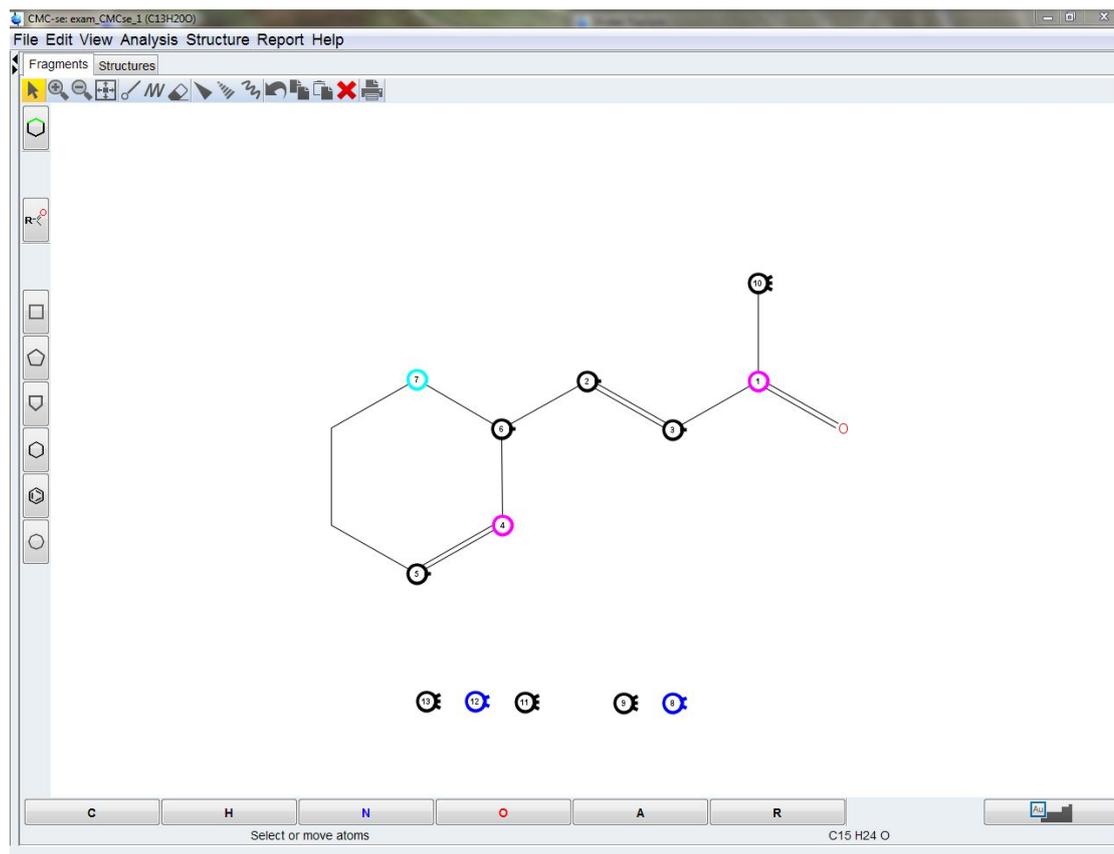


Figure 5.30: Correlations and chemical shift identify the position of the double bond in the ring and it is assigned accordingly

Going back to C4, it is shown in the figure below that there are correlations from the carbon to both C12 and C13. Therefore 2 additional assignments can be made to the ring. C12 the CH₂ group, must be the next member of the ring, and C13, is a CH₃ group that must be sticking off the ring at C4.

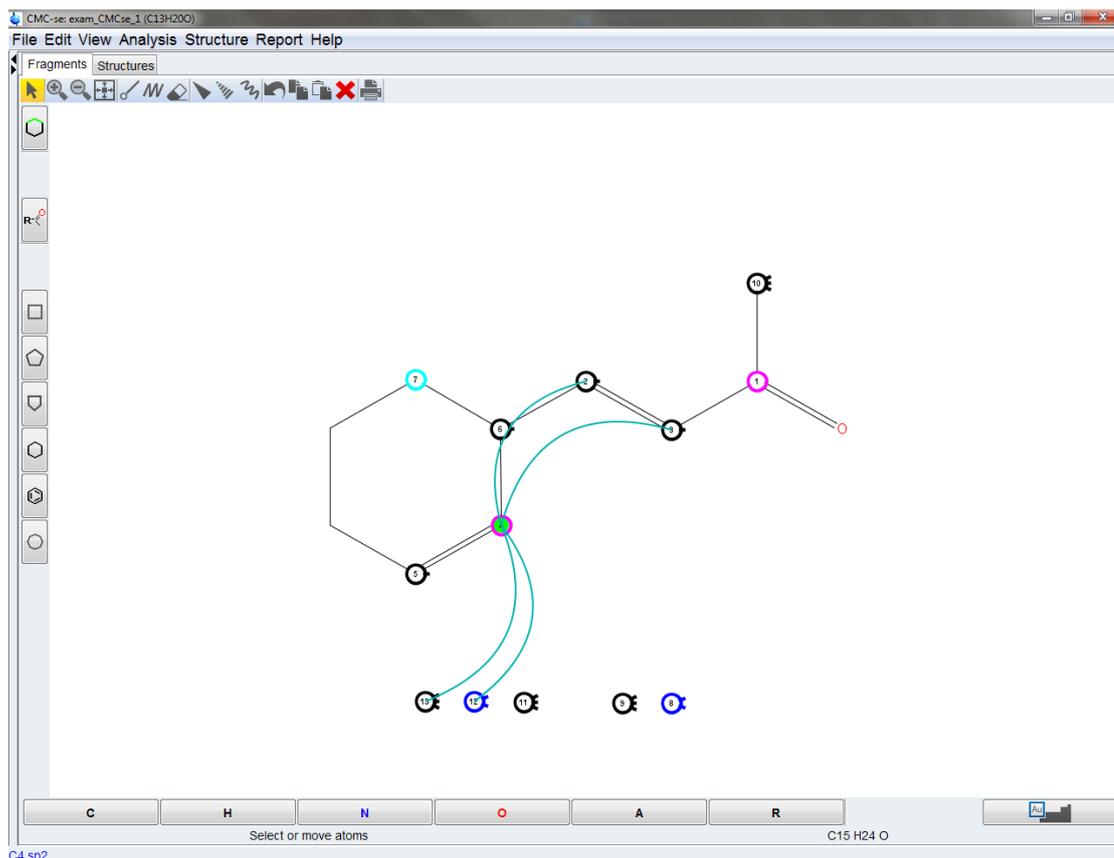


Figure 5.31: Additional HMBC correlations to further elucidate the molecular structure

As shown below, the remaining carbons can easily be assigned from here. The only position left on the ring must correspond to the only non CH₃ carbon left which is C8. Then C7 which is a 4° carbon with currently only 2 bonds must have the remaining CH₃ groups (C11, and C9) attached to it.

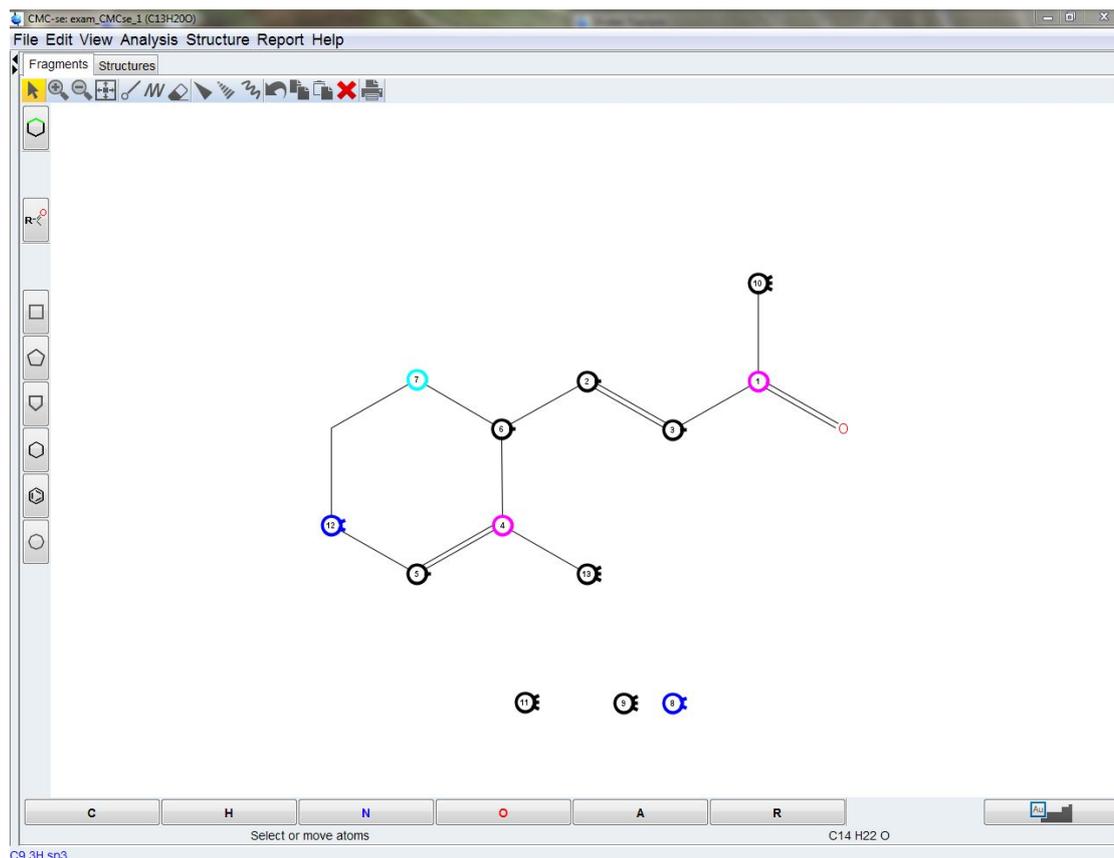


Figure 5.32: The remaining carbon on the ring can easily be assigned, and the structure completed by process of elimination

Now that the entire structure is complete and all atoms from the table have been assigned, it is possible for the user to use the chemical shift prediction routines from the CMC-se verification tool (See the chapter [The CMC-se Structure Verification Tool \[125\]](#) for details) to double check the accuracy of the manually elucidated structure. This is accomplished

through the  button, or through the pull down menu as shown in the figure below.

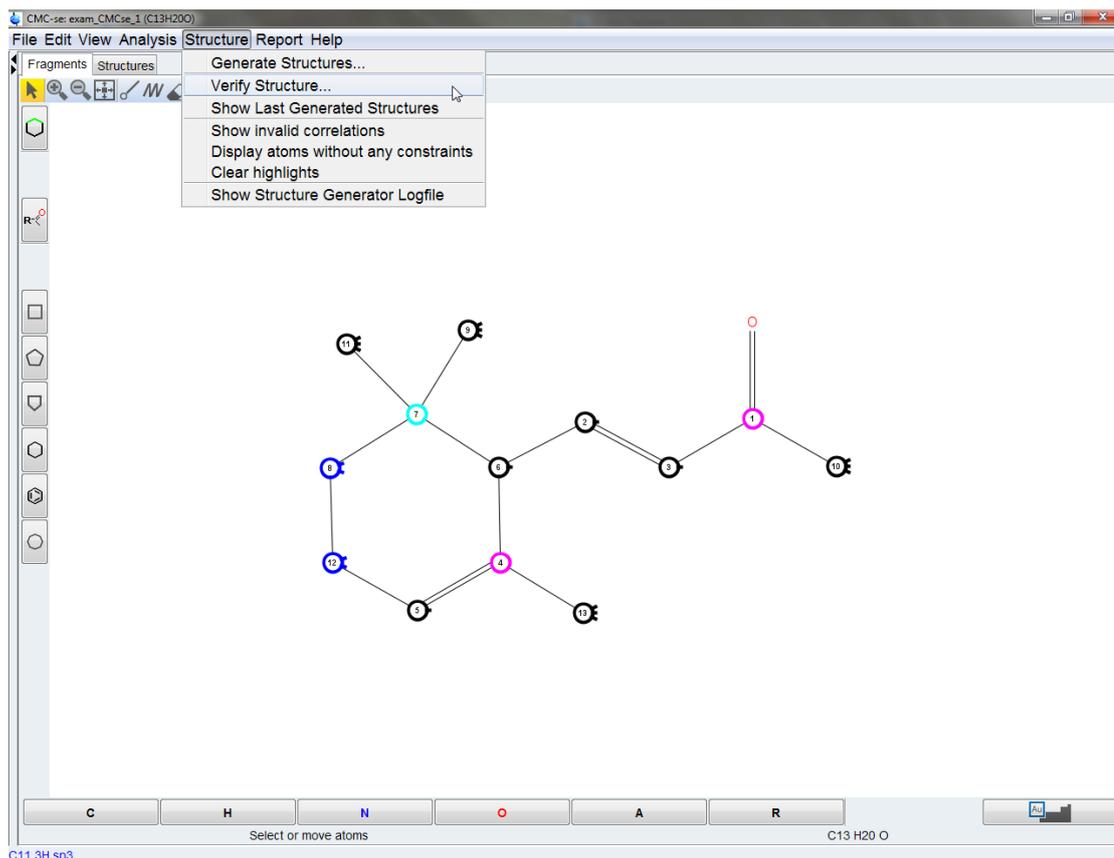


Figure 5.33: Using the CMCse verification tool to check the accuracy of the manually elucidated structure

The user needs to make sure to select the **Assigned Fragment** option from the Verification settings dialogue box as shown below and then the **OK** button to start the verification process.

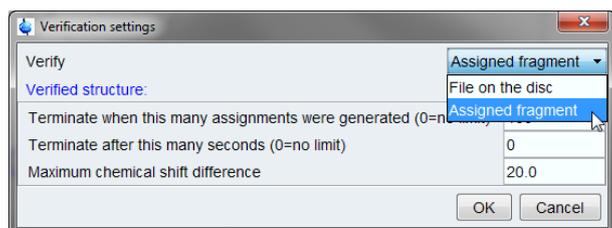


Figure 5.34: Selecting the Assigned Fragment option, so that the currently assigned fragment will be run through the verification routine

Once the verification is complete, then the structure is shown in the Structures tab along with the chemical shift prediction results just as if the structure was generated automatically by the software. This is shown in the figure below.

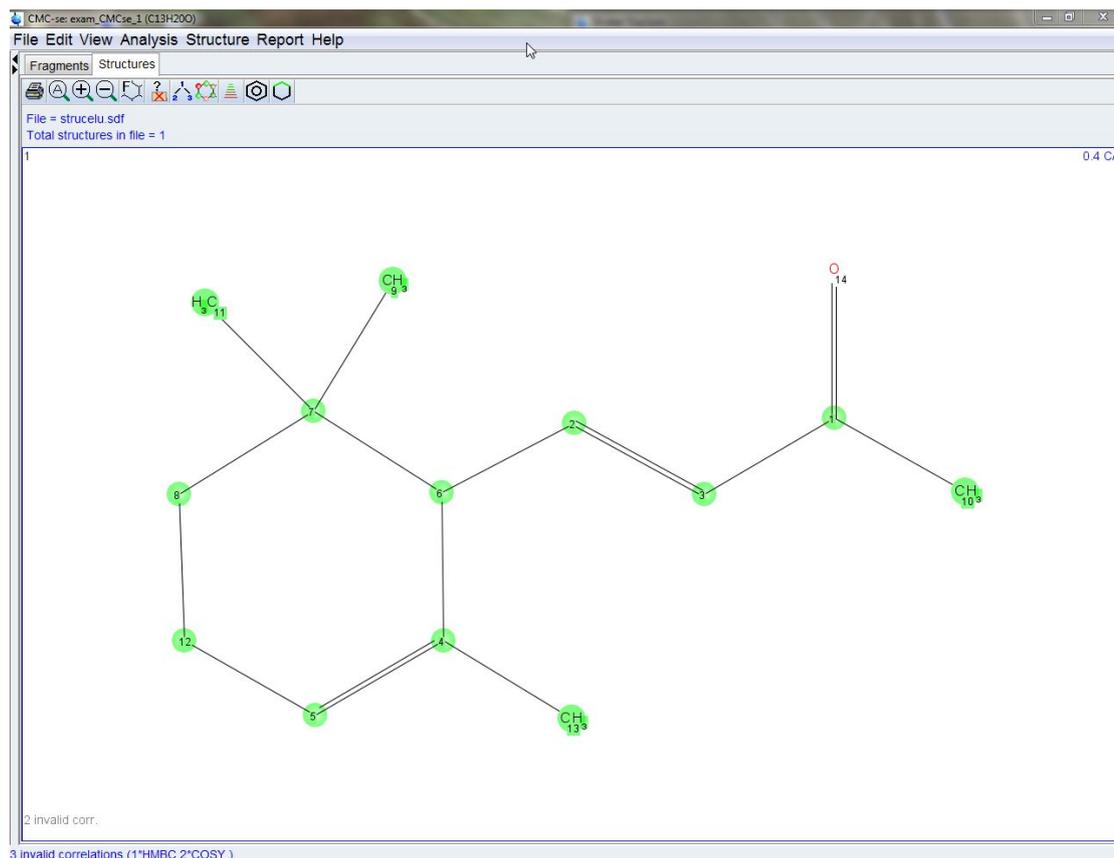


Figure 5.35: Results of the chemical shift prediction from the Verification routine are shown in the Structures tab

The user from here can then use all the regular tools as previously described in [User Validation of Generated Structure Proposals \[101 \]](#) to continue to verify the correctness of the manually elucidated structure.

5.2 Structure Elucidation for Guaiol

The next example is Guaiol (**exam_CMCse_2**). The molecular formula for this compound is $C_{15}H_{26}O$, and the sample was prepared in $CDCl_3$.

While this is still a relatively small molecule this particular data set represents a variety of common challenges (and solutions to those problems), that are likely encountered during a structure elucidation process. Additionally, it serves as an example for why adequate resolution is important in the spectra to be analyzed by the automatic analysis tools of CMC-se.

The following spectra have been acquired and processed properly:

- PROTON [1H]
- C13 [^{13}C]
- COSY [1H , 1H]
- HSQC [1H , ^{13}C]
- HMBC [1H , ^{13}C]
- HSQC-TOCSY [1H - ^{13}C]

Again a new structure elucidation project including all spectra was created (section [Creating a New CMC-se Project \[▶ 23\]](#)). The pop-up box automatic analysis options box for the Automatic Analysis for this data set is shown in the following figure:

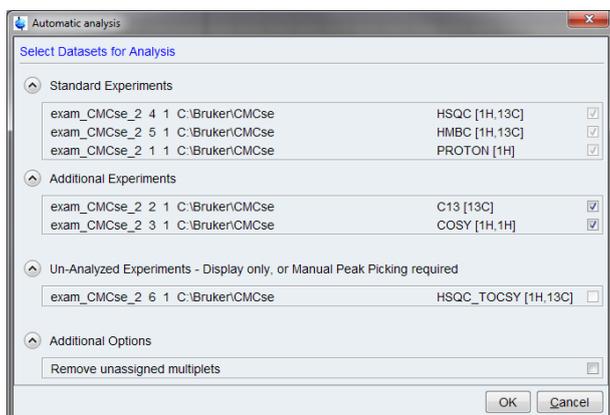


Figure 5.36: Automatic analysis options illustrating that some but not all experiments have the option of being analyzed or not.

In the above figure is the option to remove those multiplets that are not assigned is also shown, but not checked. A few other things should also be pointed out:

- **Standard Experiments:** Shows those spectra which are required for the generation (Proton, HSQC and HMBC) are automatically checked and there is not the option to not use them.
- **Additional Experiments:** Contains other experiments which are not required, but can be automatically analyzed and used for the structure generation.
- **Un-Analyzed Experiments:** The HSQC-TOCSY which is not used in the generation, but only for evaluation, is not checked, nor is there the option to do so. Many different experiments could be found in this section, both those like the HSQC-TOCSY which are not used in the generation process, and those like the INADEQUATE which are used, but require manual peak picking.
- There is the option to **Remove unassigned multiplets** which will if desired remove any peaks which are picked but cannot be assigned to the Correlation Table so that they are never seen by the user. This option is NOT used in this example. And the reason for this will be illustrated.

The results of the Automatic Analysis are shown in the figure below:

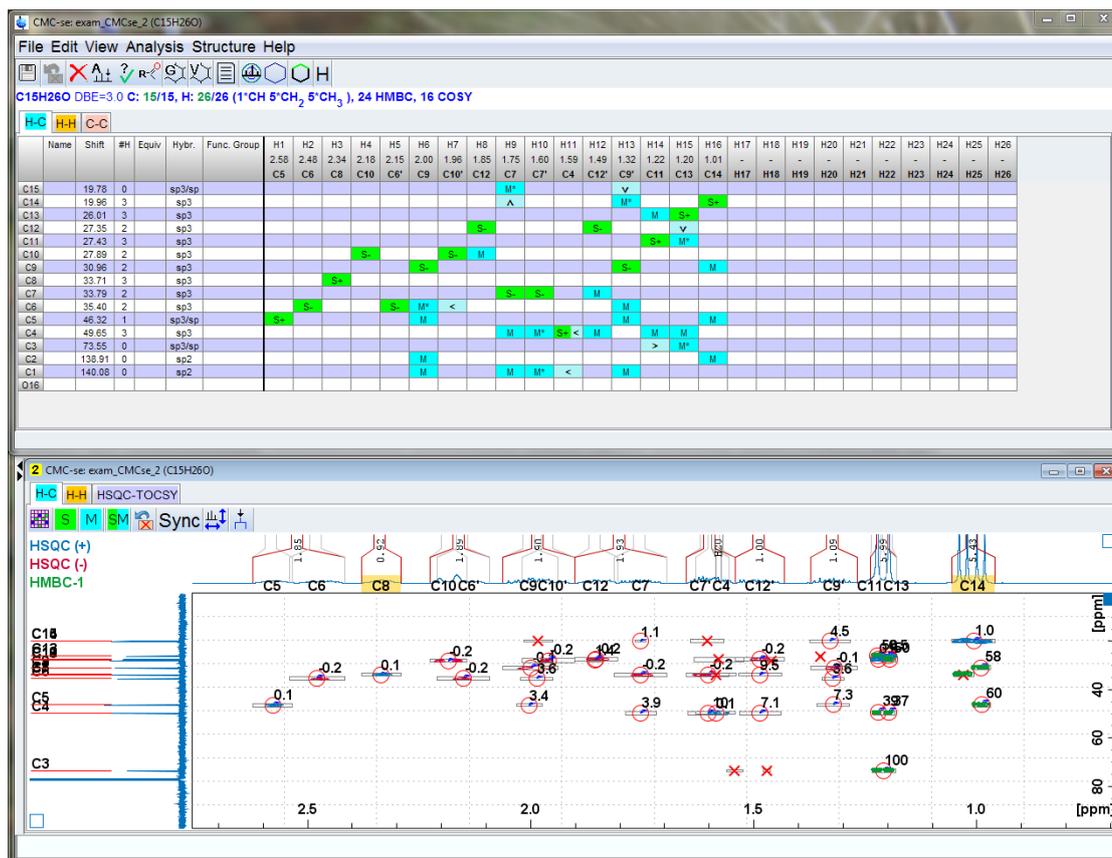


Figure 5.37: Results of the Automatic Analysis

In the figure above, the yellow highlights on the ^1H projection of the Combined Spectra Display indicate a potential problem.

Unlike the previous example, a quick examination of the Combined Spectra Display immediately indicates a potential problem. There are yellow highlights in multiple areas of the ^1H projection. This is a visual indication that the ^1H consistency check failed. The Consistency

Check Report () is shown below. In cases like this, the user should manually inspect these regions and fix any problems prior to proceeding with the generation:

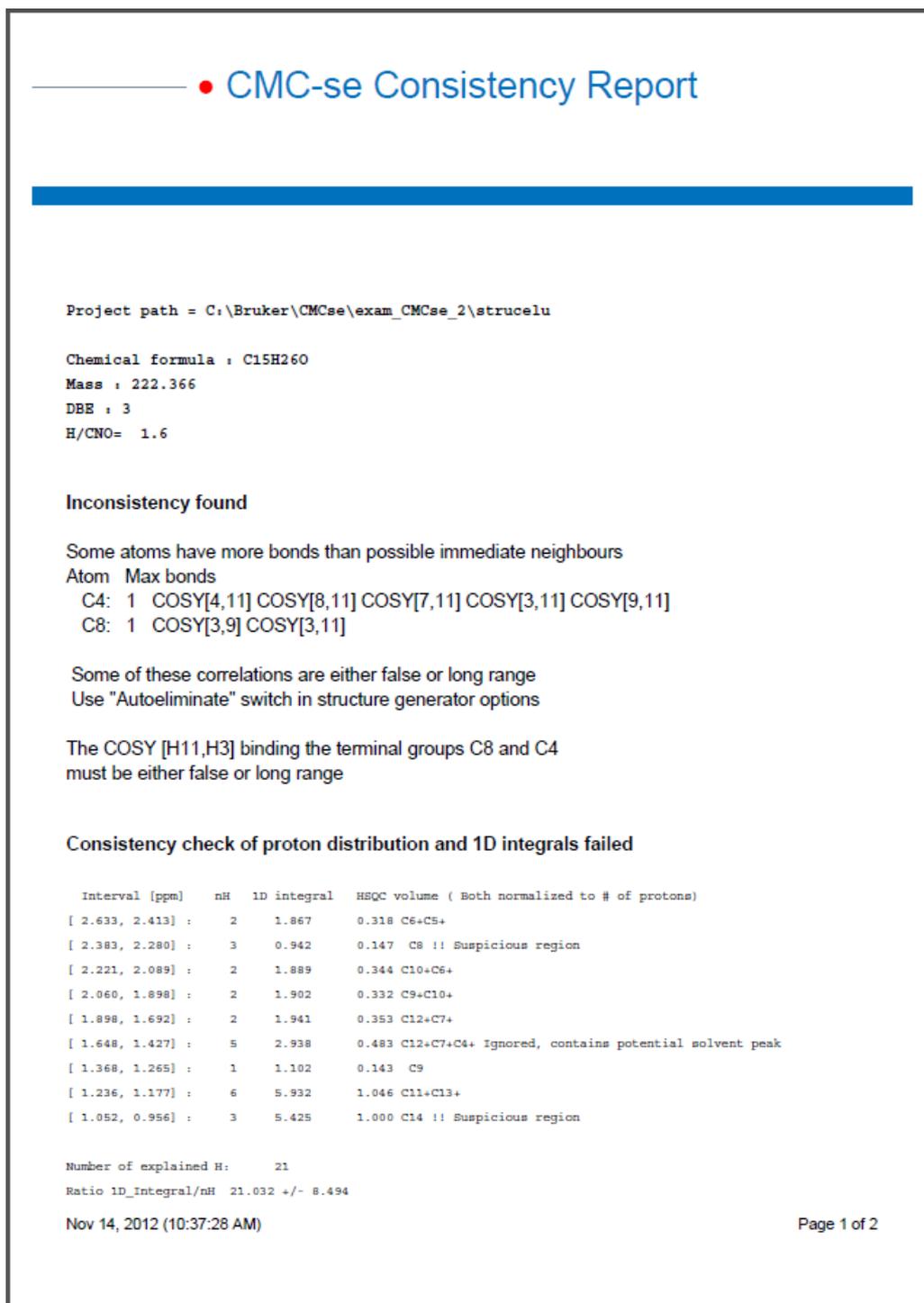


Figure 5.38: Report from the Consistency Check.

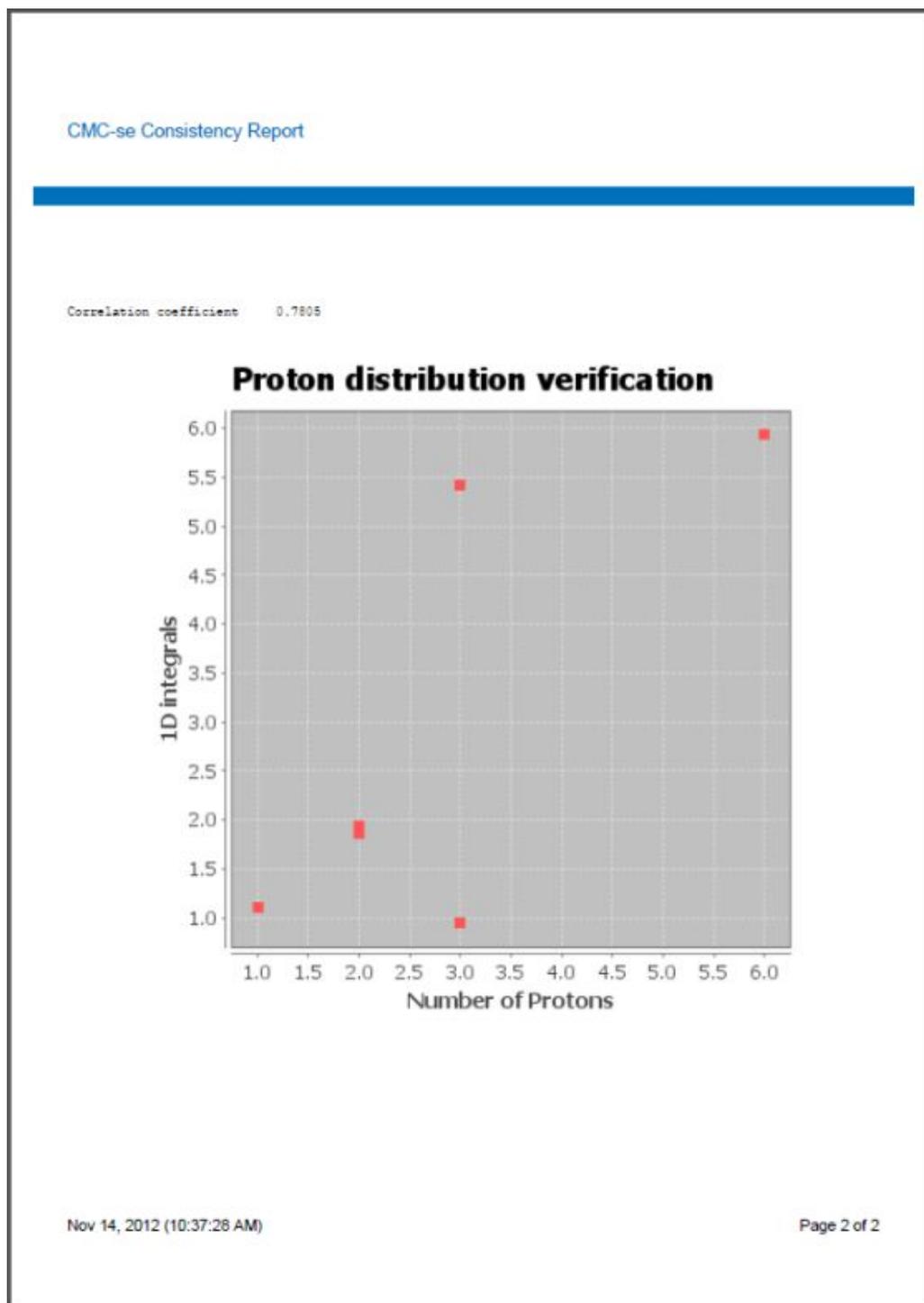


Figure 5.39: Report from the Consistency Check.

Zooming into the HSQC around the area of the first yellow highlight of C14, the cause of the inconsistency is apparent and shown in the figure below. There was not enough resolution in the HSQC spectrum, and the Automatic Analysis has picked these overlapping CH₃ groups as a single peak. This peak needs to be deleted, and the individual peaks manually picked. This is the example that was used in section [Adding and Removing Multiplets Through the Combined Spectra Display \[57\]](#), so the user is directed back to that section for instructions on how to manually delete and add the correlations.



Figure 5.40: Region of the HSQC spectrum in the Combined Spectra Display indicating the source of the inconsistency.

In the figure above, the 2 overlapping CH₃ groups have been picked as a single peak by the automatic analysis. Manual correction of this is necessary.

Even after correcting the overlapping peaks and manually picking 2 individual HSQC correlations, a quick examination of both the Correlation Table, and the combined spectra display indicate there is still a problem with the analysis. The exact situation will depend on whether the user deleted the incorrect correlation and drew 2 new ones, or resized the existing correlation and only added 1 new correlation.

As seen in the figure below, there is a light red highlighted section of the Correlation Table indicating that too many proton signals are included (27/26). And the yellow highlights on the ¹H projection of the Combined Spectra Display are still present.

If the existing correlation had been deleted instead of resized, there would still be the yellow highlights on the ¹H projection, but the number of protons would be 25/26 as the new HSQC correlations would be assigned as one ¹H each instead of the existing one which was assigned to three ¹H.

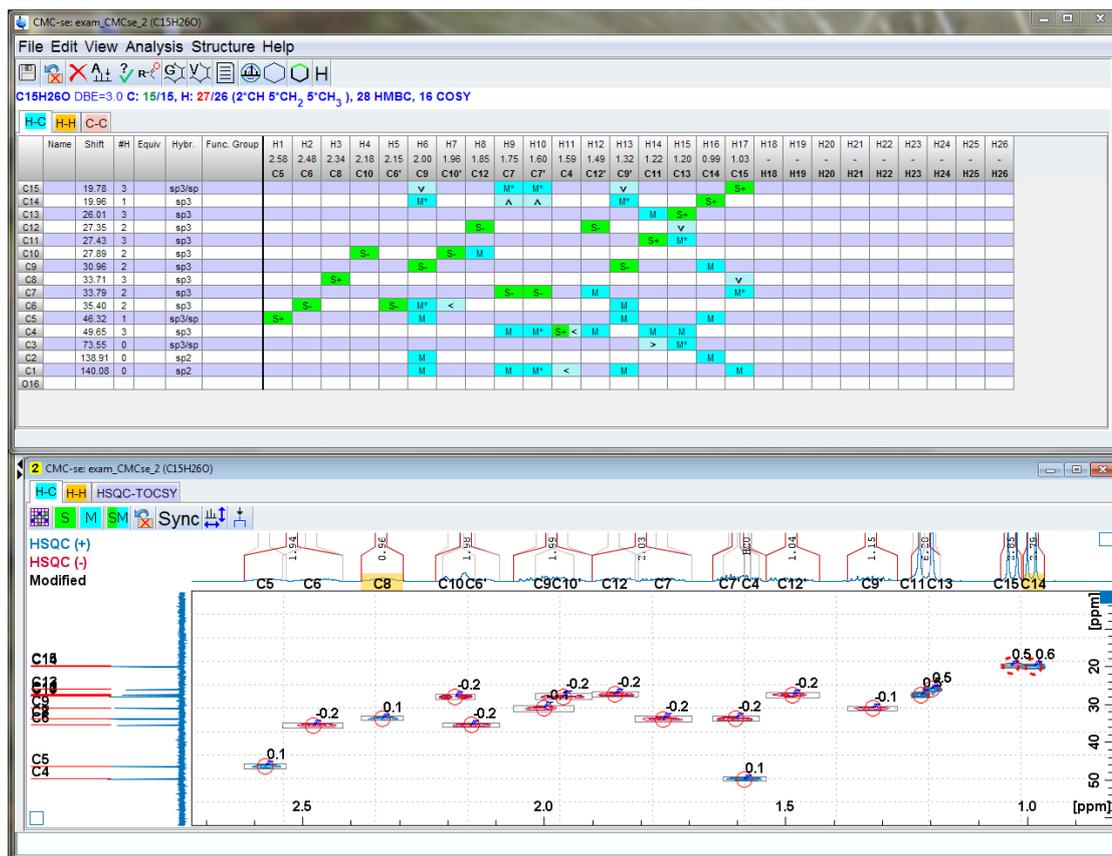


Figure 5.41: Even after manually correcting the overlapping correlations, there are still problems with the Correlation Table; too many protons are included.

As shown in the figure below, double clicking on the HSQC peak corresponding to C15 (where there is a yellow highlight to indicate the ^1H consistency check fails) brings up the Atom Properties window. It is known from the ^1H integration (also shown on the projection) that this is a CH_3 group, so this should have a value of 3 for Number of Attached H atoms where it is currently 1. However, while this is the correct change to make, it only adds more protons to the Correlation Table when the problem was that there were already too many. Therefore, other regions of the spectrum need to be examined in the same way.

The next place to look would be the correlation for C8 which is also highlighted in the figure above. That correlation was assigned by the Automatic Analysis as a CH_3 , and it is a CH . This partially fixes the problem, and while the yellow highlights disappear from the spectrum, the red error notification on the Correlation Table is still present, indicating that there is still a problem.

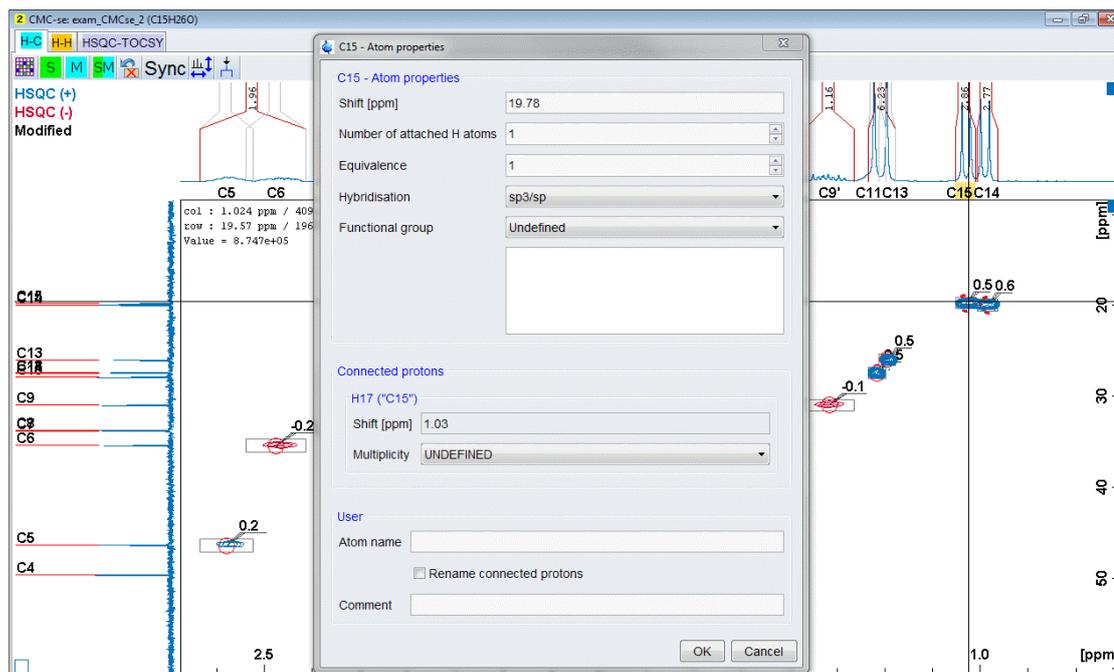


Figure 5.42: Double clicking on the HSQC correlation for C15 brings up the “Atom Properties” window, where the “Number of attached H atoms” can be corrected from the current value of “1” to the appropriate value of “3” for this CH₃ group.

The next place to check would be the region from 1.65 – 1.42 ppm that as stated in the Consistency Check Report, was ignored due to it containing a potential solvent peak. In this region, the HSQC correlation to C4 was assigned as a CH₃ by the analysis, and it should be a CH.

Once the above mentioned changes to the 3 HSQC correlations are made, there are no visible errors to the Correlation Table, or Combined Spectra Display. Note that while the ¹H count is still 25/26 this is OK because everything else is correct, and there is the possibility of an un-observable exchangeable proton. Because there are no errors, it is possible to start the structure generation at this point:

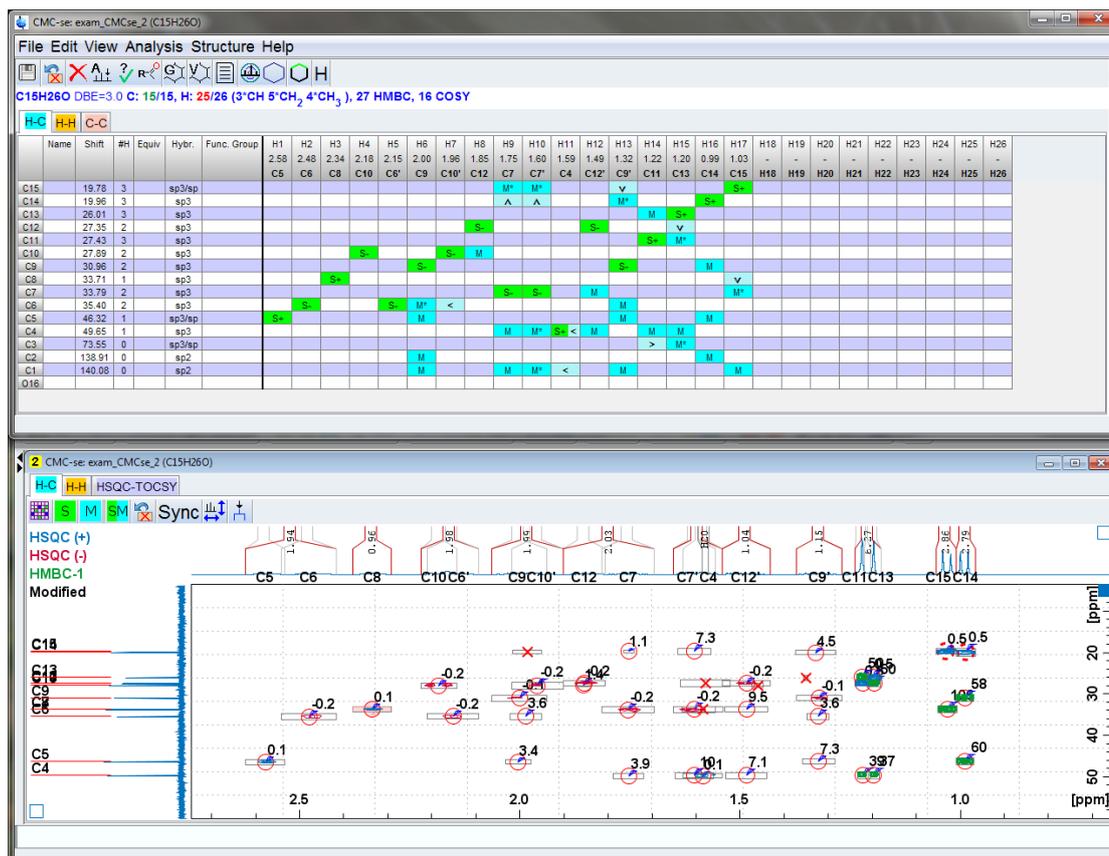


Figure 5.43: Correlation Table and Combined Spectra Display after correcting the proton distribution errors

Just as in the previous example, the Consistency Check Report did point out that there are too many correlations for at least one atom, the auto eliminate option should be activated (for both the COSY and HMBC). In this example because there were many correlations to the single C7 and C10 that were identified in the report, it will be set to 3, as shown in the figure below:

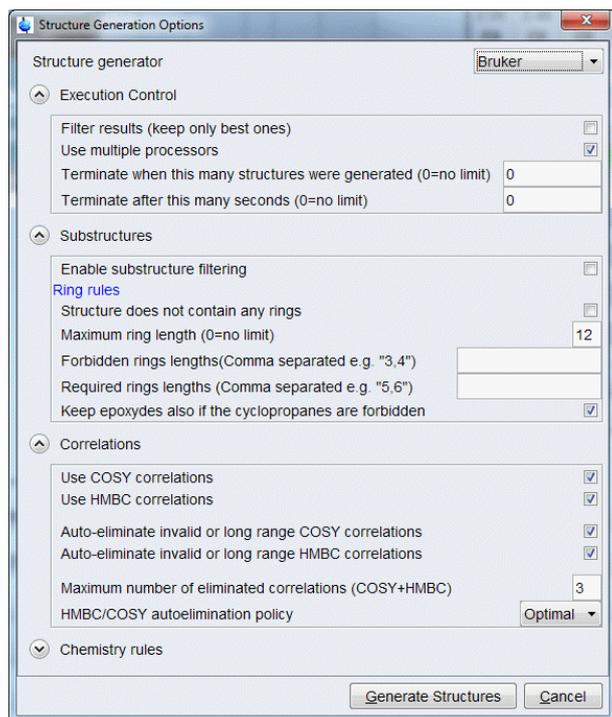


Figure 5.44: Structure generation options with the maximum number of automatically eliminated correlations (COSY and HMBC) has been set to 3.

Using these generation options, 3 structures are generated and shown in the figure below.

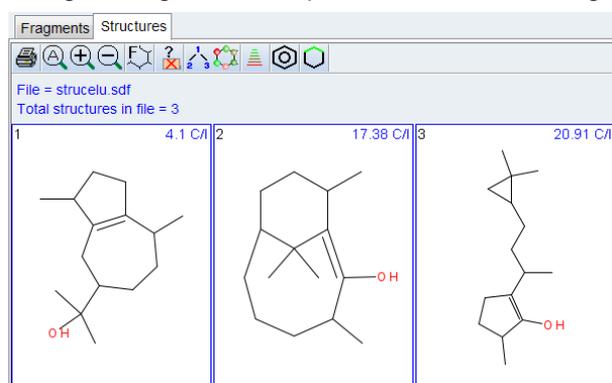


Figure 5.45: Generated structure candidates when the auto elimination # was set to 3

While the first structure of these may seem reasonable, the chemical shift prediction values are all relatively high. So especially in cases like this where no structure is standing out as being correct, it is highly recommended (sections [Correlations \[92\]](#) and [Process Overview \[120\]](#)) to now go back and increase the flexibility of the generation process by increasing the number of auto-eliminated correlations. When this is set to 4 instead of 3, there are 23 structures which are generated.

While it is possible to continue with the ranking and verification process with these 23 structures, this number could be reduced by going back and adding additional information to the Correlation Table, as suggested in the last flowchart in [Process Overview \[120\]](#).

While the HMBC may have been automatically added, this is still a good region of the spectra to start an initial examination for additional correlations that could be added.

Clicking to the H-H tab of the Combined Spectra Display, as shown in the figure below, it is seen that there is a COSY correlation that can be added to the table from this C15 that was manually added to the table. In the figure below, the cursor is poised above the button that should be pushed to enter the mode to manually define a new peak. This is the same process that was used above to manually add the HSQC correlation(s).

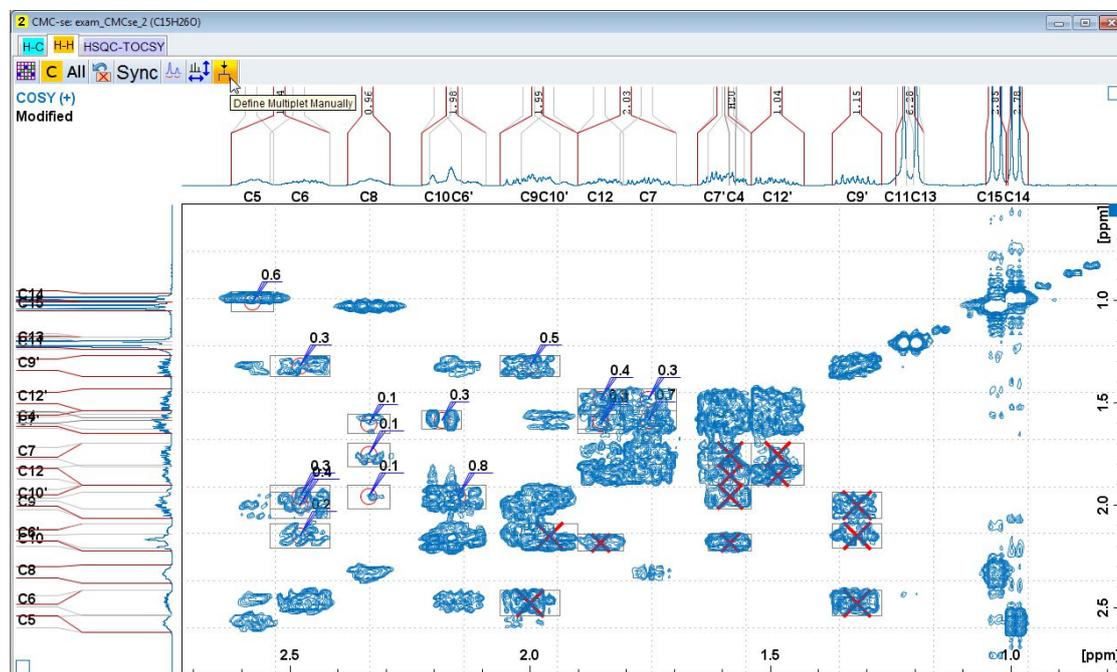


Figure 5.46: In the H-H tab, there is un-picked COSY peak between H/C15 and H/C8 that can be manually picked and added to the table.

As described in section [Adding Additional Information to the Correlation Table \[76 \]](#), another type of constraint which can be added is the multiplicity of CH₃ and aromatic protons. All four of the CH₃ signals in this molecule are clearly either doublets or singlets. Defining them as such will dictate to the generation algorithm that they are next to a carbon with a single proton, or no protons respectively.

Double clicking on the HSQC peak in the Combined Spectra Display will once again bring up the “Atom Properties” box where the multiplicity can be defined by the user as shown in the figure below.

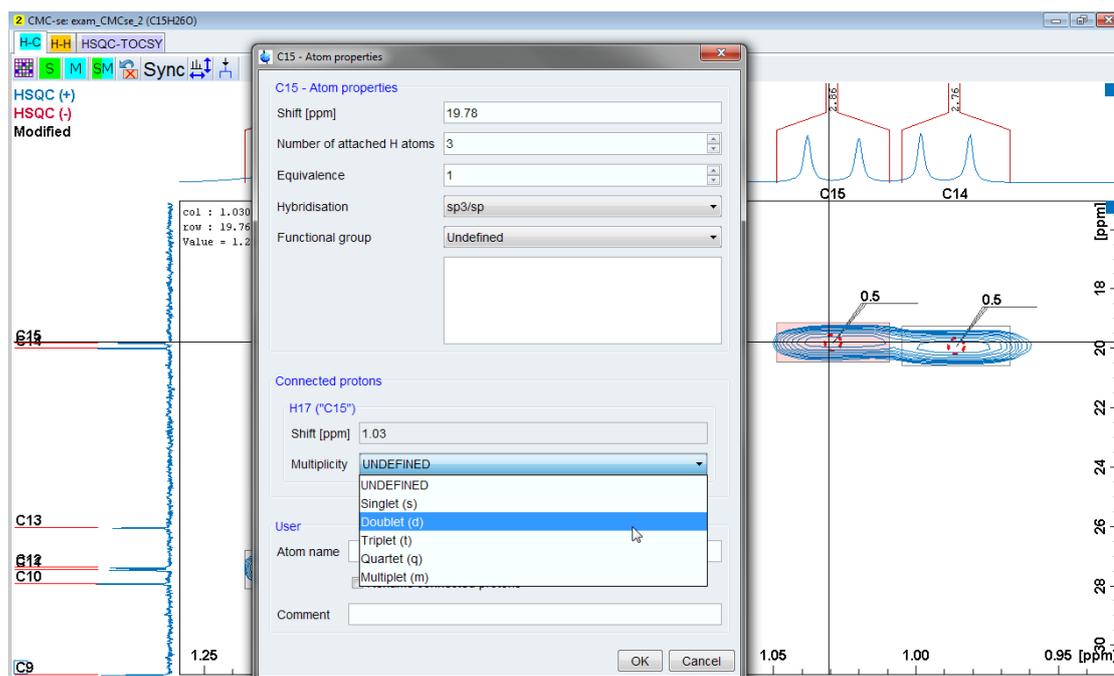


Figure 5.47: Double clicking on an HSQC peak to bring up the Atom Properties dialog box so that the proton multiplicity can be defined.

Now that these additional constraints have been added, the constraints used by the structure generation are summarized in the Correlation Table shown below. For illustration purposes, the Correlation Table has been toggled back to the old layout so that both H-C and H-C correlations are shown on the same screen, and the empty columns/rows have been removed (see section [The Correlation Table](#) [33] for more information regarding manipulating the viewing options of the Correlation Table).

CMC-se: exam_CMCse_2 (C15H26O)

File Edit View Analysis Structure Help

C15H26O DBE=3.0 C: 15/15, H: 25/26 (3^oCH 5^oCH₂ 4^oCH₃), 27 HMBC, 16 COSY

H-C C-C

	Name	Shift	#H	Equiv	Hybr.	Func. Group	H1 2.58 C5	H2 2.48 C6	H3 2.34 C8	H4 2.18 C10	H5 2.15 C6'	H6 2.00 C9	H7 1.96 C10'	H8 1.85 C12	H9 1.75 C7	H10 1.60 C7'	H11 1.59 C4	H12 1.49 C12'	H13 1.32 C9'	H14 1.22 C11	H15 1.20 C13	H16 0.99 C14	H17 1.03 C15	H18 -
C15		19.78	3		sp3/sp										M*	M*			v					S+
C14		19.96	3		sp3										Δ	Δ			M*					S+
C13		26.01	3		sp3															M	S+			
C12		27.35	2		sp3								S-						S-			v		
C11		27.43	3		sp3															S+	M*			
C10		27.89	2		sp3				S-				S-	M										
C9		30.96	2		sp3							S-							S-				M	
C8		33.71	1		sp3				S+															v
C7		33.79	2		sp3										S-	S-			M					M*
C6		35.40	2		sp3																			
C5		46.32	1		sp3/sp		S+																	
C4		49.65	1		sp3																			
C3		73.55	0		sp3/sp																			
C2		138.91	0		sp2																			
C1		140.08	0		sp2																			
O18																								

	Name	Shift	Mut	H1 2.58 C5	H2 2.48 C6	H3 2.34 C8	H4 2.18 C10	H5 2.15 C6'	H6 2.00 C9	H7 1.96 C10'	H8 1.85 C12	H9 1.75 C7	H10 1.60 C7'	H11 1.59 C4	H12 1.49 C12'	H13 1.32 C9'	H14 1.22 C11	H15 1.20 C13	H16 0.99 C14	H17 1.03 C15	H18 -		
H18	H18																						
H17	C15	1.03	d																				
H16	C14	0.99	d		c																		
H15	C13	1.20	s																				
H14	C11	1.22	s																				
H13	C9'	1.32				c																	
H12	C12'	1.49																					
H11	C4	1.59					c																
H10	C7'	1.60																					
H9	C7	1.75																					
H8	C12	1.85																					
H7	C10'	1.96																					
H6	C9	2.00																					
H5	C6'	2.15																					
H4	C10	2.18																					
H3	C8	2.34																					
H2	C6	2.48																					
H1	C5	2.58																					

Figure 5.48: Final Correlation Table that will be used for the generation for this example data set.

With these additional constraints, there are now only 19 (compared to the previous 37) structures that are generated. A subset of these structures is shown in the figure below. Looking at these structures, it is noticed that many of these structures contain 3 and 4 member rings.

If it is known ahead of time (like it is here) that the structure of interest does not contain such

rings, it is possible to filter them out with the  button. That dialog box is shown in the figure below. Note that this filter process is also part of the generation options (*Substructures / Ring Rules [91]*), and if selected prior to generation these "rejected" structures will never be shown to the user. But as shown in following figure, by filtering it after generation, it is possible to remove the filter to display all structures again (Reset to all button). This is not possible if the filter is applied during the generation step.

After applying this filter, only 10 structures remain for additional evaluation.

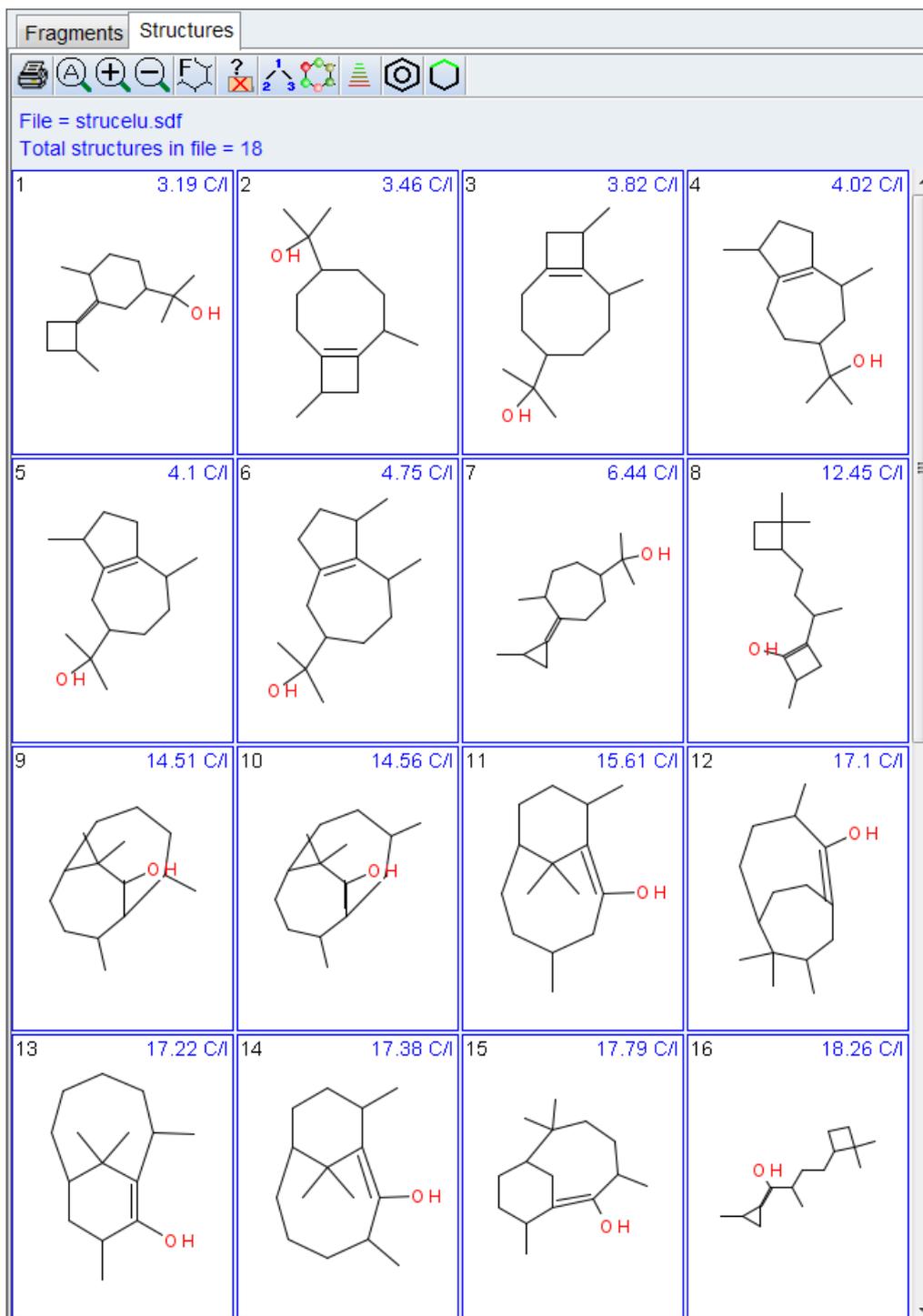


Figure 5.49: A subset of the structures that were generated when the additional constraints were added prior to filtering out the 3 and 4 member rings.

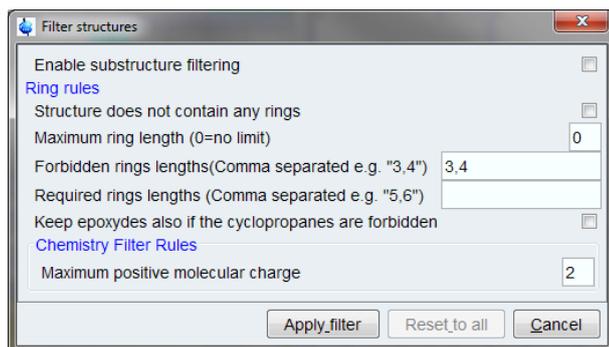


Figure 5.50: Dialog box that can be used to filter the generated structures based on user definitions, in this case excluding structures with 3, 4-member rings.

The next step in the process should now be evaluating the various correlations on the generated structure candidates. Just as in the previous example, evaluating those correlations which were eliminated automatically by the generation algorithm is a valid option.

But also, as mentioned in section [Viewing Correlations on Structures \[110\]](#), any correlations, even those not included in the Correlation Table, can be used to evaluate the proposals. Included with this example data set was an HSQC-TOCSY spectrum, and this is one of the types of spectra that are not used during the generation process, but are supported for this purpose.

Shown in the figure below, is an example of how any peak in any spectrum included with the project can be clicked on and viewed on the proposed structures. In this case a peak from the HSQC TOCSY spectrum is illustrated. It is information like this that helps distinguish the correct structure from the top 3 candidates that all have similar chemical shift prediction values are all under 6. The TOCSY correlation shown highlighted below, while present in the data, is not a long range, and possibly invalid TOCSY correlation for the “best” ranked structure, but makes more sense on the correct structure which is highlighted by the red box.

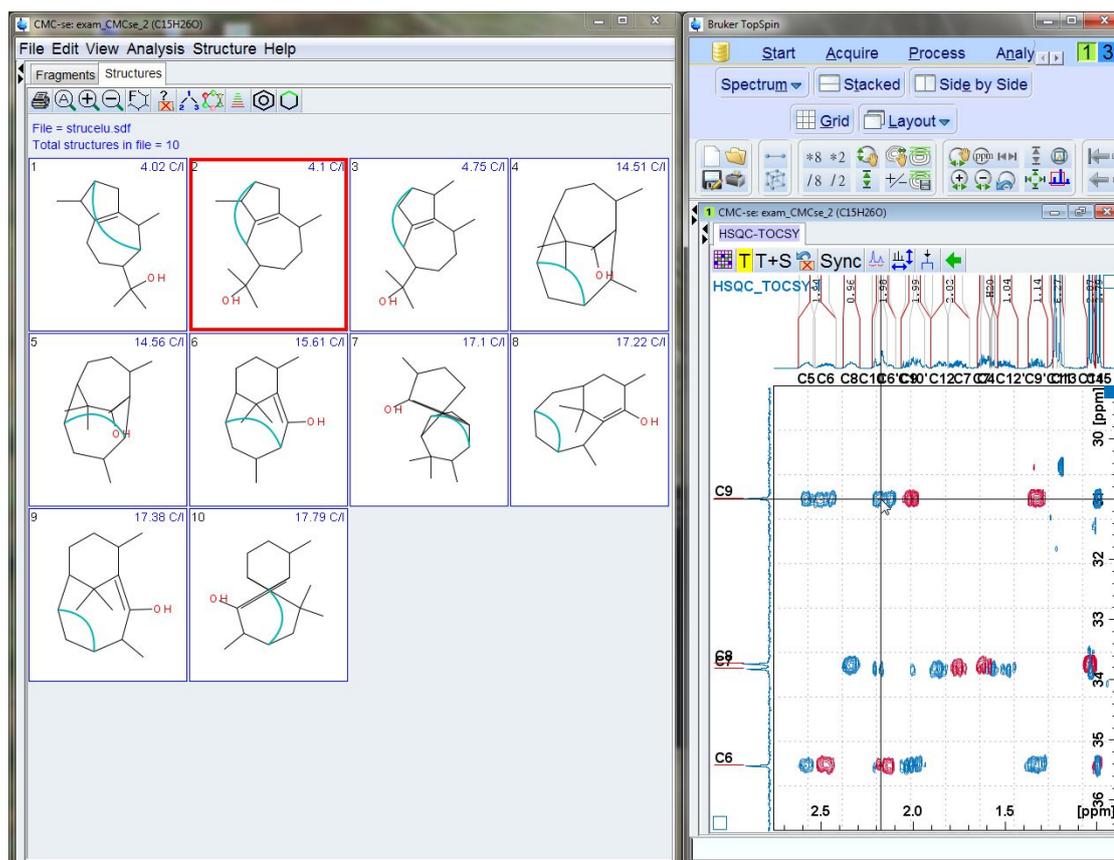


Figure 5.51: Clicking on any peak in any spectrum will result in that correlation being highlighted on the structure candidates for evaluation purposes.

In this case it is known that the correct structure is ranked 2nd among the structure proposals, and while not definitive, this single HSQC TOCSY correlation supports that conclusion as it is much more reasonable in structure #2 than #1 despite the chemical shift prediction results. There are other similar correlations that would rule out structure #3 as well.

In addition to illustrating the tools within CMC-se to modify the Correlation Table when necessary, and examine the various correlations on the structures to provide evidence beyond the chemical shift prediction values for determining the correct solution, this Guaiol data set is an example of why additional resolution is added to the CMC-se parameter sets. The data as provided was only acquired with 256 carbon increments in the HSQC rather than the suggested 400+ in the CMC-se parameter sets.

If additional linear prediction is applied to bring the final data set up to the resolution of the standard CMC-se parameter set, as shown in the figure below, the automatic analysis is able to distinguish the overlapping HSQC peaks that required manual modification above.

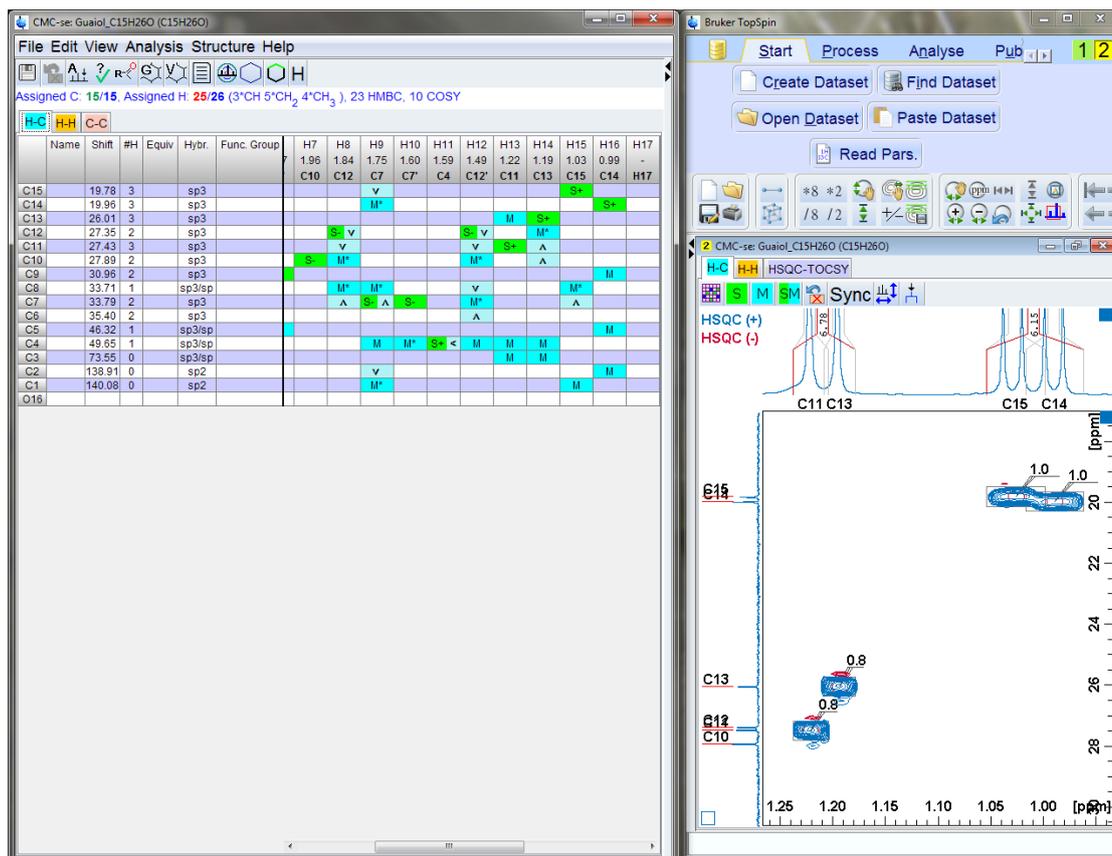


Figure 5.52: After Linear Prediction, the Automatic Analysis is able to distinguish the overlapping HSQC signals that required manual manipulation above.

However, linear prediction beyond the default 2x prediction of points should be used with caution. Shown in the figure below are the two processed spectra from this example. The one as distributed with the standard 2x linear prediction shown on the left, and the additional ^{13}C linear prediction (up to 800 points, then zero filled to 2k as would be in the current CMC-se parameter sets) is shown on the right. Both data sets are cut to the same threshold level, and it is clearly illustrated how linear prediction can increase the apparent resolution of peaks (note the internal ^{13}C projections shown on the sides of the spectra), but also an increase in the noise/artifacts which may be misinterpreted by the automatic analysis. In this particular data set, the S:N of the starting data set was high enough that these artifacts were minimal compared to the signals and thus not a problem.

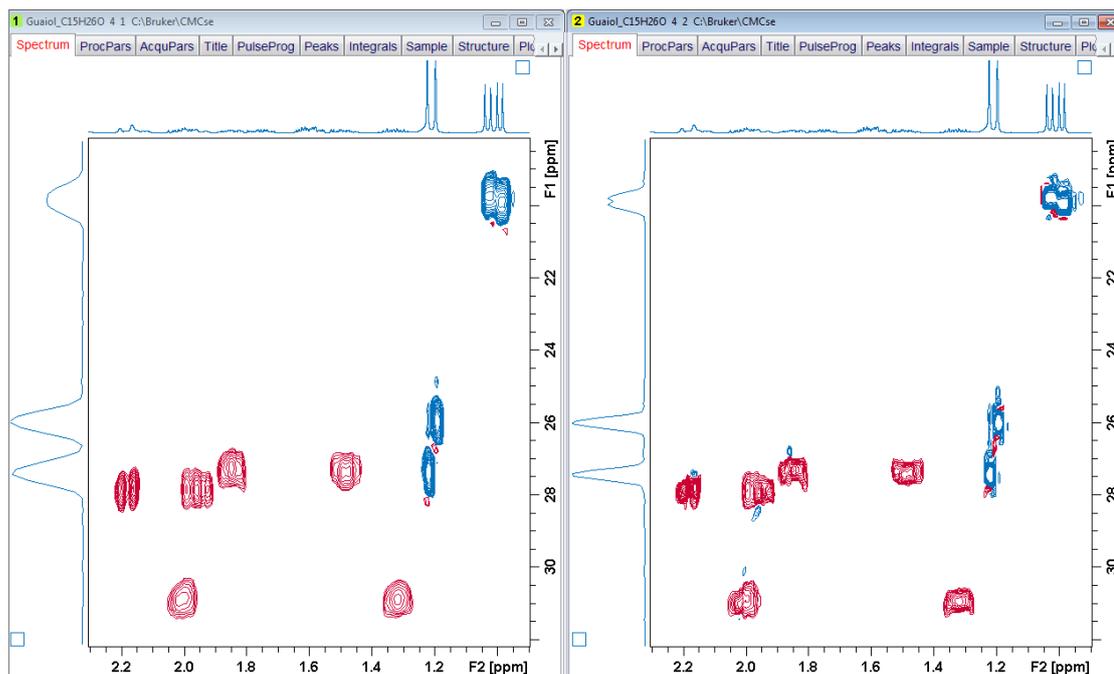


Figure 5.53: Spectra illustrating the increase of artifacts when linear prediction beyond the standard 2x is utilized.

In the figure above, the spectrum with default prediction values is on the left, and the spectrum with additional prediction is on the right.

This example was included for the following reasons:

- The results of the Automatic Analysis had to be corrected manually in HSQC spectrum due to lack of initial resolution in the collected data. Due to the high S:N of the starting spectrum, conservative use of additional linear prediction was able to resolve these peaks.
- Additional constraints were added manually to reduce the number of possible structures.
- The chemical shift prediction routine did not give a clear picture of which structure was correct, and evaluation of individual correlations on the structure was necessary.

This was a good example of how the many tools and features of the CMC-se module can be used to work with non-ideal structure elucidation problems. It was also an illustration on how the quality of the starting data set can affect the amount of manual manipulation necessary to reach a solution.

5.3 Structure Elucidation for Strychnine

The next example is Strychnine (**exam_CMCse_3**). The molecular formula for this compound is $C_{21}H_{22}N_2O_2$, and the sample was prepared in $CDCl_3$.

This is a moderately complex natural product but yet still simple enough to serve as a teaching example. This example will walk through some of the additional aspects of the software which can be used to reduce the number of structures generated, including the use of the INADEQUATE spectrum. While these additional steps may not be necessary with this data, it serves as a nice platform for introducing these concepts.

The following spectra have been acquired and processed properly:

- PROTON [1H]
- C13 [^{13}C]
- COSY [1H , 1H]

- HSQC [^1H , ^{13}C]
- HMBC [^1H , ^{13}C]
- HMBC [^1H , ^{15}N]
- INADEQUATE [^{13}C , ^{13}C]

Again a new structure elucidation project including all spectra was created (section [Creating a New CMC-se Project \[23\]](#)) and all data that can be automatically analyzed (all, except the INADEQUATE) was resulting the Correlation Table as shown in the figure below. For illustration purposes, the layout has been toggled back to the old style so that all correlations could be displayed in a single screen capture.

Assigned C: 21/21, Assigned H: 22/22 ($10^{\circ}\text{CH}_6^{\circ}\text{CH}_2$), 99 HMBC, 18 COSY

Name	Shift	#H	Equiv	Hybr.	Func. Group	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	
C21	26.87	2		sp3																								
C20	31.62	1		sp3/sp						M	M			M	M							M						
C19	42.48	2		sp3																								
C18	42.89	2		sp3																								
C17	48.23	1		sp3/sp																								
C16	50.35	2		sp3																								
C15	51.95	0		sp3/sp																								
C14	52.69	2		sp3																								
C13	60.12	1		sp3/sp																								
C12	60.20	1		sp3/sp																								
C11	64.61	2		sp3																								
C10	77.60	1																										
C9	116.20	1		sp2																								
C8	122.26	1		sp2																								
C7	124.18	1		sp2																								
C6	127.15	1		sp2																								
C5	128.51	1		sp2																								
C4	132.82	0		sp2																								
C3	140.61	0		sp2																								
C2	142.24	0		sp2																								
C1	169.27	0		sp2																								
N22	153.44																											
N23	39.07																											
O24																												

Name	Shift	Mult	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22
H22	H22																							
H21	C17	1.25																						
H20	C21	1.42																						
H19	C16	1.86																						
H18	C21	2.33																						
H17	C19	2.65																						
H16	C14	2.69																						
H15	C16	2.84																						
H14	C19	3.10																						
H13	C20	3.12																						
H12	C16	3.16																						
H11	C14	3.68																						
H10	C13	3.84																						
H9	C12	3.91																						
H8	C11	4.05																						
H7	C11	4.11																						
H6	C10	4.26																						
H5	C6	5.86																						
H4	C7	7.07																						
H3	C8	7.14																						
H2	C5	7.23																						
H1	C9	8.07																						

Figure 5.54: Starting Correlation Table after the Automatic Analysis.

As seen, there are no immediate problems with the analysis as there was with the previous example. And a quick inspection of the Combined Spectra Display, the picked peaks all look reasonable. Therefore it is possible to start the structure generation without modification to the table.

Because there are over 100 correlations in the table, it is likely (though not always) that a somewhat large number of correlations will need to be automatically eliminated. Starting with a value of 5 results in 0 structures; increasing this to 10 still results in 0 structures.

At this point, it is a good idea to take a closer look at the Combined Spectra Display to make sure there are no errors that are causing 0 structures to be generated.

One place to start looking is where atoms are overlapped or barely resolved. In this case, Carbons 12 and 13 are less than 0.1 ppm apart. Looking at the Combined Spectra Display arranged so that only the HSQC is showing (see [Adding Additional Information to the Correlation Table \[76\]](#) for options to manipulate the view of the Combined Spectra Display), an experienced spectroscopist would likely be able to notice that the carbon assignments for these 2 HSQC peaks is incorrect. As shown in the figure below, the software has assigned the HSQC correlation from the downfield ^1H to C13, when this chemical shift, is actually more in the center of the other HSQC correlation. These 2 peaks should be corrected.

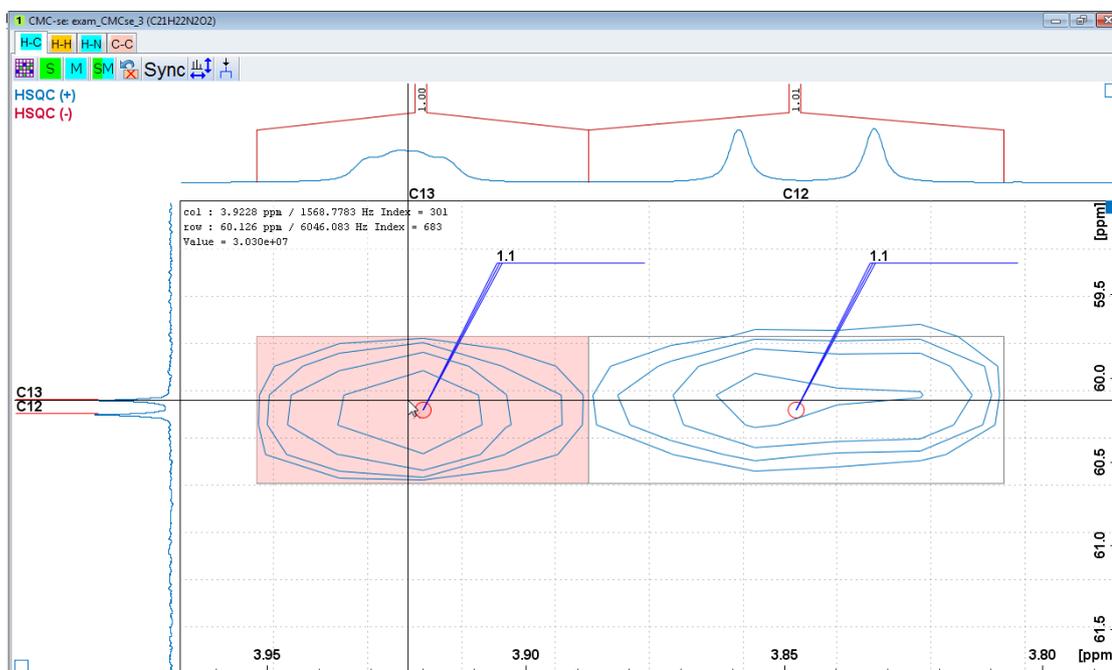


Figure 5.55: Closer inspection reveals a miss-assignment of the carbon resonances for 2 of the HSQC correlations.

Because the Automatic Analysis has assigned these in this manner, it is likely that deleting and re-picking the correlations through the Combined Spectra Display would result in the same situation. It is probably easier to simply make the change through the Correlation table by right clicking in the appropriate cell, and deleting/adding the necessary correlation (see section [Adding Additional Information to the Correlation Table \[76\]](#), and the figures below).

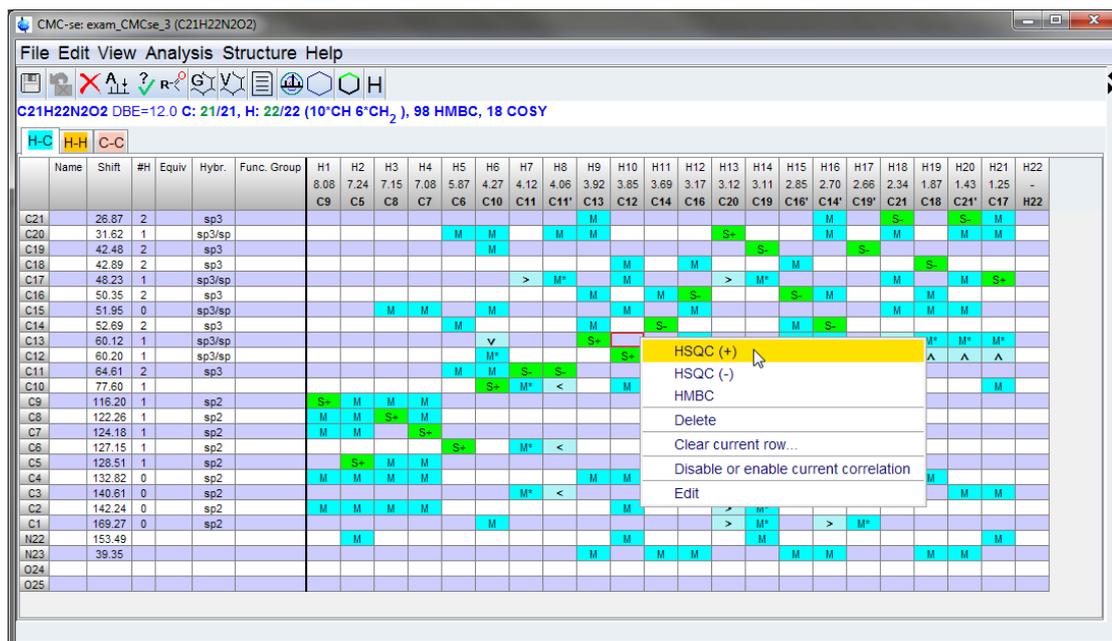


Figure 5.56: Manually changing the HSQC correlations for C12 and C13.

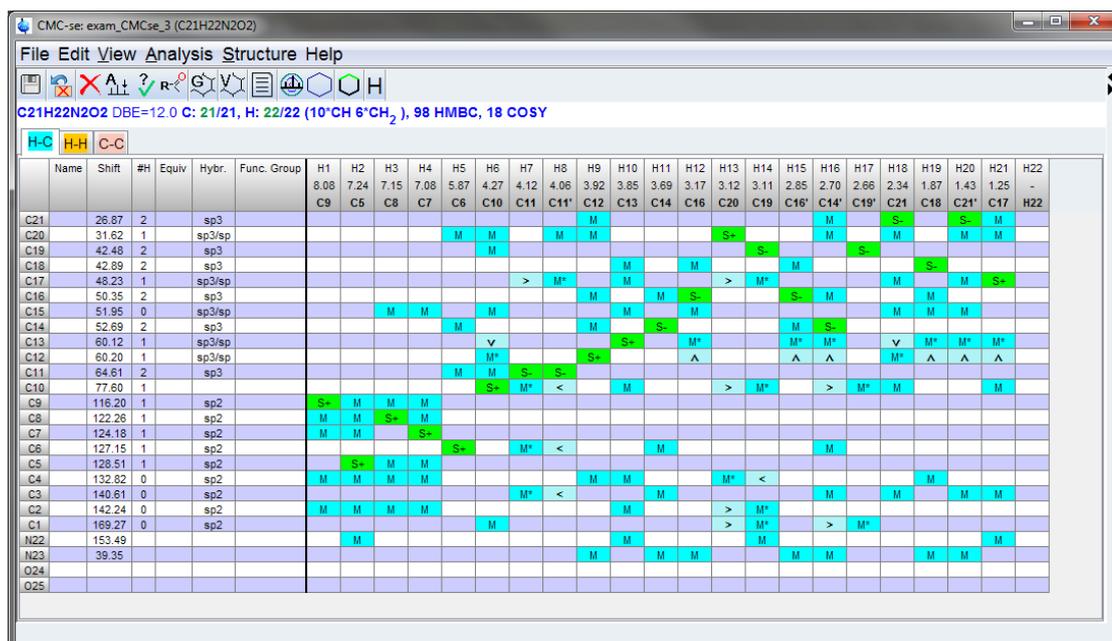


Figure 5.57: Manually changing the HSQC correlations for C12 and C13.

Further inspection of the data reveals that there are a number of HMBC correlations that were not picked and included by the Automatic Analysis.

However, manually picking them as shown in the figure below (dotted red circles to indicate a manually picked correlation) it is apparent that these correlations are an order of magnitude less intense than the other correlations for the corresponding protons. **Therefore they are likely long-range, and should not be included at this time.**



In some rare cases, when all other attempts to generate structures fail, it might be beneficial to come back and add correlations like these, but this is not usually the best option, especially before other things (as will be discussed below) have not been tried.



Figure 5.58: Manually peak picking weaker HMBC correlations that were not originally included by the Automatic Analysis.

Because these correlations are so much weaker than the others for the corresponding protons, this is NOT usually the recommended solution.

Because there are no other changes that appear to be necessary to the Correlation table, it is possible to re-start the structure generation. However, even after making this change to the HSQC correlations of C12/C13, running the structure generation with **Number of Correlations to Auto Eliminate** set to 10 results in 0 structures being generated.

Another option that could be explored is that this might be a molecule with charged atoms. Because the Double Bond Equivalent value is not highlighted in red, it is not likely that this is a charged molecule, or a "salt". But it is possible, especially with all of the Nitrogen and Oxygen atoms that there are internal charged atoms.

If the "Chemistry Rules" section of the Structure Generation Options is expanded, it will be seen that the default of the generator is to not allow charged atoms or a molecular charge (see the figure below).

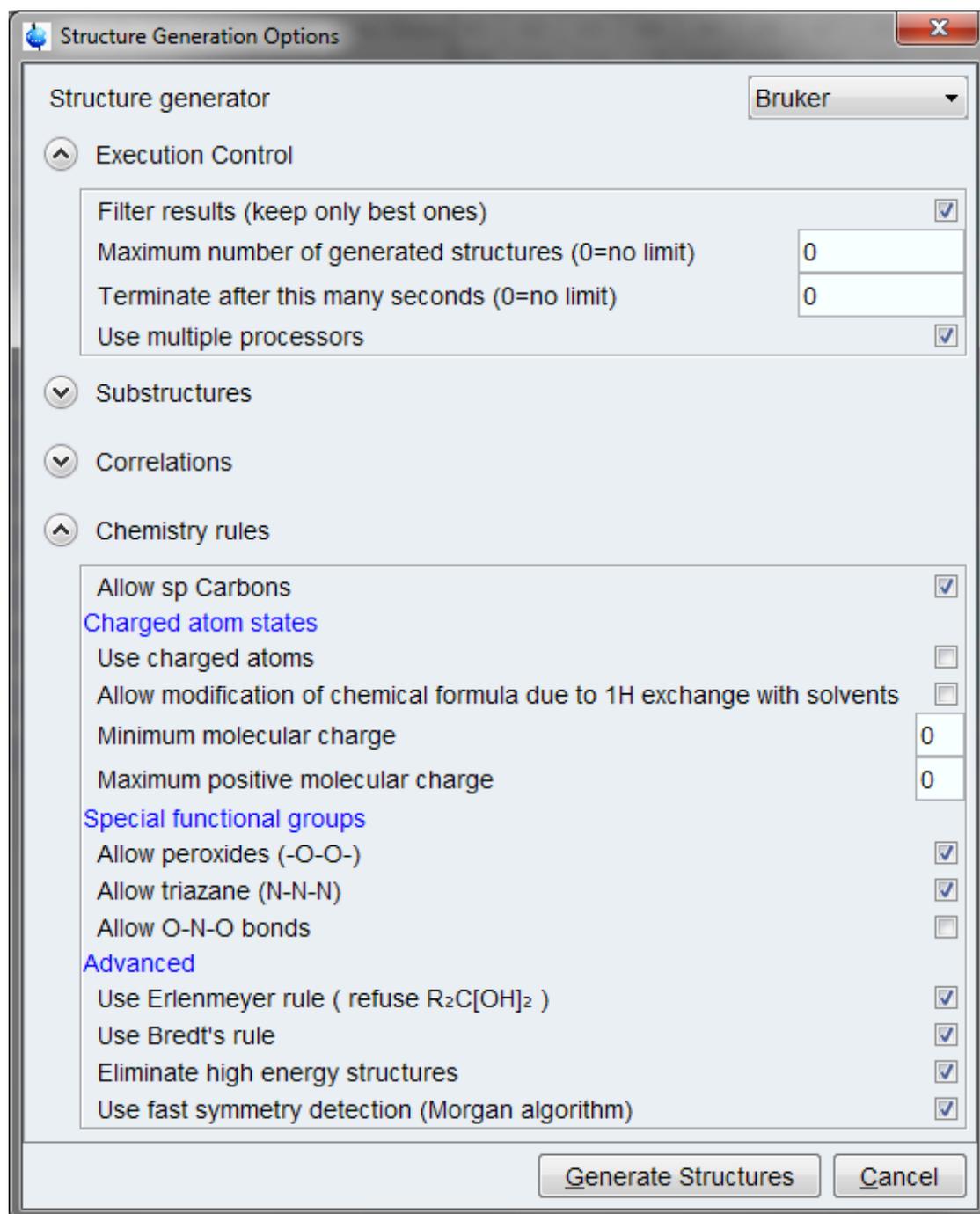


Figure 5.59: The “Charged atom States” section of the “Chemistry Rules” by default does not allow charged atoms or molecules.

In cases where no structures are being generated, the option shown in the figure above can be changed.

Even with checking the *Use charged atoms* option, and setting the **Maximum and Minimum molecular Charge** values to 0 (there is no indication that this is molecule with an overall charge), while taking longer than previous runs, still only auto-eliminating 10 correlations results in 0 structures.

It should be noted that using the charged atoms or molecules, will increase the time of the structure generation process. This was one of the reasons it was moved to an option rather than default.

Examples

If the user wants to save time during this example, they could uncheck the option, as it is known that Strychnine does not have any charged atoms. **However, the figures in the remainder of the chapter will have it left on.**

At this point, the next step should be to increase even further the number of correlations to be auto-eliminated. When this is increased to 15, multiple structure candidates are generated, a subset of which are shown in figure below.

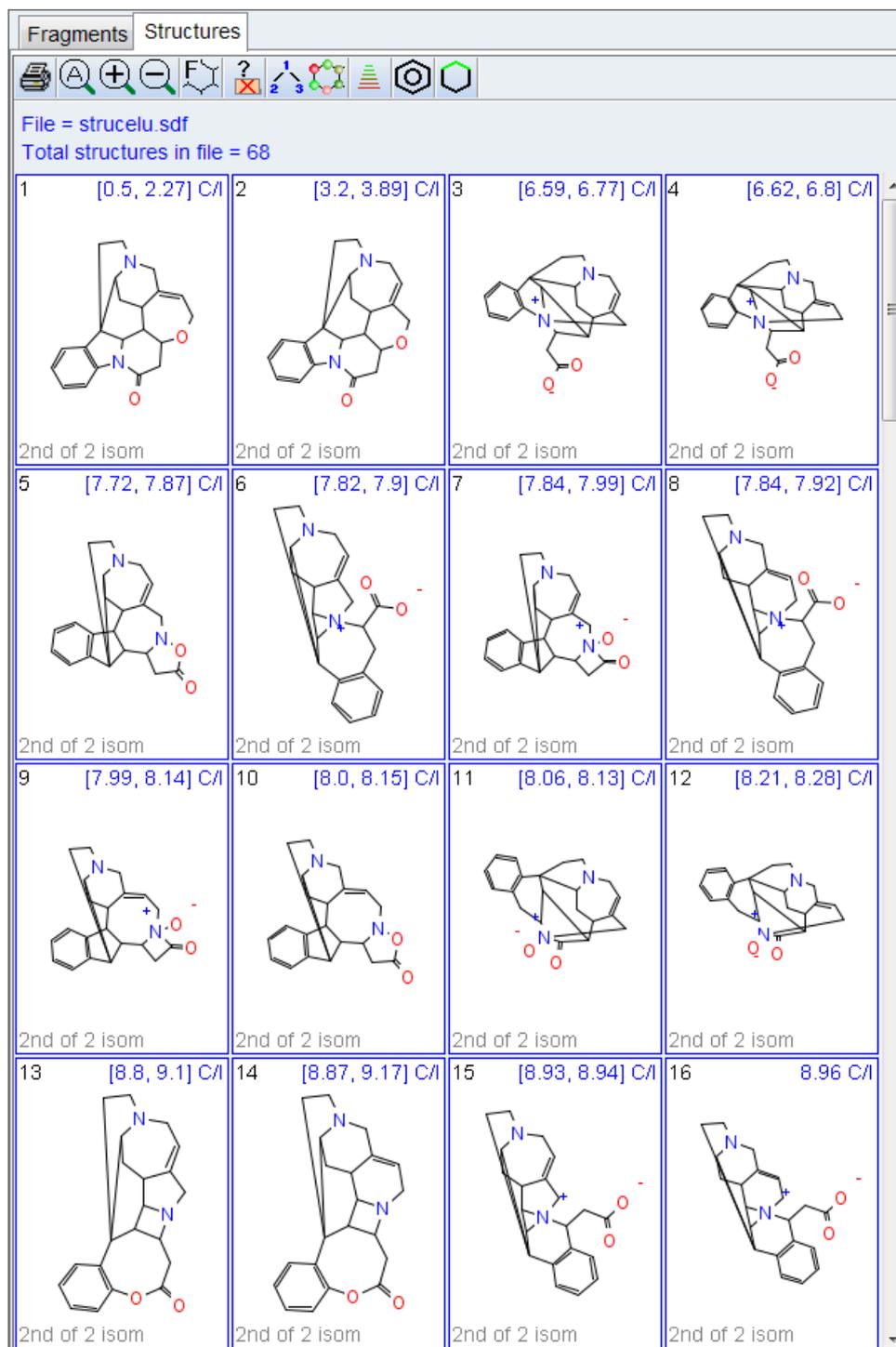


Figure 5.60: The 68 structures that are generated with the above fixed Correlation Table, the auto elimination value set to 15, and allowing charged atoms.

As mentioned previously in this manual, including the previous examples, the next step would usually be to add additional freedom in the auto-elimination of correlations. However in the interest of brevity for this section of the manual that will be skipped this time around.

From the structures above, there are a couple reasonable candidates shown. Any of the previous methods of verifying the correct structure (section [User Validation of Generated Structure Proposals \[101\]](#)) are still reasonable at this point. However, because this example data set also includes the INADEQUATE spectrum, it can be added to the structure constraints to even further reduce the number of possible structures.

As mentioned previously, the INADEQUATE spectrum must be manually peak picked and added to the Correlation Table. Instructions for this process are in section [Adding INADEQUATE Correlations \[68\]](#).

With careful peak picking, especially where neighboring carbons are close (especially C12 & C13 as well as C18 & C19) the C-C tab of the correlation table is updated and should be like that shown in the figure below.

Name	Shift	#H	Equiv	Hybr.	Func. Group	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	N22	N23	N24	N25
C21	26.67	2		sp3																										
C20	31.62	1		sp3/sp																										
C19	42.48	2		sp3																										
C18	42.89	2		sp3																										
C17	48.23	1		sp3/sp																										
C16	50.35	2		sp3																										
C15	51.95	0		sp3/sp																										
C14	52.69	2		sp3																										
C13	60.12	1		sp3/sp																										
C12	60.20	1		sp3/sp																										
C11	64.61	2		sp3																										
C10	77.60	1		sp2																										
C9	116.20	1		sp2																										
C8	122.26	1		sp2																										
C7	124.18	1		sp2																										
C6	127.15	1		sp2																										
C5	128.51	1		sp2																										
C4	132.82	0		sp2																										
C3	140.61	0		sp2																										
C2	142.24	0		sp2																										
C1	169.27	0		sp2																										
N22	153.49																													
N23	39.35																													
O24																														
O25																														

Figure 5.61: The C-C tab of the Correlation Table after careful manual peak picking of the INADEQUATE spectrum.

Using the same generation options as previous (Auto-eliminate 15 correlations), and allow charged atoms) the number of structures has now been reduced from 68 to 16 candidates. These are shown in the figure below.

When examining the structures as those below, it should be kept in mind again, that the CMC-se module does not take into account 3D geometry, and these are all flattened structures, so just because some look a little different, they should not be completely discounted without additional reasons.

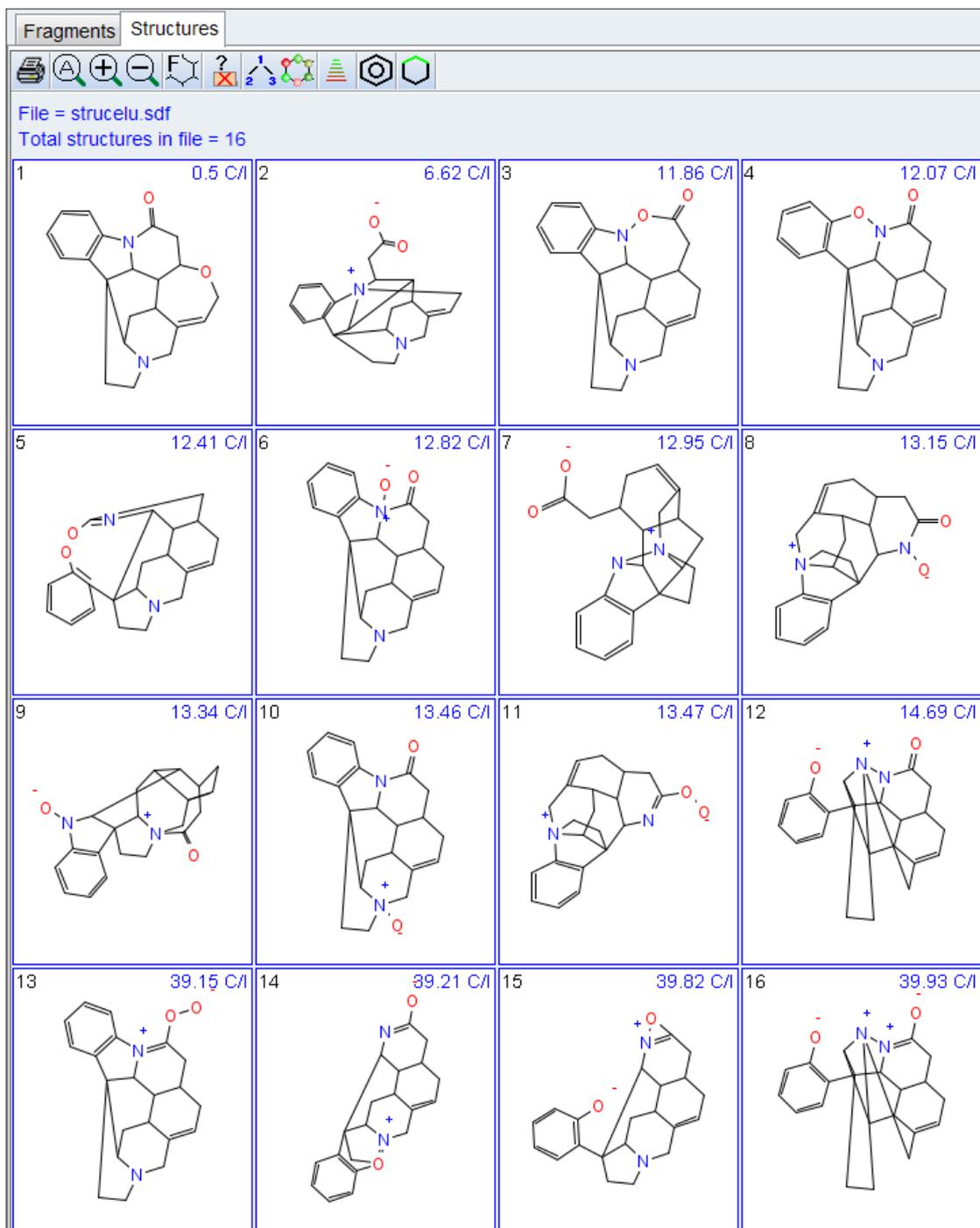


Figure 5.62: The 16 structures that are now generated with the inclusion of the INADEQUATE for additional constraints.

From these newly generated structures compared to the initial set, it is a bit more obvious based on chemical intuition and chemical shift prediction values which of these structures is correct.

Again, the previous methods of viewing the correlations on the structures can be used to confirm this structure as being correct.

The fragment tool that was used in section [Manual Structure Elucidation of \$\alpha\$ -lonon using the Fragment Tool \[143\]](#), may also be used in cases where there is difficulty generating a reasonable number of structures to evaluate. It is possible to instead of generating the entire structure, to generate small fragments which serve as building blocks for the computer generation algorithms.

If the user were to examine the Combined Spectra Display as in the figure below, it might be obvious to an experienced spectroscopist that this type of pattern of 2 doublets, and 2 triplets in the aromatic region of the spectrum would correspond to a 1,2-substituted benzene ring. Using the Fragment Tool of CMC-se (section [The Fragment Tool \[81\]](#)), this can not only be defined as a required structural element, but also assignments can be made to this fragment, so that it serves as a structural building block thus reducing the number of structures, and time required to reach that solution.

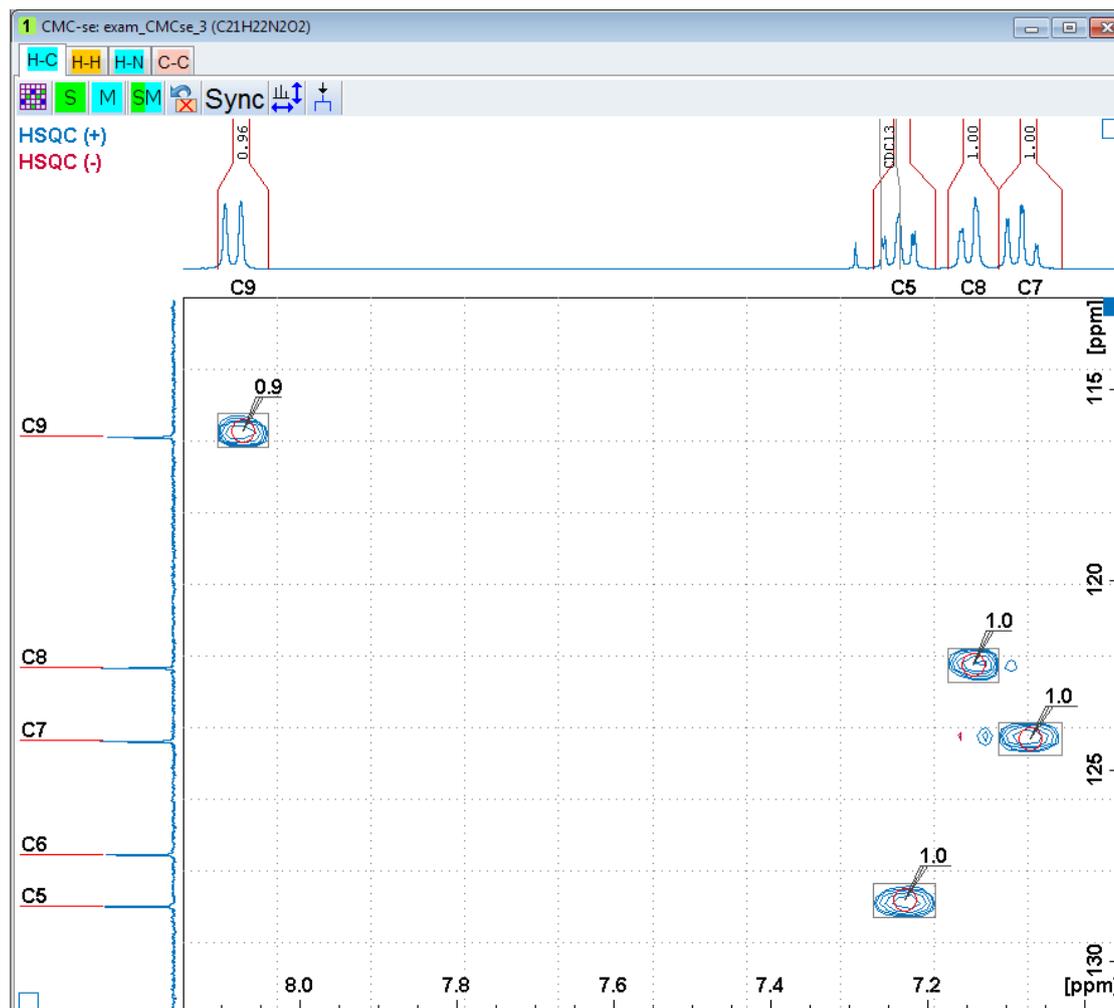


Figure 5.63: Aromatic region of the HSQC spectrum in the Combined Spectra Display window.

Prior to starting this process in this example, it will simplify things for sake of explanation, if

the user deletes the INADEQUATE correlations. This is done with the  button in the main section of the Correlation table. From the resulting Pop-Up window, the user can define which of the “non-required” data sets to delete from the Project. As shown in the following figure, there is also the option to delete *All Correlations*, but in this example, just the *INADEQUATE* correlations will be removed.

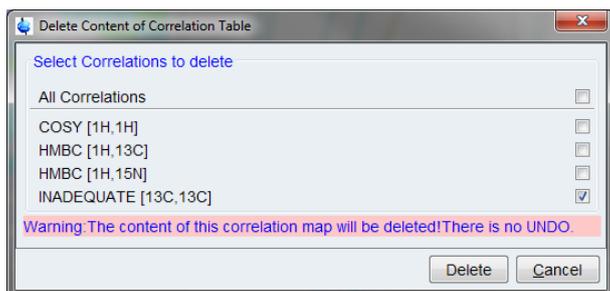


Figure 5.64: Pop-Up Dialogue box to allow the user to delete non-required, or all of the content from the Correlation Table.

As described in section [The Fragment Tool \[81\]](#), the process to start defining a required fragment begins in the *Fragment tool* window. This can be opened by pressing the  button from either the main section of the Correlation Table. In this window it is possible to draw fragments. But in this case, the desired fragment is already in the library and simply right clicking on the button  opens the *Fragment selection* window. Defining it as **Require** as shown in the figure below, will add this fragment to the Fragments Tab, where there are graphical tools to facilitate assigning the atoms of this fragment.

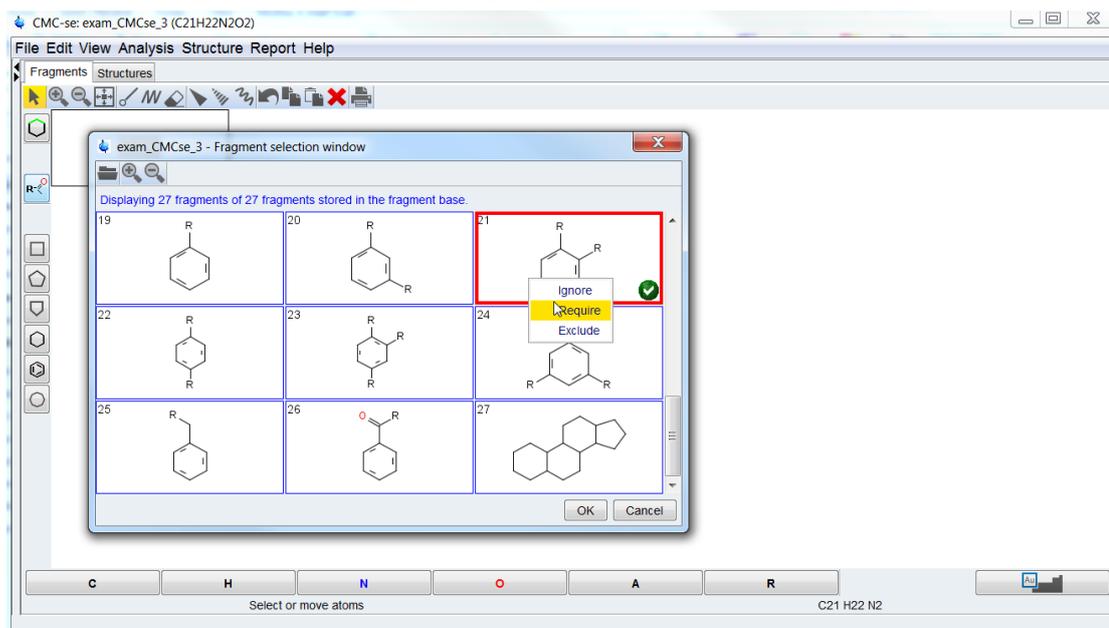


Figure 5.65: Opening the Fragment Management tool to define the required fragment for this structure generation.

Once the fragment(s) have been defined as required they show up in the Fragments Tab as shown in the following figure. Also added to this window now are colored spheres corresponding to all of the Carbon and hetero atoms in the table. They are color coded based on hybridization, and have small bullets to indicate the number of attached protons.

As illustrated below, these spheres can simply be dragged onto the fragment, and the various correlations can also be highlighted on the fragment, thus simplifying the assignment process.

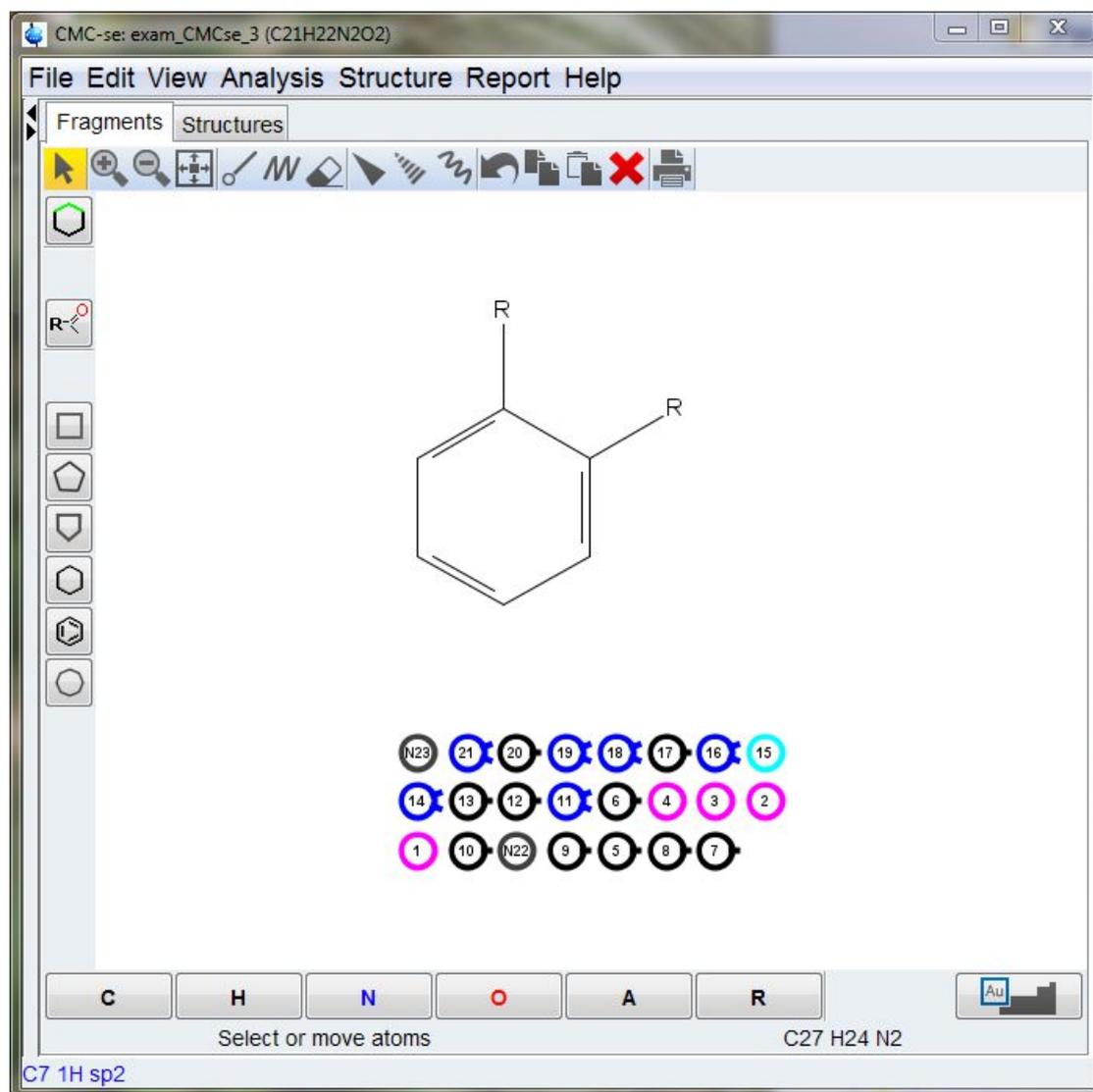


Figure 5.66: A required fragment has been added to the Fragments Tab and is ready for atom assignments if available.

Looking at the spectra in the figure *Aromatic region of the HSQC spectrum in the Combined Spectra Display window* above, two of the protons are doublets which indicates that they are next to a single protonated carbon. It doesn't matter which one is which at this point, but either the sphere belonging to C9 or C8 can be clicked on and dragged to either of these positions on the fragment as shown in the figure below.

Once the atom is assigned on the fragment, clicking on it will result in all of the correlations for that atom being illustrated. From here, it is easy to see that the next carbon in the ring should be C5 due to the COSY correlation (brown line) that is observed between these 2 atoms.

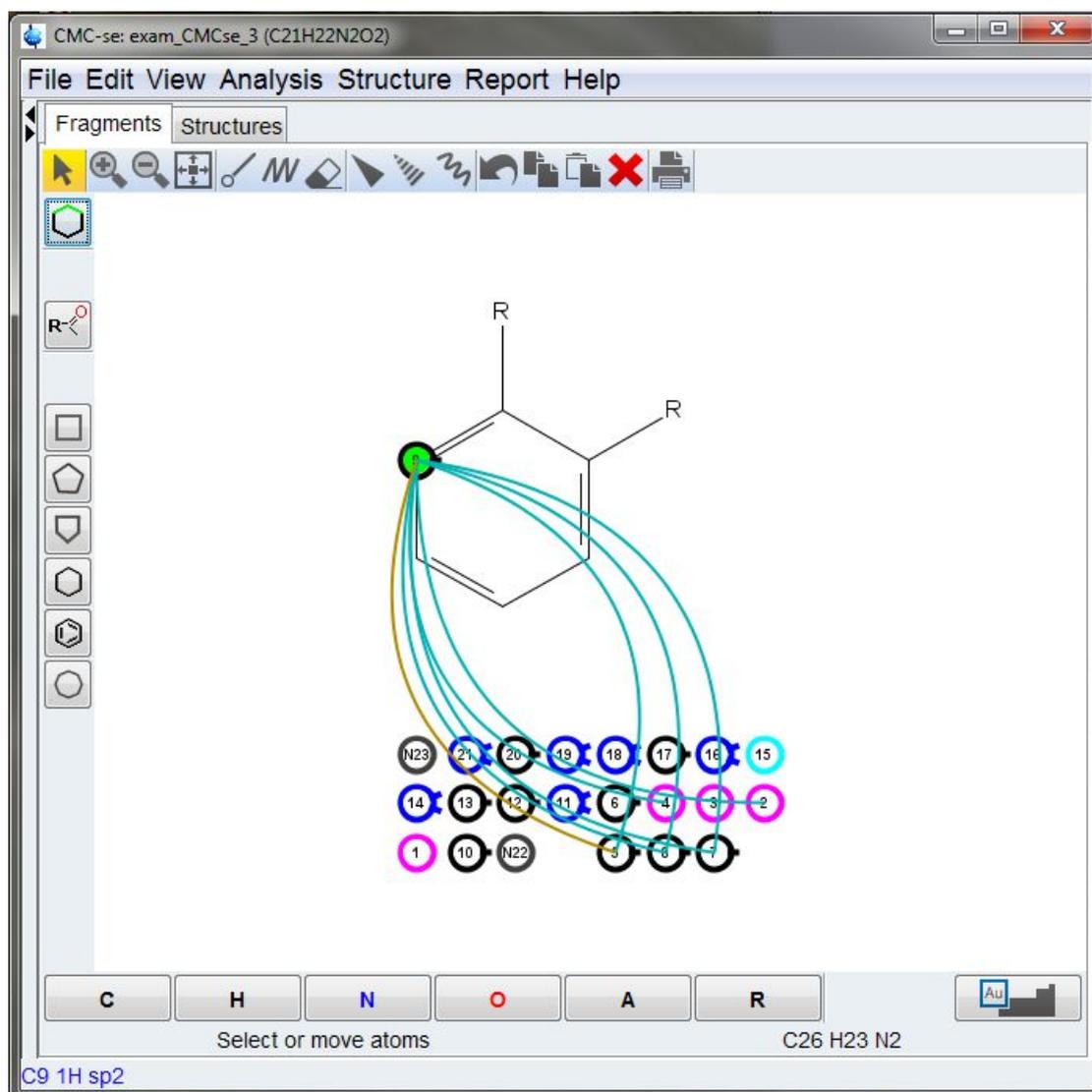


Figure 5.67: Once an atom is assigned to a position in the fragment by clicking and dragging the sphere accordingly, the experimental correlations to that atom can also be viewed.

As shown in the figure below, these correlations are not only shown between various spheres in the heading, but when multiple atoms are assigned to the fragment, they will be indicated on the fragment as well; thus providing more confidence to the assignments as they progress.

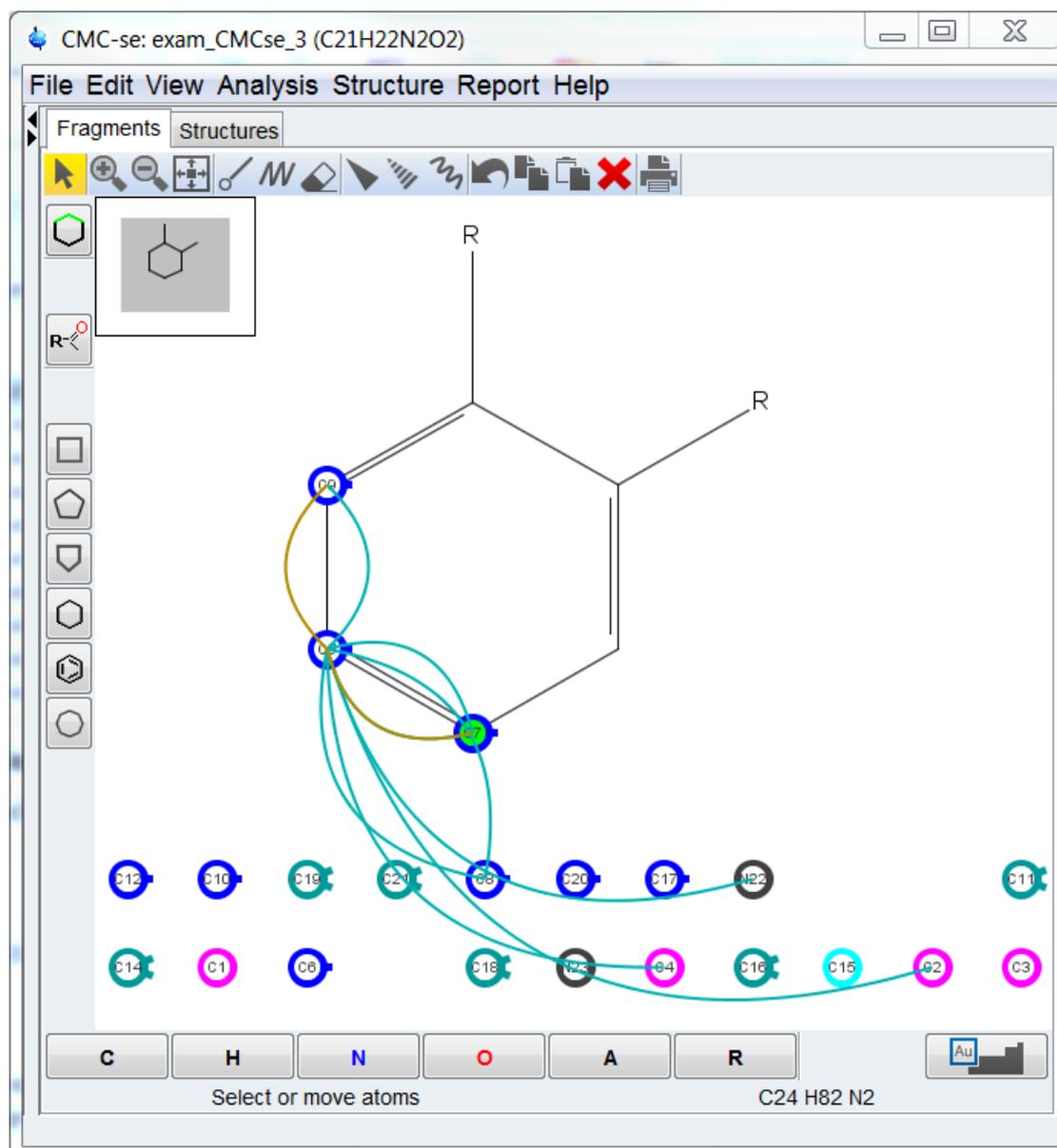


Figure 5.68: The correlations are also shown on the fragment when more than one atom is assigned.

As illustrated above, it might be a case where the next atom in the ring is missing the COSY correlation, making it difficult to know which one should be next in the ring. While in this case it is relatively easy because there is only one atom left from the 4 protons that had been identified as being in this ring. But as shown in the following figure, it is also possible to examine the spectra and visualize the correlations on the fragment even if they are not included in the table.

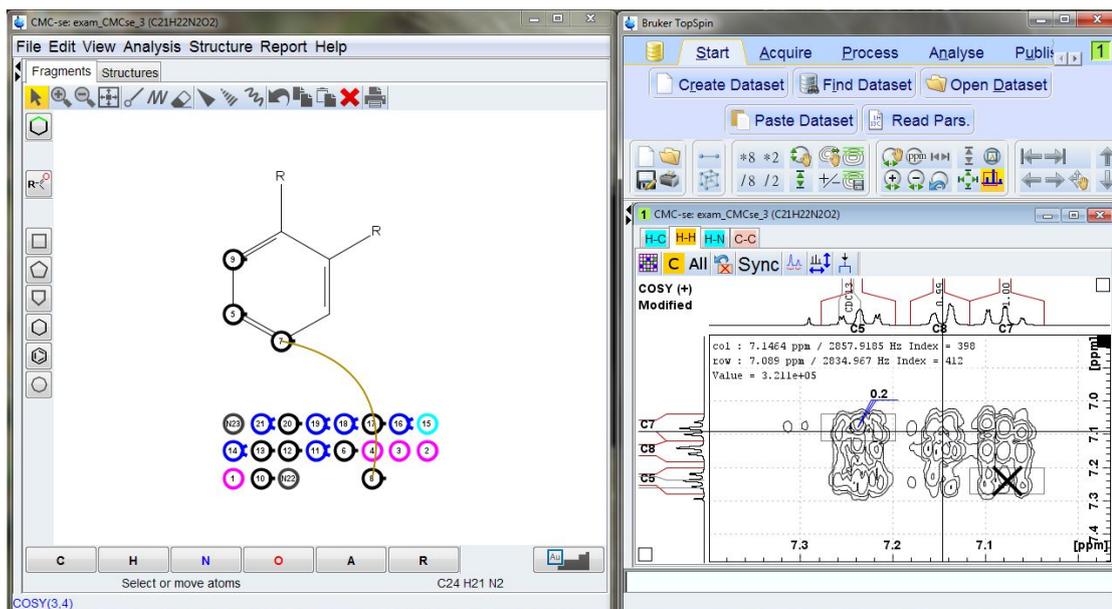


Figure 5.69: Even correlations not included in the table can be visualized on the fragments.

Also, as described in section [The Fragment Tool \[81\]](#), with the new Fragment Tool, additional atoms can be added to these assigned fragments. For example in the figure below, there is an HMBC from C5 to N22. Therefore, one could conclude that as long as this is a real correlation and not a mistake, one of the R groups, must be a nitrogen and this is a long range correlation.

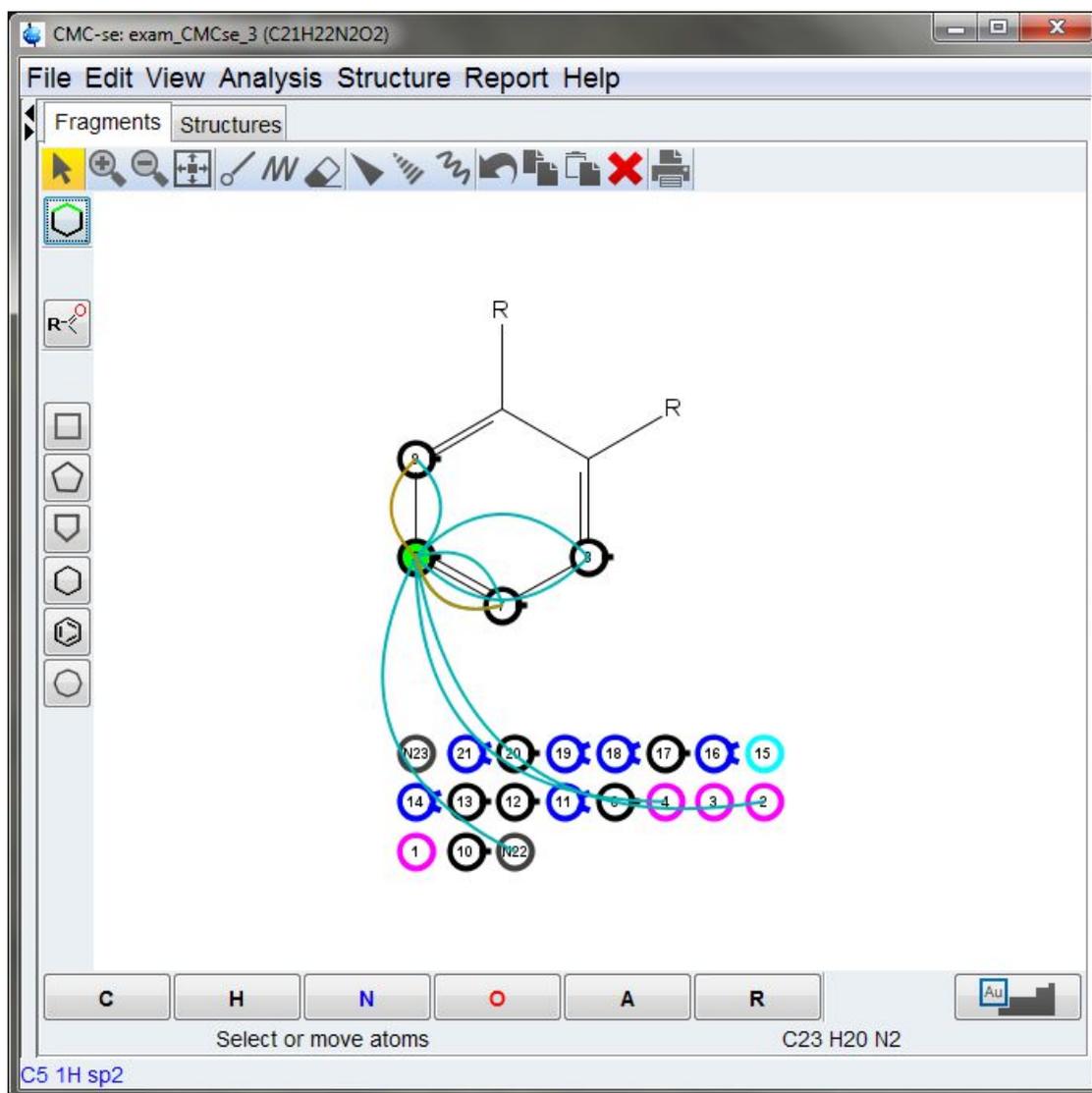


Figure 5.70: There is an HMBC from C5 not only to the 4^o carbons in this ring, but also to N22

Using the Fragment Tool, this current fragment can be expanded to include these additional atoms.

In the figure below, one of the R groups was replaced by the nitrogen.

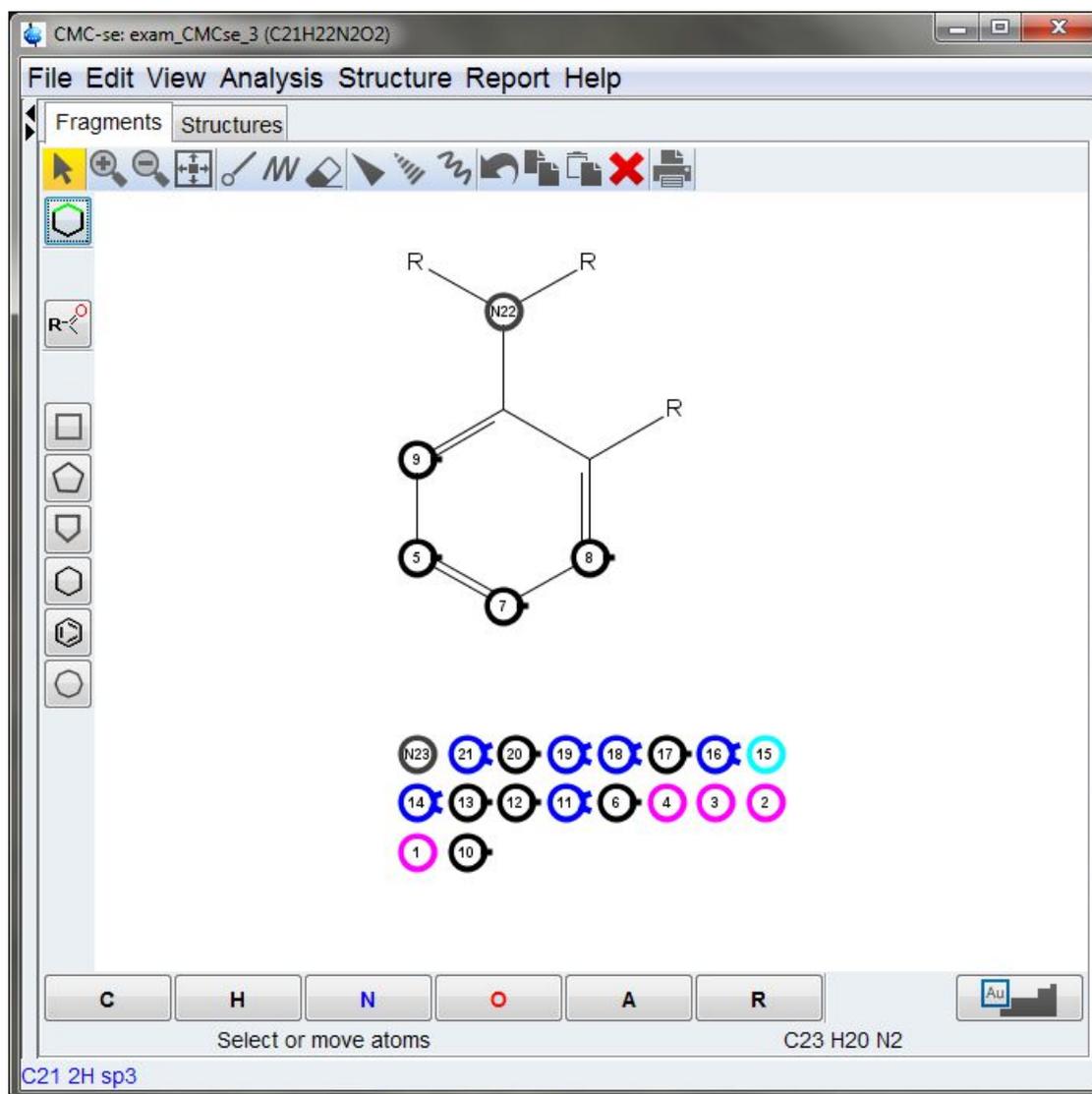


Figure 5.71: Using the Fragment Tool to modify existing assigned fragments.

Another place that constraints can be added to the generation process is to define a simple functional group for a given atom instead of an entire fragment (section [Adding Additional Information to the Correlation Table \[76\]](#)). This process along with some of the book keeping features (section [Bookkeeping Tools within the CMC-se Module \[78\]](#)) of the CMC-se software, are illustrated below.

As shown before, double clicking on any HSQC peak in the Combined Spectra Display will bring up the *Atom Properties* dialog box for that atom.

From this box, it is possible to designate a functional group for this atom. Note that the pre-defined pull down list guides the user to possible functional groups based on chemical shift.

Also in this box, as indicated in the following figure, it is possible for the user to add a comment. While this can be anything, in this case it might be beneficial to use this as a way of reminding the user or some other reviewer of the data why this was defined the way it was.

When a functional group is defined in this manner, as shown in the figure below, it is automatically added to the Fragments Tab so it can be treated just like other fragments, including adding additional assignments, and modification as was demonstrated previously with the ring.

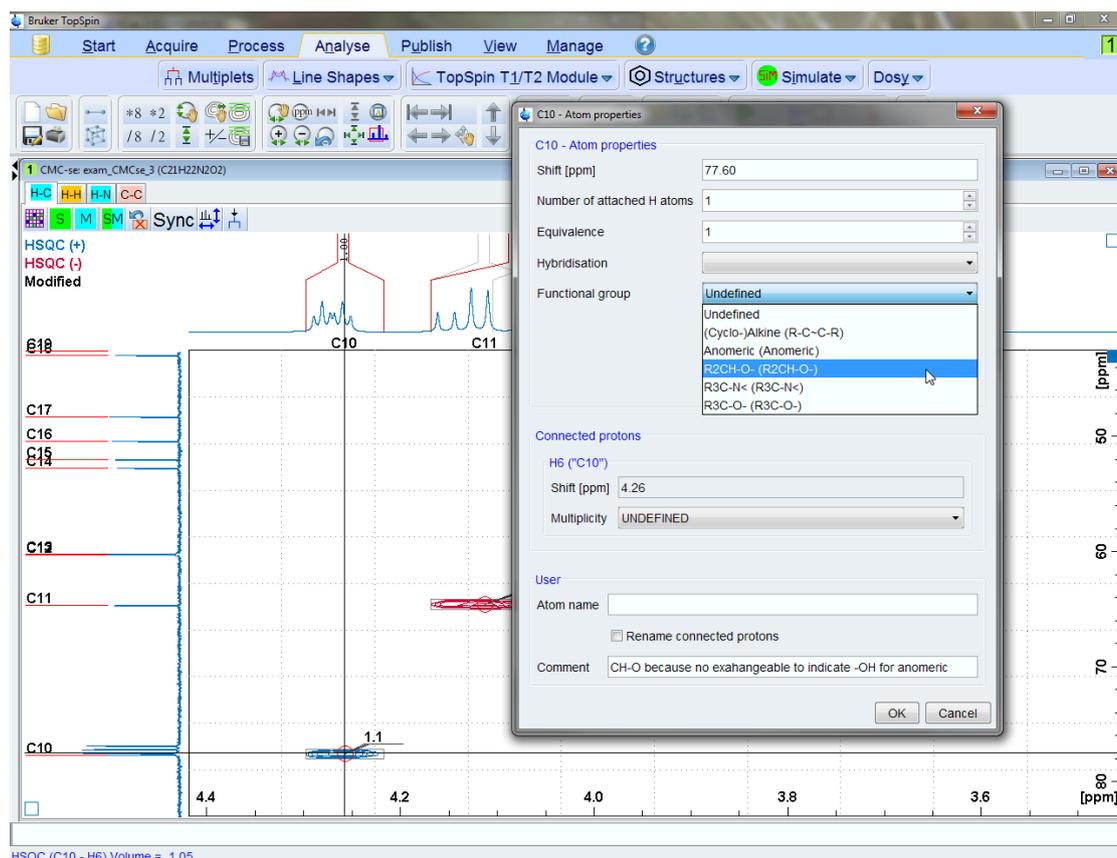


Figure 5.72: Double clicking on any HSQC correlation in the Combined Spectra Display will bring up the "Atom Properties" dialog box where the user can define things like Functional group, and even add a comment.

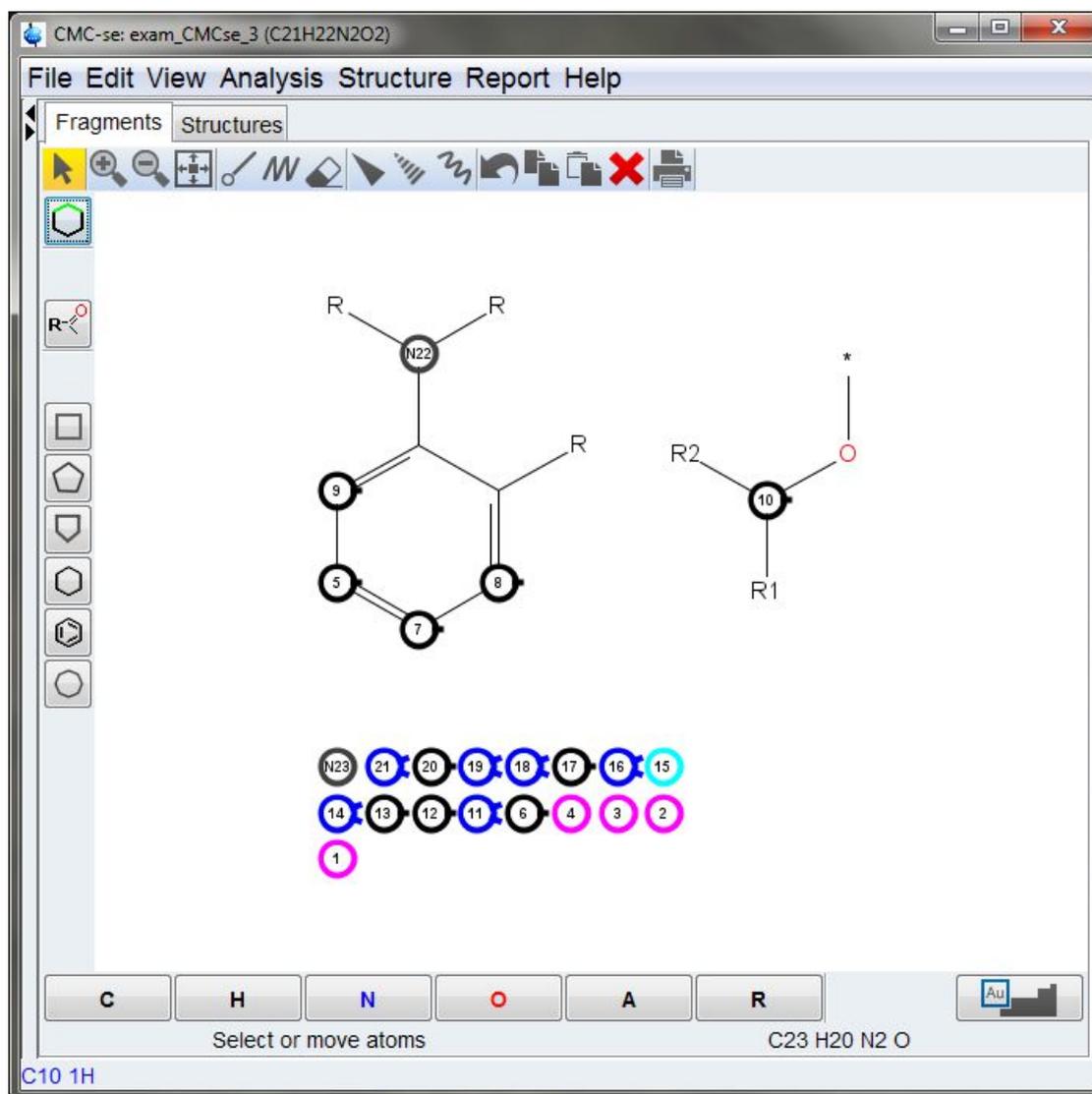


Figure 5.73: Once a functional group is defined, it is added to the Fragments Tab, for additional assignments, modifications as desired by the user.

In addition to adding a user comment in the above Atom Properties box, it is also possible to right click on any correlation (including HMBC, COSY etc.) and select the properties option and bring up the properties for that correlation. In this box, it is possible to define an annotation which can be displayed on the spectrum as shown in the figure below.

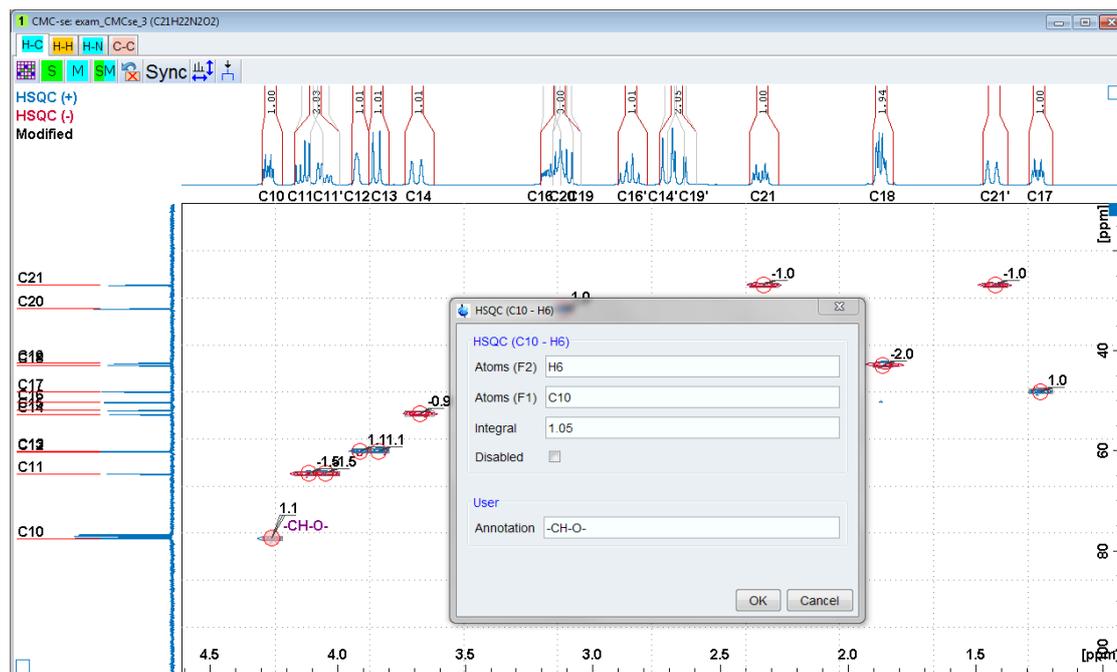


Figure 5.74: Right clicking on any correlation will bring up a properties box in which an annotation can be entered and then shown on the spectrum in the Combined Spectra Display.

With the above additions to the Correlation Table, the assigned fragment, and the functional group definition. The structure generation can be started again. This time, as shown in the figure below, only 2 structures are generated.

It should be pointed out that the same Generations Options were used in this generation, as were used previously prior to the addition of the INADEQUATE data. And how adding these few fragment assignments, improves not only those initial results, but is also improved over those results from the INADEQUATE.

In much less time than it took to acquire the INADEQUATE spectrum, a little bit of manual work with the data resulted in a few additional constraints that provided better results than the inclusion of the INADEQUATE spectrum.

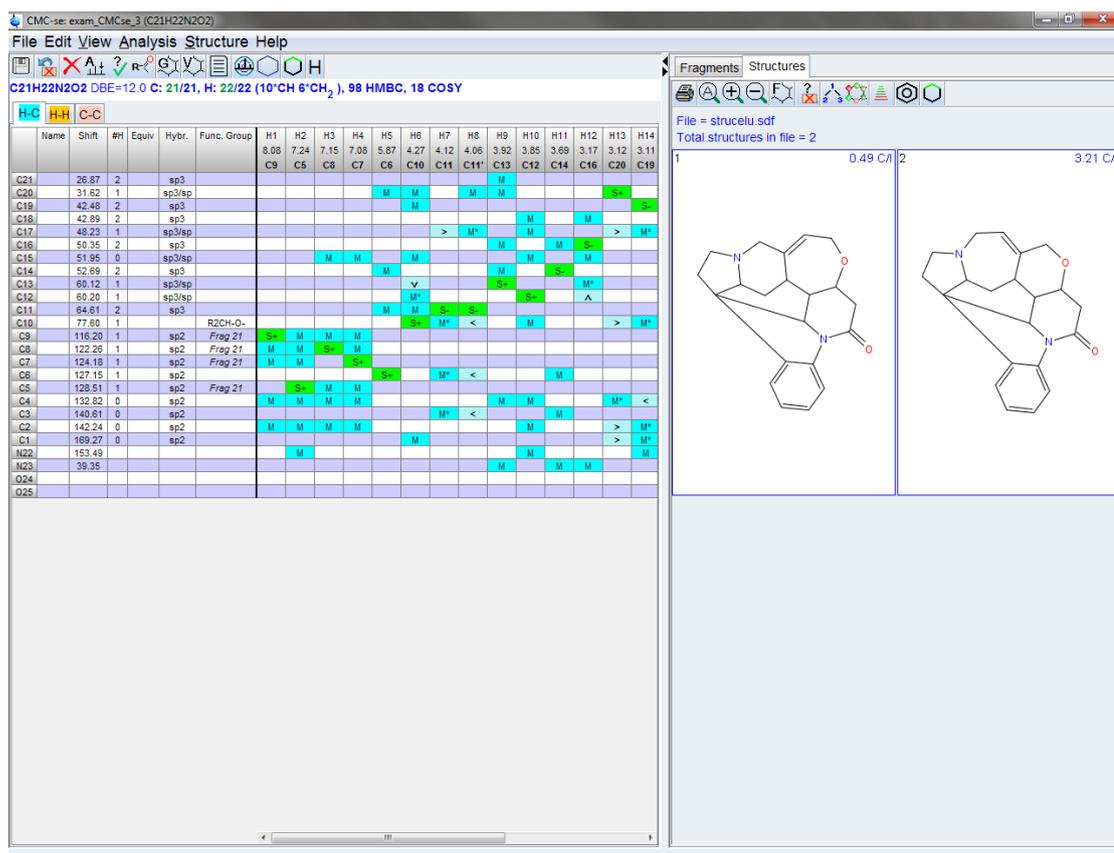


Figure 5.75: Resulting 2 structures that were generated after the deletion of the INADEQUATE correlations, but addition of the assigned fragments.

While the lowest ranked structure is known to be correct in this case, as has been shown in the last example, that is not always the case, and additional verification (possibly through additional spectra such as TOCSY, HSQC-TOCSY etc.) should be done by the user if this were a true unknown molecule.

This example was included in the manual and example data for multiple reasons:

- Using just the basic data did result in a reasonable number of structures, but it was necessary to auto-eliminate a rather large number of correlations.
- The use of the fragment tool, defining a functional group, and the book-keeping tools were covered, which while they may not have been necessary in this example their benefit of reducing the number of structures and the required time was greatly decreased.
- The handling of the INADEQUATE spectrum was introduced. In this particular example, it did not prove to be as useful as the simple addition of an assigned fragment, and functional group definition, but in other cases, the un-ambiguous C-C connectivity of this experiment can be highly beneficial.

6 Contact

Manufacturer

Bruker BioSpin GmbH
Silberstreifen 4
D-76287 Rheinstetten
Germany

E-Mail: nmr-support@bruker.com

<http://www.bruker.com>

WEEE DE43181702

Bruker BioSpin Hotlines

Contact our Bruker BioSpin service centers.

Bruker BioSpin provides dedicated hotlines and service centers, so that our specialists can respond as quickly as possible to all your service requests, applications questions, software or technical needs.

Please select the service center or hotline you wish to contact from our list available at:

<https://www.bruker.com/service/information-communication/helpdesk.html>

List of Figures

Figure 2.1:	Custom Install window highlighting the minimum packages necessary for CMC-se installation.....	12
Figure 3.1:	Referencing several 1D and 2D data sets in TopSpin multiple display.	20
Figure 3.2:	Pop-up window indicating a potential problem with the consistency of the SR values between different spectra	20
Figure 3.3:	Structure Elucidation Project Selection Window.....	22
Figure 3.4:	The Project Directory	23
Figure 3.5:	Project dialog, illustrating the information automatically filled in upon starting CMC-se with an active data set.	24
Figure 3.6:	Project window illustrating the balloon help pop-up window upon entering an incorrect molecular formula – In this case, a lowercase “o” was used instead of the required “O”.	25
Figure 3.7:	Example of using “+” for mixed solvent system. The data analysis detects and excludes solvent peaks from all components.	26
Figure 3.8:	Find data window to assist in adding more spectra to the project.	27
Figure 3.9:	New CMC-se project window, containing an empty Correlation Table with appropriate number of rows and columns are present, and the fragment tool on the right side of the window.	29
Figure 3.10:	Initial Combined Spectra Display within the main TopSpin window.	30
Figure 3.11:	Data browser showing the Structure Elucidation (CMC-se) projects in the data tree... ..	31
Figure 3.12:	Project Selection dialog box in the cases that more than one CMC-se project is included with a data set.	32
Figure 3.13:	Help how to use the CMC-se commands.	32
Figure 3.14:	Main project window of the CMC-se module with short explanation of the contents of the “H-C” tab Correlation Table.	33
Figure 3.15:	A figure with brief descriptions of the “H-H” tab for the homonuclear correlations.	34
Figure 3.16:	C-C tab of a project that contains an INADEQUATE spectrum that has been manually picked for the correlations.	35
Figure 3.17:	Correlation Table after it has been toggled back to the old layout.....	37
Figure 3.18:	Structure Elucidation Report.....	38
Figure 3.19:	Automatic analysis dialog window.	39
Figure 3.20:	Automatic Analysis Dialogue box when data sets have previously been analyzed.	40
Figure 3.21:	Delete Content of Correlation Map Options discriminate between “all correlations” or multiplets and correlations from selected experiments.....	41
Figure 3.22:	First page of the Consistency Check Report	43
Figure 3.23:	Graphical representation of a potential failed analysis of proton distribution.....	44
Figure 3.24:	Combined Spectra Display highlighting the failed 1H Integral consistency check.....	44
Figure 3.25:	CSEARCH request dialog.....	45
Figure 3.26:	CSEARCH result example.....	46
Figure 3.27:	Combined Spectra Display (right), illustrating how the cursor is tied the Correlation Table (left).....	46
Figure 3.28:	Clicking on a cell in the Correlation Table will result in that peak being highlighted in the data.....	47

List of Figures

Figure 3.29:	Screen shot of the Combined Spectra Display in the “old” layout. In this mode, more types of spectra are visible at once.	48
Figure 3.30:	Illustration of how the “Old layout” can be modified to allow closer examination of multiple types of spectra at once, but yet simplified so as not to be viewing too much information at once.	49
Figure 3.31:	Combined Spectra Display window enlarged to highlight features used for manipulating the display options.	50
Figure 3.32:	Peak-actions pop-up window	53
Figure 3.33:	Resize multiplet box function to change box size or position.	53
Figure 3.34:	Using the “Disable or enable current correlation” to add a currently non-assigned correlation to the table.	54
Figure 3.35:	Right clicking to enable a picked, but not assigned correlation will add both the atom and the correlation to the table.	54
Figure 3.36:	Pop-up window indicating that the selected correlation can’t be assigned to the current Correlation Table.	55
Figure 3.37:	Multiplet Properties window for an HSQC correlation.....	56
Figure 3.38:	The Atom Properties Dialog Box	56
Figure 3.39:	Combined Spectra Display prepared for manual addition of correlations to table	58
Figure 3.40:	Correlation Table, and Combined Spectra Display, after deleting the incorrect HSQC correlation.	59
Figure 3.41:	Manual definition of a multiplet in the Combined Spectra Display window.....	59
Figure 3.42:	Correlation Table, and Combined Spectra Display after manually adding the HSQC peak that corresponds to C15 that was previously missing.....	60
Figure 3.43:	Additional Modification Actions Table.	61
Figure 3.44:	Manual remove options	62
Figure 3.45:	Dialogue box asking the user the extent of the information to be removed through this action.....	62
Figure 3.46:	Starting situation where the automatic analysis has picked the solvent peak (indicated by the cursor) but missed HMBC signals from a quaternary carbon (bottom of the figure).....	63
Figure 3.47:	Removing incorrect atoms from the Correlation table, with the “Clear Atom Correlations” through the Correlation Table.	64
Figure 3.48:	The manual addition of an atom and corresponding HMBC correlations to the table through the “Add an Atom to the Table”, option in the Combined Spectra Display.	66
Figure 3.49:	Situation of large ^{19}F - ^{13}C coupling which requires manual modification.	67
Figure 3.50:	The result of merging the split multiplets of a ^{13}C coupled to ^{19}F	68
Figure 3.51:	The INADEQUATE spectrum must be manually referenced appropriately prior to peak picking in CMC-se.	69
Figure 3.52:	In the C-C tab of CMC-se, there is an additional cross-hair that indicates the corresponding carbon chemical shift of a correlation, allowing for quicker visual analysis of the INADEQUATE spectrum.	70
Figure 3.53:	When the INADEQUATE spectrum is folded, the CMC-se module will recognize this and adjust the cross-hair accordingly. It still indicates the corresponding ^{13}C chemical shift, but the cursor is red to indicate that it is a folded correlation.	70
Figure 3.54:	Defining the Multiplet	71
Figure 3.55:	Pop up box prompting the user to double check the position of the cursor and vertical cross hairs because the software can’t find a carbon to correspond to the current position.	71

Figure 3.56:	Dialogue box asking the user the extent of the information to be removed through this action.....	72
Figure 3.57:	Pop Up menus resulting from right clicking in the Correlation Table. Both menus have options for clearing content of the Correlation Table.	73
Figure 3.58:	Example pop-up menu to manipulate HSQC/HMBC correlations.	74
Figure 3.59:	Pop-up menu to manipulate COSY correlations.....	74
Figure 3.60:	Correlation Table displaying disabled COSY and HMBC correlations	75
Figure 3.61:	Pop-up dialog box to edit correlation properties.	75
Figure 3.62:	Atom Properties window for a selected heteroatoms	77
Figure 3.63:	Atom Properties window for a 1H atom	78
Figure 3.64:	Atom Properties pop-up, and resulting changes in Correlation Table, and Combined Spectra Display when an atom is given a name by a user.	79
Figure 3.65:	Correlation Table highlighting atom properties when a user Comment is added	80
Figure 3.66:	Combined Spectra Display after providing user Annotations for an HMBC correlation	81
Figure 3.67:	Fragment Tool within the main CMCse project window.....	82
Figure 3.68:	Fragment editor in selection mode	83
Figure 3.69:	Drawing Normal Bonds	83
Figure 3.70:	Drawing chain bonds	84
Figure 3.71:	Docking Recognition for Binding Sites.....	84
Figure 3.72:	Illustration for the Work with Assignments.....	85
Figure 3.73:	Illustration of how the experimental correlations can be used to graphically assign fragments to be used as building blocks for the structure generation.	85
Figure 3.74:	Illustration for the adding of fragments through the assignment of functional groups. .	86
Figure 3.75:	Fragment Filter dialog window	87
Figure 3.76:	Structure Generation Options window highlighting the default generation options.	89
Figure 3.77:	Execution Control Options	90
Figure 3.78:	Substructure and Ring Rules Options	91
Figure 3.79:	Correlation Control Options	92
Figure 3.80:	Error Message for an Incorrect Bond Count.....	92
Figure 3.81:	Chemistry Rules Options	95
Figure 3.82:	Structure of Betaine (trimethylglycine) generated in CMC-se 2.1.....	95
Figure 3.83:	A chemical formula resulting in a fractional or negative Double Bond Equivalent value.	96
Figure 3.84:	Generating structures progress display	97
Figure 3.85:	Generating Structures Progress Display	98
Figure 3.86:	Updated Correlation Table with Generated Structures	99
Figure 3.87:	Structure Display Window	102
Figure 3.88:	The filter structure dialog box that allows user definable elements to be filtered for or against in the generated structure set.	103
Figure 3.89:	Pop-up window for Structure Options	103
Figure 3.90:	The 5 pages from the report generated from the above "save" options.	105
Figure 3.91:	Structure panel illustrating multiple isomorphs of a generated structure	107
Figure 3.92:	The structures have been color coded based on the chemical shift prediction results.	109
Figure 3.93:	Properties of a Structure Selected from Structure Panel.....	110

List of Figures

Figure 3.94:	Linking of atoms in a structure to the contents of the Correlation Table	111
Figure 3.95:	Illustration of highlighting a single “fuzzy” HMBC correlation from the table on the structure.....	112
Figure 3.96:	Clicking on a peak in the Combined Spectra Display will also result in that correlation being highlighted on the structure candidates for evaluation.....	113
Figure 3.97:	Popup Highlight Correlations	114
Figure 3.98:	Highlighting the invalid correlations found for the first structure on the rest of the structures.....	115
Figure 3.99:	Displaying invalid correlations on structure, Correlation Table, and how this information can be evaluated in the Combined Spectra Display with the Correlated Cursor... ..	116
Figure 3.100:	Displaying atoms without any constraints.....	117
Figure 3.101:	Highlight Correlations ROESY	118
Figure 3.102:	Stereochemistry information from proton correlations	118
Figure 3.103:	CSEARCH dialog.....	119
Figure 3.104:	CSEARCH Robot Referee Result.....	120
Figure 3.105:	Proposed CMC-se Work Flow	121
Figure 3.106:	Flow Chart for the case when zero or unreasonable structures are generated.....	122
Figure 3.107:	Flow Chart for the case when too many structures generated	123
Figure 4.1:	Verification Settings	125
Figure 4.2:	Selecting multiple structure candidates for the verification to decide which, if any, are consistent with the Correlation Table.	126
Figure 4.3:	The results of the CMC-se verification routine.....	127
Figure 4.4:	Signed structure in Fragment editor for further verification.....	128
Figure 4.5:	Fragment pane in CMC-se window	128
Figure 4.6:	Verification settings.....	129
Figure 4.7:	Pop-up window for Structure Options	129
Figure 4.8:	The 5 pages from the report generated from the above “save” options.	130
Figure 5.1:	Correlation Table after automatic analysis. The top panel is the H-C tab, and the bottom is the H-H tab.....	132
Figure 5.2:	The Combined Spectra Display is updated with the peaks that are picked during the automatic analysis.	133
Figure 5.3:	An alternative view option which when working with simple spectra like this example allows all data to be examined at the same time (Arrange Frames).....	134
Figure 5.4:	CMC-se Consistency Report.	135
Figure 5.5:	Results of Consistency Check.....	136
Figure 5.6:	Default Structure Generation Options.....	137
Figure 5.7:	Error message resulting from structure generation with default options.....	138
Figure 5.8:	Checking the quality of the “Inconsistent” COSY correlations	138
Figure 5.9:	Structures generated when the auto-elimination options are activated.....	139
Figure 5.10:	Closer look at the Generated and Ranked Structures	140
Figure 5.11:	Graphical representation of the chemical shift prediction results	141
Figure 5.12:	Top ranked structures highlighting the correlations that were automatically eliminated for structure #2 on all structures and in the table.....	142
Figure 5.13:	Correlation Table with the Fragments Window Open for Manual Elucidation	143
Figure 5.14:	Using the Combined Spectra Display to clarify Fuzzy Correlations when Possible	144

Figure 5.15:	Modifying the properties of a correlation to remove the alternate atom assignments ..	144
Figure 5.16:	Defining C1 as a ketone through the Atom Properties dialogue box.	145
Figure 5.17:	Fragments window automatically updated with the newly defined fragment.....	146
Figure 5.18:	Clicking on an atom will highlight the experimental constraints for that atom.....	147
Figure 5.19:	Clicking and dragging another atom to an existing functional group to make additional assignments.	148
Figure 5.20:	Using the NMR data to help extend the current functional group into a larger fragment.....	149
Figure 5.21:	The smaller functional group has been extended to a larger fragment	150
Figure 5.22:	Identifying the next atom in the fragment by the COSY correlations	151
Figure 5.23:	No longer clear which atom is next due to too many correlations to C6.....	152
Figure 5.24:	Using the Correlation Table to determine additional fragments that can be defined	153
Figure 5.25:	Using the templates to add a 6 member ring to the Fragments tool workspace.....	154
Figure 5.26:	Selecting and merging two fragments into one in the Fragments tool.....	155
Figure 5.27:	Continuing the assignments of the ring based on HMBC correlations	156
Figure 5.28:	Two additional carbons assigned to the ring	157
Figure 5.29:	HMBC correlations identify C5 as one of the next members of the ring	158
Figure 5.30:	Correlations and chemical shift identify the position of the double bond in the ring and it is assigned accordingly	159
Figure 5.31:	Additional HMBC correlations to further elucidate the molecular structure	160
Figure 5.32:	The remaining carbon on the ring can easily be assigned, and the structure completed by process of elimination	161
Figure 5.33:	Using the CMCse verification tool to check the accuracy of the manually elucidated structure.....	162
Figure 5.34:	Selecting the Assigned Fragment option, so that the currently assigned fragment will be run through the verification routine	162
Figure 5.35:	Results of the chemical shift prediction from the Verification routine are shown in the Structures tab	163
Figure 5.36:	Automatic analysis options illustrating that some but not all experiments have the option of being analyzed or not.....	164
Figure 5.37:	Results of the Automatic Analysis	165
Figure 5.38:	Report from the Consistency Check.....	166
Figure 5.39:	Report from the Consistency Check.....	167
Figure 5.40:	Region of the HSQC spectrum in the Combined Spectra Display indicating the source of the inconsistency.	168
Figure 5.41:	Even after manually correcting the overlapping correlations, there are still problems with the Correlation Table; too many protons are included.	169
Figure 5.42:	Double clicking on the HSQC correlation for C15 brings up the "Atom Properties" window, where the "Number of attached H atoms" can be corrected from the current value of "1" to the appropriate value of "3" for this CH3 group.	170
Figure 5.43:	Correlation Table and Combined Spectra Display after correcting the proton distribution errors.....	171
Figure 5.44:	Structure generation options with the maximum number or automatically eliminated correlations (COSY and HMBC) has been set to 3.	172
Figure 5.45:	Generated structure candidates when the auto elimination # was set to 3	172
Figure 5.46:	In the H-H tab, there is un-picked COSY peak between H/C15 and H/C8 that can be manually picked and added to the table.	173

Figure 5.47:	Double clicking on an HSQC peak to bring up the Atom Properties dialog box so that the proton multiplicity can be defined.	174
Figure 5.48:	Final Correlation Table that will be used for the generation for this example data set.	175
Figure 5.49:	A subset of the structures that were generated when the additional constraints were added prior to filtering out the 3 and 4 member rings.	176
Figure 5.50:	Dialog box that can be used to filter the generated structures based on user definitions, in this case excluding structures with 3, 4-member rings.	177
Figure 5.51:	Clicking on any peak in any spectrum will result in that correlation being highlighted on the structure candidates for evaluation purposes.	178
Figure 5.52:	After Linear Prediction, the Automatic Analysis is able to distinguish the overlapping HSQC signals that required manual manipulation above.	179
Figure 5.53:	Spectra illustrating the increase of artifacts when linear prediction beyond the standard 2x is utilized.	180
Figure 5.54:	Starting Correlation Table after the Automatic Analysis.	181
Figure 5.55:	Closer inspection reveals a miss-assignment of the carbon resonances for 2 of the HSQC correlations.	182
Figure 5.56:	Manually changing the HSQC correlations for C12 and C13.	183
Figure 5.57:	Manually changing the HSQC correlations for C12 and C13.	183
Figure 5.58:	Manually peak picking weaker HMBC correlations that were not originally included by the Automatic Analysis.	184
Figure 5.59:	The "Charged atom States" section of the "Chemistry Rules" by default does not allow charged atoms or molecules.	185
Figure 5.60:	The 68 structures that are generated with the above fixed Correlation Table, the auto elimination value set to 15, and allowing charged atoms.	186
Figure 5.61:	The C-C tab of the Correlation Table after careful manual peak picking of the INADEQUATE spectrum.	187
Figure 5.62:	The 16 structures that are now generated with the inclusion of the INADEQUATE for additional constraints.	188
Figure 5.63:	Aromatic region of the HSQC spectrum in the Combined Spectra Display window.	189
Figure 5.64:	Pop-Up Dialogue box to allow the user to delete non-required, or all of the content from the Correlation Table.	190
Figure 5.65:	Opening the Fragment Management tool to define the required fragment for this structure generation.	190
Figure 5.66:	A required fragment has been added to the Fragments Tab and is ready for atom assignments if available.	191
Figure 5.67:	Once an atom is assigned to a position in the fragment by clicking and dragging the sphere accordingly, the experimental correlations to that atom can also be viewed.	192
Figure 5.68:	The correlations are also shown on the fragment when more than one atom is assigned.	193
Figure 5.69:	Even correlations not included in the table can be visualized on the fragments.	194
Figure 5.70:	There is an HMBC from C5 not only to the 4° carbons in this ring, but also to N22.	195
Figure 5.71:	Using the Fragment Tool to modify existing assigned fragments.	196
Figure 5.72:	Double clicking on any HSQC correlation in the Combined Spectra Display will bring up the "Atom Properties" dialog box where the user can define things like Functional group, and even add a comment.	197
Figure 5.73:	Once a functional group is defined, it is added to the Fragments Tab, for additional assignments, modifications as desired by the user.	198

Figure 5.74:	Right clicking on any correlation will bring up a properties box in which an annotation can be entered and then shown on the spectrum in the Combined Spectra Display. ...	199
Figure 5.75:	Resulting 2 structures that were generated after the deletion of the INADEQUATE correlations, but addition of the assigned fragments.	200





Bruker Corporation

info@bruker.com
www.bruker.com

Order No: H149082