RZ 3916 (# ZUR1701-090) 1/30/2017
Physical Sciences 38 pages

Research Report

Heavy oil based mixtures of different origins and treatments studied by AFM

Bruno Schuler,[1] Shadi Fatayer,[1] Gerhard Meyer,[1] Estrella Rogel,[2] Michael Moir,[2] Yunlong Zhang,[3] Michael R. Harper,[3] Andrew E. Pomerantz,[4] Kyle D. Bake,[4] Matthias Witt,[5] Diego Peña,[6] J. Douglas Kushnerick,[3] Oliver C. Mullins,[4] Cesar Ovalles,[2] Frans G. A. van den Berg,[7] and Leo Gross[1]

[1]IBM Research – Zurich, 8803 Rüschlikon, Switzerland
[2]Petroleum and Materials Characterization Unit, Chevron Energy Technology Company, 100 Chevron Way, Richmond CA 94801, USA
[3]ExxonMobil Research and Engineering Company, Annandale NJ 08801, USA
[4]Schlumberger-Doll Research, Cambridge MA 02139, USA
[5]Bruker Daltonik GmbH, Fahrenheitstrasse 4, 28359 Bremen, Germany
[6]Centro de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS) and Depto. de Química Orgánica, Universidade de Santiago de Compostela, Santiago de Compostela 15782, Spain
[7]Shell Global Solutions International B. V., Grasweg 31, 1031 HW Amsterdam, The Netherlands

Heavy oil based mixtures of different origins and treatments studied by AFM

Bruno Schuler,[*,†] Shadi Fatayer,[†] Gerhard Meyer,[†] Estrella Rogel,[‡] Michael Moir,[‡] Yunlong Zhang,[¶] Michael R. Harper,[¶] Andrew E. Pomerantz,[§] Kyle D. Bake,[§] Matthias Witt,[‖] Diego Peña,[⊥] J. Douglas Kushnerick,[¶] Oliver C. Mullins,[§] Cesar Ovalles,[‡] Frans G. A. van den Berg,[#] and Leo Gross[*,†]

†*IBM Research – Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland*
‡*Petroleum and Materials Characterization Unit, Chevron Energy Technology Company, 100 Chevron Way, Richmond CA 94801, USA*
¶*ExxonMobil Research and Engineering Company, Annandale NJ 08801, USA*
§*Schlumberger-Doll Research, Cambridge MA 02139, USA*
‖*Bruker Daltonik GmbH, Fahrenheitstrasse 4, 28359 Bremen, Germany*
⊥*Centro de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS) and Departamento de Química Orgánica, Universidade de Santiago de Compostela, Santiago de Compostela 15782 Spain*
#*Shell Global Solutions International B. V., Grasweg 31, 1031 HW Amsterdam, The Netherlands*

E-mail: bschuler@lbl.gov; lgr@zurich.ibm.com

Abstract

Heavy oil molecular mixtures were investigated on the basis of single molecules resolved by atomic force microscopy. The eight different samples analyzed include asphaltenes and other heavy oil fractions of different geographic/geologic origin and processing steps applied. The collected AFM data of individual molecules provide information about the molecular geometry, aromaticity, the content of non-hexagonal rings, typical types and locations of heterocycles, occurrence, length and connectivity of alkyl side chains and ratio of archipelago- versus island-type architectures. Common and distinguishing structural motifs for the different samples could be identified. The measured size distributions and the degree of unsaturation by scanning probe microscopy is consistent with mass spectrometry data presented herein. The results obtained reveal the complexity, properties and specifics of heavy oil fractions with implications for upstream oil production and downstream oil processing. Moreover, the identified molecular structures form a basis for modeling geochemical oil formation processes.

Introduction

Understanding chemical transformations of organic matter under different environmental conditions forms the foundation of organic geochemistry. This knowledge helps to engineer industrial processes to separate and refine natural molecular mixtures. Thereby the structure of the molecules itself is a key indicator to trace the state and physicochemical conditions prevailing in such transformations. Often, however, the structural and compositional com-

plexity of a natural product impedes a detailed inspection of the molecular architecture.

Recently, high-resolution noncontact atomic force microscopy (AFM) with CO-functionalized tips[1] was introduced as a general analytical tool to study the structure of organic samples with single-molecule sensitivity.[2,3] It could be shown that even complex mixtures such as asphaltenes – the toluene soluble but n-heptane insoluble fraction of organic mixtures[a] – can be studied by AFM.[5] This preceding work characterized in detail coal derived asphaltenes, which serve as model compounds that are highly suited for AFM due to their planarity and small content of flexible side chains. Petroleum asphaltenes were found to be less planar, more complex and thus more challenging to assign by AFM.[5] Recently we established fingerprinting of molecular moieties such as alkyl groups and aliphatic rings[6] improving the assignment of individual hydrocarbons by AFM. Now we studied petroleum-based samples from real world value chains in upstream and downstream oil production. Mastering the structural and compositional complexity of such compounds is a great scientific challenge with huge economic relevance.[7–12]

Table 1: Measured samples.

A1	Crude oil asphaltene
A2	Asphaltene from deposit from A1 oilfield
B1	Vacuum resid asphaltene from a heavy oil
B2	Asphaltene from hydroconverted product from B1 oil
C1	Asphaltene from steam cracker tar
C2	Vacuum residue after asphaltene removal
D1	Shale oil bitumen asphaltene (Green River)
D2	Shale oil bitumen asphaltene (Eagle Ford)

[a] This is the current, generally accepted definition of asphaltenes. Methods of asphaltene preparation vary by solvent, time, and temperature depending on the study or intended use.[4] In this paper some samples called asphaltenes are precipitated in n-heptane but not dissolved in toluene (see methods section).

Here we studied four pairs of heavy oil mixtures (see Table 1): Asphaltenes **A1** obtained from a crude oil with precipitation behavior and **A2** the asphaltene fraction of a deposit formed by this crude oil obtained from the oilfield. Comparing **A1** and **A2** will help to understand what molecules tend to form deposits. The asphaltene fraction of a feedstock **B1** extracted from a virgin heavy oil (oil sand) and the product **B2** after passing a residue hydroconversion unit in order to understand the chemical transformation after catalytic hydrocracking. Asphaltene **C1** was isolated from a tar sample produced by heavy feed steam cracking and represents products from severe thermal cracking of petroleum. **C1** was thought to be well-suited for CO-tip AFM analysis due to its similarity to coal asphaltenes (c.f. Ref. 5), namely the high aromaticity / planarity and low degree of methylation. **C2** is a heavy aromatic fraction of a virgin vacuum residue from which the asphaltenes have been removed. Sample **C2** was selected due to its contrast in chemical properties from **C1**, e.g. the presence of naphthenic rings and alkyl chains, as a means to bookend the hydrocarbon systems studied. **C2** also enabled the study of structural differences between asphaltenic and non-asphaltenic molecules. **D1** and **D2** represent asphaltenes from immature organic-rich shales. Sample **D1** comes from the R1 zone of the Garden Gulch member of the Green River formation in Colorado, USA.[13] Sample **D2** comes from an unweathered outcrop of the Lower Eagle Ford formation in Del Rio, Texas, USA.[14] These immature shale asphaltenes (**D1** and **D2**) were chosen to compare against petroleum asphaltenes (samples **A**-**C**) because immature shale asphaltene contain larger hydrogen content (H/C $\approx$ 1.4) than petroleum asphaltenes (H/C $\approx$ 1.1).[15,16]

These samples span a wide range of crude sources, petroleum processes, chemical compositions, and aromatics contents. Our main incentive is to obtain individual molecular structures for each sample and to investigate the structural differences between the samples, in particular between each sample pair. The

single-molecule approach enables the correlation between molecular size, structural and compositional information. Especially, characteristics of the molecular architecture of the different samples are obtained. Additionally, we compare the results obtained by scanning tunneling microscopy (STM) and AFM to mass spectrometry (MS) data to verify a consistent interpretation.

Results and Discussion

The measured molecules in each mixture were randomly selected, i.e., by measuring all molecules of overview images taken at random positions. We measured about $50-150$ molecules per mixture. First, we obtained STM topography images and determined the STM footprint of each molecule as an approximate value of its size (see Supplemental Information). In Figure 1 the distribution of STM footprints are shown for each sample. Measured molecules from **A1**, **A2**, **B2** and **C2** show a comparably broad distribution with values ranging from $20\,\text{Å}^2$ up to $800\,\text{Å}^2$, indicating a large variety in molecular sizes. On the contrary, molecules sampled from **C1** and **D1** peak around $150-200\,\text{Å}^2$ with a more confined distribution. The naturally-occurring asphaltenes from shale (**D1** and **D2**) are generally smaller than naturally-occurring asphaltenes from crude oil (**A1** and **B1**). By comparing each sample pair we find that the area distribution is very similar for the crude oil **A1** and the deposit **A2**. The distribution of the hydro-treated sample **B2**, on the other hand, expands towards larger sizes as compared to the feedstock material **B1**. Also for samples **C1**/**C2** and **D1**/**D2** the latter in both sets is on average larger than the former. The similar distribution found for **A** and the larger average size of the second sample in **B**, **C** and **D** is also supported by the average molecular weight obtained by MS (see Table S1 in the Supplemental Information).

The variability of the distribution can also be used to assess the quality of the sampling. In our case the quality is mainly determined by the number of molecules measured (as the selection process is inherently random). We find that about $50-100$ molecules are a lower limit to represent the spread of molecules in the present samples. Most importantly the good qualitative agreement between the area distributions observed with STM with the mass distribution measured by MS corroborates that a consistent set of molecules is measured in our experiments. Unlike in MS, however, with AFM we can address the structure of individual constituents. Next we discuss atomically-resolved AFM images with CO functionalized tips[1] recorded for more than 400 randomly selected molecules.

A selection of such AFM measurements of single molecules found in the different samples is shown in Figure 2. Some derived chemical drawings thereof are illustrated in Figure 3. We observe a tremendous diversity of different molecular structures in all samples. Nevertheless, we find common and distinguishing structural motifs that are discussed in the following. Note that the characterization of each sample is based on the complete set of AFM images obtained for it. The complete set represents a statistically relevant, unbiased subset of the respective sample (as argued before). Although not all structures can be assigned based on their respective AFM image, they contain important information on size, planarity, compactness of each molecule and content and size of polycyclic aromatic hydrocarbons (PAHs), alipahtic side chains and methyl side groups. The complete sets of AFM images and their analysis are detailed in the Supplemental Information.

Common to all samples is that the 'island'-type architecture[17–20] , i.e. a central aromatic core with attached side-chains is predominant over 'archipelago'-type molecules (that were also detected but only with occurrences of less than 10% in all samples, e.g. **B2.1**). On the structural moiety level, we frequently observe fluoranthene and fluorene-type motifs. For the latter, we can distinguish at least four different AFM patterns, which we tentatively assign to fluorene, dibenzothiophene, fluorenone and car-

bazole. Note, however, that cyclopentadienone and furan are difficult to differentiate by AFM[21] and both might be considered as a possible option in the assignments of Figure 3. Several methyl groups, identified by AFM before,[5] attached to the PAH are common for all eight samples, too. Interestingly many fluorene-type moieties have such a methyl group at their 1 or 8 position for instance molecules **A1.4**, **A2.2**, **C1.1**, **C1.2**, **C1.4** and **D2.3**, which might indicate a specific pathway for their formation. Five-membered rings also occur in fused pairs (**C1.5**, **D2.3**). Occasionally we find saturated 5-membered (**A1.3**, **C1.4**, **C2.2**) and 6-membered (**C2.1**) rings fused to aromatic rings. These groups could be assigned based on a recent study of aliphatic model compounds.[6] If longer aliphatic chains exist in a sample they tend to be present either as a long unbranched strand attached to a PAH (along with methyl groups) (**A1.6**, **B2.3** left, **D1.2** right) or just as an isolated alkane (**A1.2**, **A1.5**, **A2.1**, **B1.3**, **B1.4**, **D1.4**). Such chains can be $5-70$ Å long and are typically buckled or twisted in certain reoccurring patterns. However, often a (at least partially) straight conformation is adopted such that a characteristic zig-zag pattern can be recognized.[5,6] The isolated alkane chains represent coprecipitated waxes that are not asphaltenes by definition but can be found in asphaltene samples if they are not extensively purified.[?]

Next, we summarize some observations made for each pair of samples based on AFM measurements of individual molecules. In samples **A** we find the comparably largest diversity of different structures: Molecules with an extended aromatic core, isolated aliphatic chains and more three-dimensional ('bulky') molecules with roughly equal proportionality. This was characteristic for the **A** mixtures whereas other samples were typically dominated by one or two out of these three types of molecule structures. In tendency, the deposit **A2** contains less 'bulky' molecules compared to the crude **A1**, which is consistent with earlier works showing that the average molar H/C ratio for **A1** is greater than the ratio for **A2**.[22]

Particular for sample **B1** is that its molecules can be essentially classified into aromatics, among which we find even fully aromatic molecules like **B1.1**, and isolated aliphatic chains with similar abundance. The hydroprocessed sample **B2** in contrast contains more 'bulky' molecules at the expense of the aromatic fraction. This can be rationalized by the hydrotreatment, which adds hydrogens to the molecules. Surprisingly, we still find longer aliphatic side chains (**B2.3** left), indicating an incomplete cracking process.

The steam cracked tar asphaltene **C1** is very well suited for AFM-based structure identification as most of the molecules were highly aromatic. Accordingly, a large portion of molecules could be assigned completely. They are more homogeneous in size compared to the other samples and typically feature methyl substituted aromatic hydrocarbons. Nonetheless, we found also an extremely large molecule (**C1.5**) with more than 60 carbon rings including 5-, 6-, 7- and 8-membered rings. Also for other samples we observed such large molecules (see Figure 1) but they were usually not as planar and therefore more difficult to evaluate. The vacuum residue **C2** on the other hand is rather inhomogeneous and clearly less aromatic than **C1**.

The shale oil bitumen asphaltenes **D1** and **D2** were strikingly different as compared to all other samples. Especially the low frequency of aromatic units or linear aliphatic chains observed was striking. This might be related to their low maturity.[15] In these samples we also detected a significant amount of very small molecules of about $2-6$ atoms that give rise to a manipulation pattern (e.g. **D1.1**). In such images the molecule is moved by the tip around a pinning site on the surface and therefore reflects the hexagonal Cu(111) substrate symmetry rather than the molecular structure. In much lower concentration such molecules were also found in all other samples. The small molecules likely represent enclosed gas. Such gas might be preferentially captured by the samples **D** because of their three-dimensional

molecule architecture potentially forming organic matter hosted pores.[23]

On the basis of the structure assignments inferred from AFM measurements we also evaluated the double bond equivalent (DBE)[b] as a function of the carbon number of samples **A** and compared it to atmospheric pressure photoionization Fourier transform ion cyclotron resonance MS (APPI FT-ICR MS) data (see Figure 4). Different mass spectrometry techniques can give some differences in overall asphaltene mass distributions.[20,24] For the comparison to our data derived from AFM of single molecules, that is of very limited resolution in mass and DBE in cases where a significant part of the molecule structure could not be identified, we believe that this MS technique is sufficient. Most importantly the scatter of the molecules in the scanning probe experiments spans the entire distribution measured by MS in both samples. This fact gives us a high confidence that with AFM we measure a representative set of molecules.

Conclusions

In summary, we studied a wide range of different heavy oil hydrocarbon mixtures by high-resolution scanning probe microscopy and resolved individual molecular structures contained in each sample. We could identify repeating and distinguishing structural motifs by AFM, in particular typical molecular moieties and their relative occurrence and typical positions within the molecules. Moreover, statistical analysis of molecular footprints obtained by STM and comparison of AFM data with MS indicate the statistical significance of the sampling. The obtained images support an asphaltene molecular architecture with a central aromatic core with peripheral alkane substituents. Differentiating features between the four pairs of samples and within each pair of samples were detected. The identified molecular structures form a foundation for first principles molecular dynamics and thermodynamic models to understand the highly complex process of oil formation, maturation, its behavior in reservoirs and during refining. The measurements shown here open the prospects to analyze complex geological formation processes and chemical reactions on the single molecule level.

[b] DBE = C – H/2 + N/2 + 1

Materials and Methods

Heavy oil hydrocarbon samples

Samples A1 and A2. A field asphaltene deposit and the corresponding crude oil were obtained from an oilfield produced by CO_2 flooding. Asphaltenes **A1** were extracted from the crude oil using a modification of the ASTM D6560 test.[25] The deposit was separated by solubility according to a fractionation procedure using an Accelerated Solvent Extractor Dionex 300. This procedure yields five asphaltene fractions. Asphaltene **A2** represents 70 wt% of the recovered asphaltenes. Further details of the procedure are presented in the SI. Detailed conventional analysis of **A1** and **A2** are found elsewhere.[22]

Samples B1 and B2. Asphaltenes **B1** and **B2** were obtained from a commercial hydroconversion unit operating at high conversion (typical conversion of feedstock vacuum residue about 80 wt%). The feedstock for this unit consisted of vacuum distilled residue from Athabasca Oilsands. Asphaltenes **B1** from the feedstock (10.5 wt%) were isolated by precipitation in heptane in a 1:40 ratio, followed by Soxhlet extraction. Asphaltenes **B2** (11.7 wt%) from the converted product were isolated following the same procedure.

Samples C1 and C2. Tar was obtained from a heavy feed steam cracker, and asphaltene **C1** was precipitated with 10:1 v/v n-heptane at ambient temperatures (25°C), and collected with filtration with thorough washing with n-heptane. Asphaltene **C1** represents 24.1 wt% from steam-cracked tar sample. **C2** was isolated from a virgin vacuum residue. Asphaltene was precipitated with 10:1 v/v n-heptane and removed, and the n-heptane soluble fraction (maltenes) was concentrated to remove solvent, and then separated using silica gel chromatography. Sample **C2** represents 13.7 wt%

of the vacuum residue.

Samples D1 and D2. The shale asphaltenes were obtained from thermally immature source rocks (Green River[13,15] and Eagle Ford[14,16]). Rock samples were crushed and extracted for several days with a 9:1 methylene chloride:methanol mixture in a Soxhlet extractor to remove soluble organic matter (bitumen). The bitumen was then diluted 1:40 in n-heptane to precipitate the asphaltenes. The asphaltenes were then washed for several days with n-heptane in a Soxhlet extractor to remove coprecipitants. When sufficient sample was available, the purified bitumen asphaltenes were then redissolved in toluene to exclude toluene-insoluble contaminants. Thus, the immature shale asphaltenes represent the toluene soluble, n-heptane insoluble fraction of the immature shales.

STM/AFM experiments

The scanning probe microscopy experiments were carried out using a home-built combined STM/AFM in ultrahigh vacuum ($p \approx 10^{-10}$ mbar) at low temperatures ($T \approx 5$ K) using a qPlus sensor.[26] AFM measurements were acquired in the frequency-modulation mode[27] ($f_0 \approx 30$ kHz and $A = 0.5$ Å) at constant-height and zero sample bias. For both the STM and AFM experiments the tip was functionalized with a CO molecule (CO tip).[1,28] As a substrate a Cu(111) single crystal partially covered by islands of bilayer NaCl[29] was used. The molecules were evaporated from the solid phase *in situ* by rapid resistive heating of a oxidized Si wafer directly onto the substrate at $T \approx 10$ K.[6] The molecules were either measured on the Cu or the NaCl substrate.

The molecule structures were identified by their CO-tip AFM contrast at different scan heights.[2,3] The assignment is based on multiple studies of synthetic[1,6,30–33] and natural molecules[5,34,35] with CO-tip AFM while considering the influence of the molecular adsorption geometry.[30]

FT-ICR MS analysis

Samples A1 and A2. The samples **A** were analyzed using a solariX XR FT-ICR mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) equipped with a 12 T refrigerated actively shielded superconducting magnet (Bruker Biospin, Wissembourg, France) and the ParacellTM analyzer cell. Apollo II Dual ESI/MALDI ion source was used. Samples were analyzed using positive ion mode APPI. Each fraction was dissolved in a toluene/methanol blend at 50 ppm for analysis. The prepared samples were directly injected to the APPI source with a syringe pump at a flow rate of 10 μL/min. The resolving power was 1,300,000 at m/z 400. Further details can be found elsewhere.[36]

Acknowledgments

We thank R. Allenspach for comments. We acknowledge financial support from the ERC Grants CEMAS (agreement no. 291194) and AMSEL (682144), the EU project PAMS (610446), the Spanish Ministry of Science and Competitiveness (MAT2016-78293-C6-3-R), the Consellería de Cultura, Educación e Ordenación Universitaria (Centro singular de investigación de Galicia accreditation 2016-2019, ED431G/09) and the European Regional Development Fund (ERDF).

Supporting Information Available

This material is available free of charge via the Internet at http://pubs.acs.org.

Figures

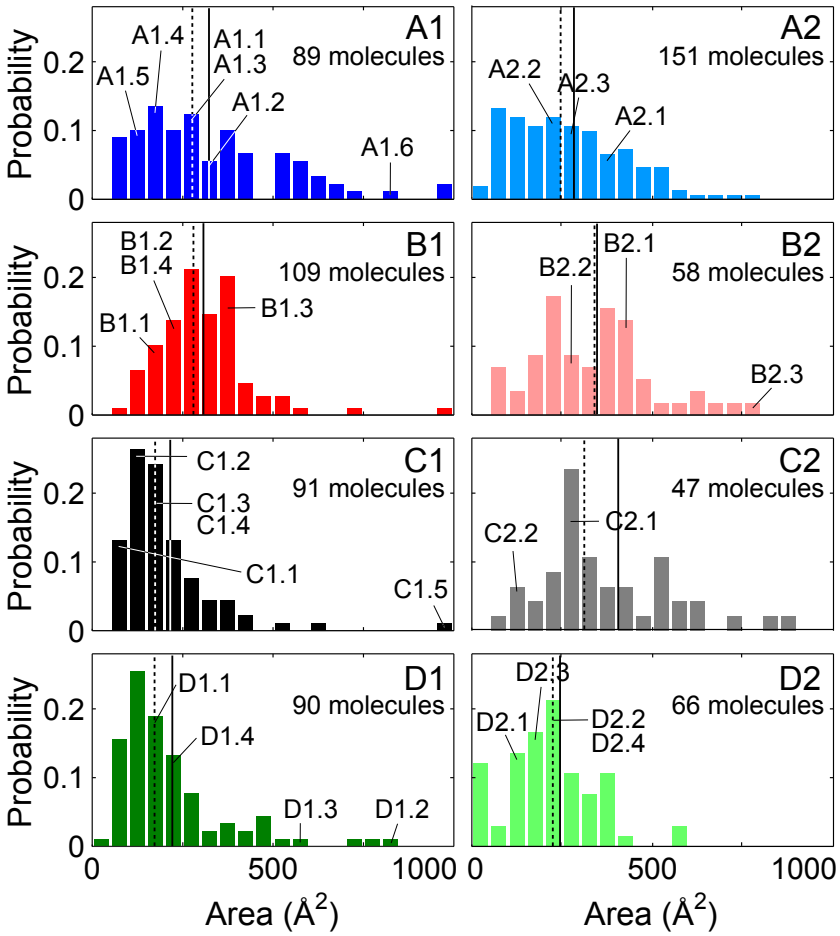


Figure 1: **Area Analysis.** Distribution of the measured STM areas for each sample pair. The mean and median area is indicated by a vertical solid and dashed line, respectively. The number of considered measurements and the assignments of the molecules shown in Figure 2 are indicated in the panel.

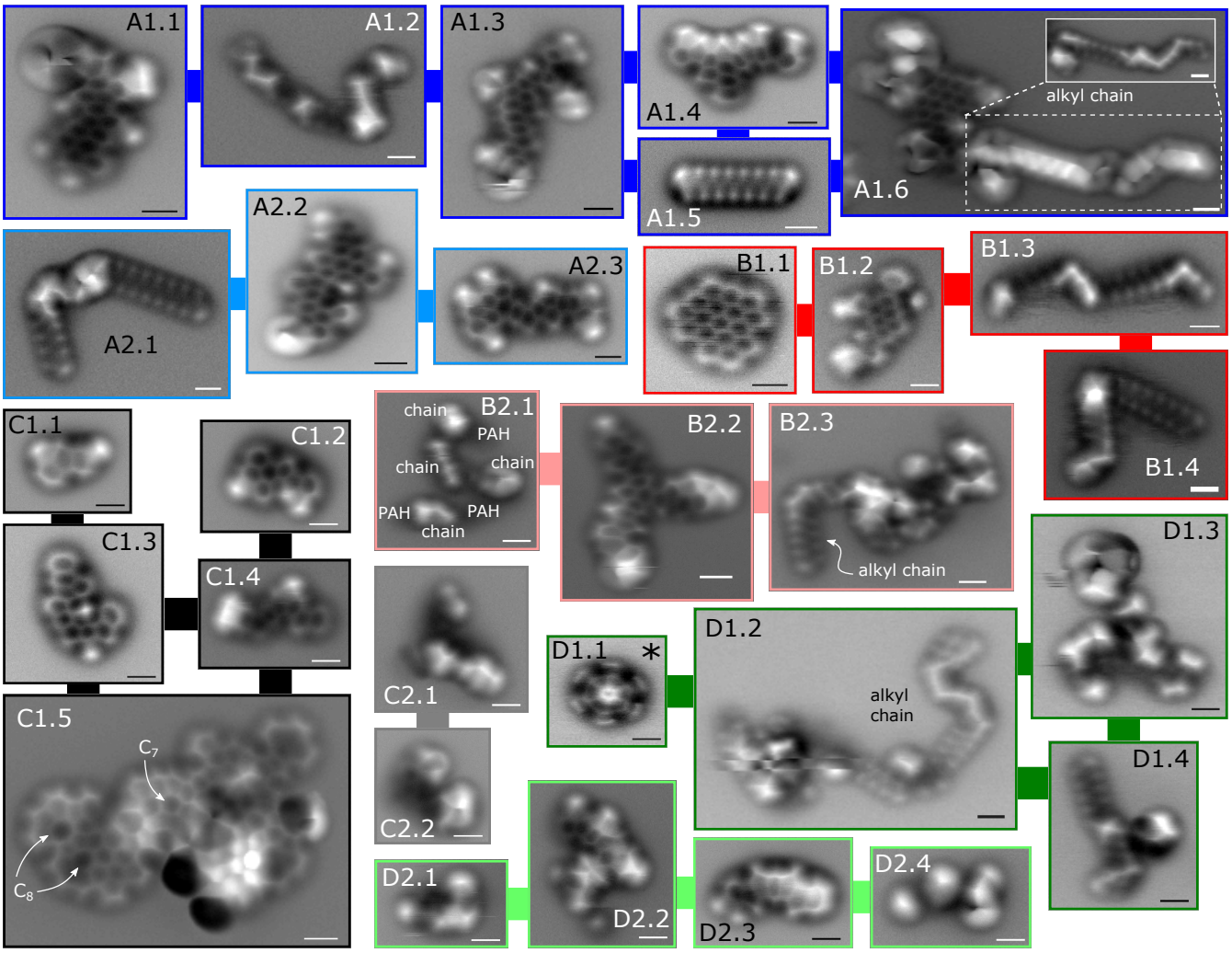


Figure 2: **AFM measurements.** Characteristic selection of CO-tip AFM raw data of individual heavy oil molecules. The labels indicate the sample name and molecule number. For molecule **A1.6** the inset shows a part of the same frame recorded at a larger scan height. The asterisk indicates a manipulation image where the molecule rotates around a pinning site. A more comprehensive set of measurements for each sample can be found in the Supplemental Information. Scale bars: 5 Å.

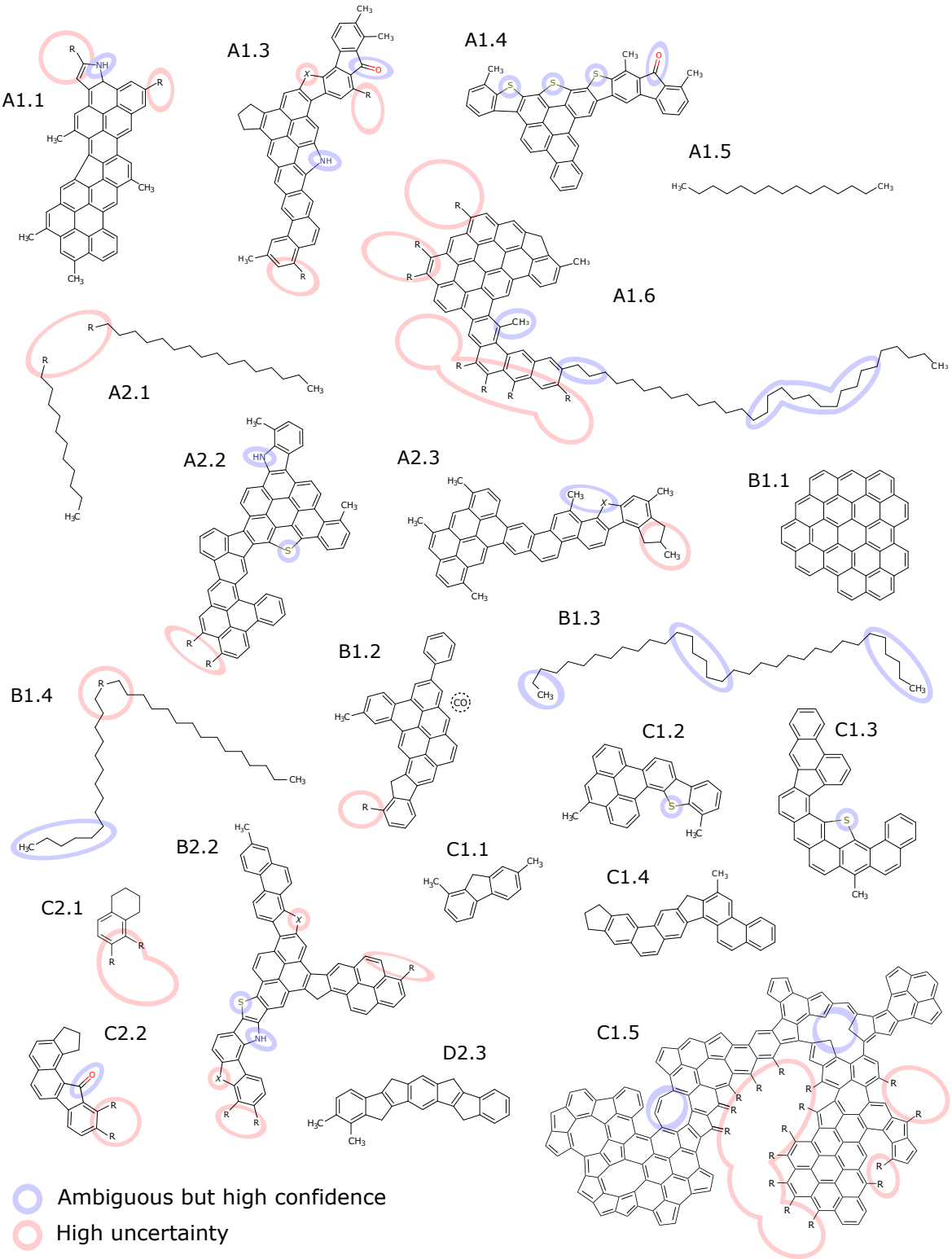


Figure 3: **Structure assignment.** Structure proposals based on AFM measurements. Note that besides the data shown in Figure 2, measurements at different scan heights were considered to derive the structures. X and R denote unknown heteroatoms and side chains, respectively. Regions where a full structure assignment was not possible or was uncertain are highlighted with red and blue circles, respectively.

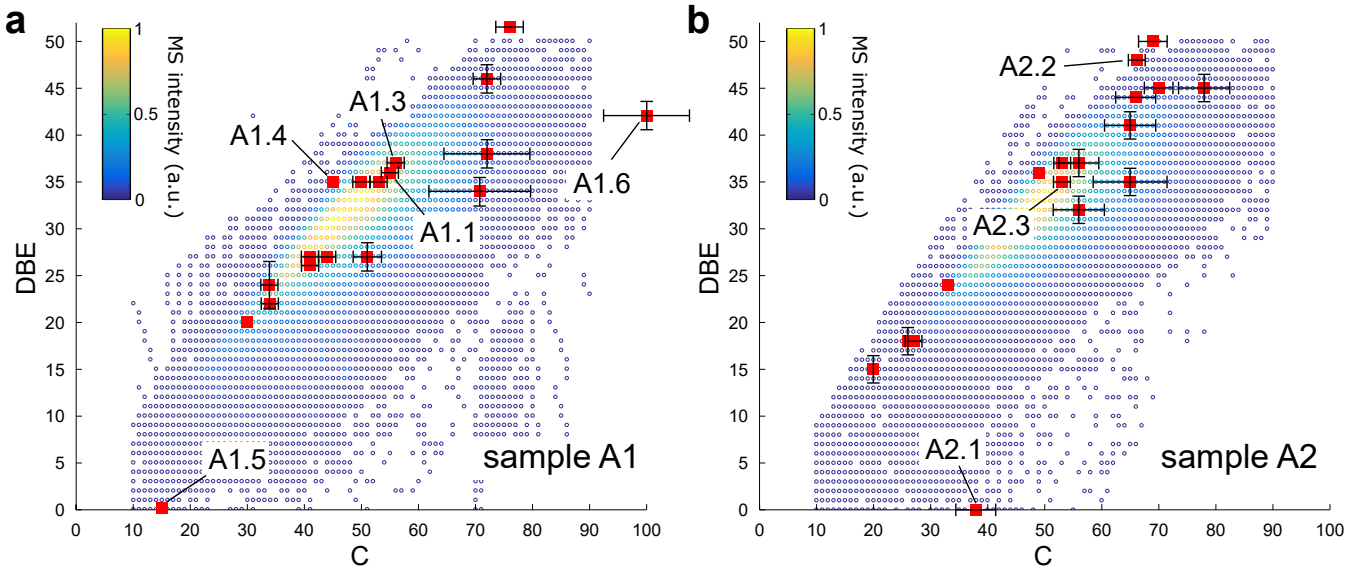
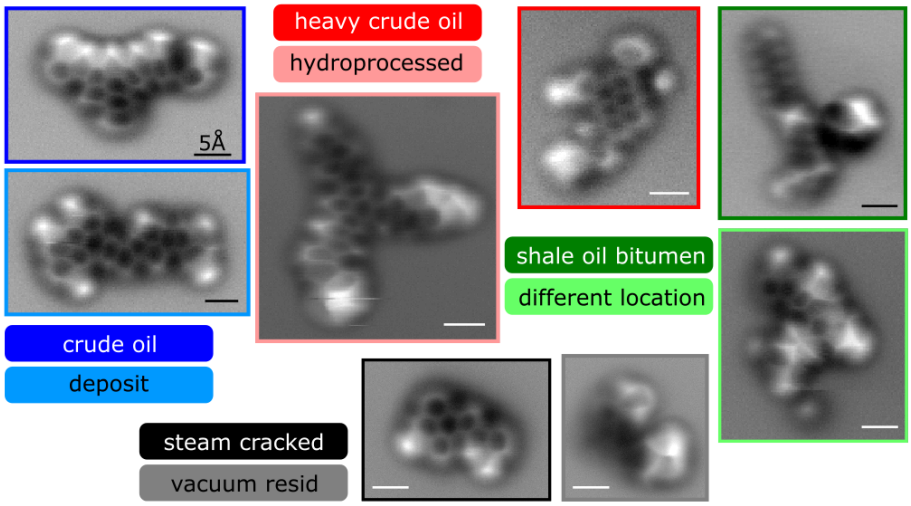


Figure 4: **Carbon number vs double bond equivalent (DBE).** Comparison of mass spectrometry data (colored scatter plot)[36] with structure proposals deduced by AFM (red squares). **a,b** Data obtained for samples **A1** and **A2**, respectively. Data points that refer to structure proposals shown in Figure 3 are labeled.



For Table of Contents Only

References

(1) Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G. *Science* **2009**, *325*, 1110–1114.

(2) Gross, L.; Schuler, B.; Mohn, F.; Moll, N.; Repp, J.; Meyer, G. In *Noncontact Atomic Force Microscopy: Volume 3*; Morita, S., Giessibl, J. F., Meyer, E., Wiesendanger, R., Eds.; Springer International Publishing: Cham, 2015; pp 223–246.

(3) Schuler, B.; Mohn, F.; Gross, L.; Meyer, G.; Jaspars, M. *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*; The Royal Society of Chemistry, 2015; Vol. 1; pp 306–320.

(4) Speight, J. G.; Long, R. B.; Trowbridge, T. D. *Fuel* **1984**, *63*, 616–620.

(5) Schuler, B.; Meyer, G.; Peña, D.; Mullins, O. C.; Gross, L. *J. Am. Chem. Soc.* **2015**, *137*, 9870–9876.

(6) Schuler, B.; Zhang, Y.; Collazos, S.; Fatayer, S.; Meyer, G.; Pérez, D.; Guitián, E.; Harper, M. R.; Kushnerick, J. D.; Peña, D.; Gross, L. **2017**, DOI: 10.1039/C6SC04698C.

(7) Mullins, O. C.; Sheu, E. Y.; Hammami, A.; Marshall, A. G. *Asphaltenes, Heavy Oils, and Petroleomics*; Springer, 2007; Vol. 1.

(8) Larter, S.; Bowler, B.; Li, M.; Chen, M.; Brincat, D.; Bennett, B.; Noke, K.; Donohoe, P.; Simmons, D.; Kohnen, M.; Allan, J.; Telnaes, N.; Horstad, I. *Nature* **1996**, *383*, 593–597.

(9) Head, I. M.; Jones, D. M.; Larter, S. R. *Nature* **2003**, *426*, 344–352.

(10) Aitken, C. M.; Jones, D.; Larter, S. *Nature* **2004**, *431*, 291–294.

(11) Ancheyta, J.; Trejo, F.; Rana, M. S. *Asphaltenes: Chemical Transformation during Hydroprocessing of Heavy Oils*; CRC Press, 2010.

(12) Sheng, J. *Modern Chemical Enhanced Oil Recovery: Theory and Practice*; Gulf Professional Publishing, 2010.

(13) Le Doan, T.; Bostrom, N.; Burnham, A.; Kleinberg, R.; Pomerantz, A.; Allix, P. *Energy Fuels* **2013**, *27*, 6447–6459.

(14) Pomerantz, A. E.; Bake, K. D.; Craddock, P. R.; Kurzenhauser, K. W.; Kodalen, B. G.; Mitra-Kirtley, S.; Bolin, T. B. *Org. Geochem.* **2014**, *68*, 5–12.

(15) Pomerantz, A. E.; Le Doan, T. V.; Craddock, P. R.; Bake, K. D.; Kleinberg, R. L.; Burnham, A. K.; Wu, Q.; Zare, R. N.; Brodnik, G.; Lo, W. C. H.; Grayson, M.; Mitra-Kirtley, S.; Bolin, T. B.; Wu, T. *Energy Fuels* **2016**, *30*, 7025–7036.

(16) Rogel, E.; Ovalles, C.; Bake, K. D.; Zuo, J. Y.; Dumont, H.; Pomerantz, A. E.; Mullins, O. C. *Energy Fuels* **2016**, *30*, 9132–9140.

(17) Groenzin, H.; Mullins, O. C. *J. Phys. Chem. A* **1999**, *103*, 11237–11245.

(18) Groenzin, H.; Mullins, O. C. *Energy Fuels* **2000**, *14*, 677–684.

(19) Sabbah, H.; Morrow, A. L.; Pomerantz, A. E.; Zare, R. N. *Energy Fuels* **2011**, *25*, 1597–1604.

(20) Pomerantz, A. E.; Wu, Q.; Mullins, O. C.; Zare, R. N. *Energy Fuels* **2015**, *29*, 2833–2842.

(21) Mohn, F.; Gross, L.; Meyer, G. *Appl. Phys. Lett.* **2011**, *99*, 053106.

(22) Rogel, E.; Miao, T.; Vien, J.; Roye, M. *Fuel* **2015**, *147*, 155–160.

(23) Löhr, S.; Baruch, E.; Hall, P.; Kennedy, M. *Org. Geochem.* **2015**, *87*, 119–132.

(24) McKenna, A. M.; Donald, L. J.; Fitzsimmons, J. E.; Juyal, P.; Spicer, V.; Standing, K. G.; Marshall, A. G.; Rodgers, R. P. *Energy Fuels* **2013**, *27*, 1246–1256.

(25) American Society for Testing and Materials, Standard Test Method for Determination of Asphaltenes (Heptane Insolubles) in Crude Petroleum and Petroleum Products IP143/01 ASTM D6560. 2005.

(26) Giessibl, F. J. *Appl. Phys. Lett.* **1998**, *73*, 3956.

(27) Albrecht, T. R.; Grutter, P.; Horne, D.; Rugar, D. *J. Appl. Phys.* **1991**, *69*, 668.

(28) Mohn, F.; Schuler, B.; Gross, L.; Meyer, G. *Appl. Phys. Lett.* **2013**, *102*, 073109.

(29) Bennewitz, R.; Barwich, V.; Bammerlin, M.; Loppacher, C.; Guggisberg, M.; Baratoff, A.; Meyer, E.; Güntherodt, H.-J. *Surf. Sci.* **1999**, *438*, 289–296.

(30) Schuler, B.; Liu, W.; Tkatchenko, A.; Moll, N.; Meyer, G.; Mistry, A.; Fox, D.; Gross, L. *Phys. Rev. Lett.* **2013**, *111*, 106103.

(31) van der Heijden, N. J.; Hapala, P.; Rombouts, J. A.; van der Lit, J.; Smith, D.; Mutombo, P.; Švec, M.; Jelinek, P.; Swart, I. *ACS Nano* **2016**, *10*, 8517–8525.

(32) Majzik, Z.; Cuenca, A. B.; Pavliček, N.; Miralles, N.; Meyer, G.; Gross, L.; Fernández, E. *ACS Nano* **2016**, *10*, 5340–5345.

(33) Kawai, S.; Sadeghi, A.; Okamoto, T.; Mitsui, C.; Pawlak, R.; Meier, T.; Takeya, J.; Goedecker, S.; Meyer, E. *Small* **2016**, *12*, 5303–5311.

(34) Gross, L.; Mohn, F.; Moll, N.; Meyer, G.; Ebel, R.; Abdel-Mageed, W. M.; Jaspars, M. *Nature Chem.* **2010**, *2*, 821–825.

(35) Hanssen, K. Ø. et al. *Angew. Chem. Int. Ed.* **2012**, *51*, 12238–12241.

(36) Rogel, E.; Witt, M. *Energy Fuels* **2016**, *30*, 915–923.

Supporting Information:
Heavy oil based mixtures of different origins and treatments studied by AFM

Bruno Schuler,[1,*] Shadi Fatayer,[1] Gerhard Meyer,[1] Estrella Rogel,[2] Michael Moir,[2] Yunlong Zhang,[3] Michael R. Harper,[3] Andrew E. Pomerantz,[4] Kyle D. Bake,[4] Matthias Witt,[5] Diego Peña,[6] J. Douglas Kushnerick,[3] Oliver C. Mullins,[4] Cesar Ovalles,[2] Frans G. A. van den Berg,[7] and Leo Gross[1,†]

[1]*IBM Research – Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland*

[2]*Petroleum and Materials Characterization Unit, Chevron Energy Technology Company, 100 Chevron Way, Richmond CA 94801, USA*

[3]*ExxonMobil Research and Engineering Company, Annandale NJ 08801, USA*

[4]*Schlumberger-Doll Research, Cambridge MA 02139, USA*

[5]*Bruker Daltonik GmbH, Fahrenheitstrasse 4, 28359 Bremen, Germany*

[6]*Centro de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS) and Departamento de Química Orgánica, Universidade de Santiago de Compostela, Santiago de Compostela 15782 Spain*

[7]*Shell Global Solutions International B. V., Grasweg 31, 1031 HW Amsterdam, The Netherlands*

(Dated: 23rd May 2017)

CONTENTS

I. STM SUPPORTING MEASUREMENTS

To estimate the area footprint of the molecules on the surface we determined the border of the molecule from STM measurements recorded with a CO tip. We defined the border of the molecule as the contour line of the STM topography at half the height of a reference point on the molecule. In Figure S1 an exemplary STM image of molecule **A2.1** is given with its border indicated. The overlay of the same contour line on the AFM image shows that the enclosed area by the border is a good estimate of the molecule footprint. The distribution of all measured molecule areas is shown in Figure 1 of the main article. Please note that the STM area will in general overestimate the footprint of molecules, but can serve as a reference to evaluate the size distribution of comparably flat molecules.

II. AFM SUPPORTING MEASUREMENTS

In this section additional AFM measurements of samples **A1-D2** along with proposed structure assignments are provided and discussed to support the main conclusions drawn in the manuscript. Furthermore, information on the derivation of Figure 4 of the main text is given.

For ease of description we classify the observed molecules in three types: polycyclic aromatic hydrocarbon (PAHs), aliphatic chains and 'bulky' molecules. As 'bulky' we describe

* bschuler@lbl.gov

† lgr@zurich.ibm.com

molecules that have mainly non-planar moieties, which leads to a more corrugated adsorption geometry. In general, the molecule will be a combination of all three types. AFM is especially suited to identify highly aromatic molecules. Aliphatic chains can typically be readily recognized but because of their internal structural degrees of freedom, they may not be easily unambiguously assigned [1]. Finally, 'bulky' molecules are most difficult to characterize by constant-height AFM. Although a structural assignment is in such cases often not possible, the AFM images comprise valuable information about the size and planarity of certain molecule parts. For these reasons, the molecule drawings shown in the manuscript and the Supplement do *not* represent the entire molecule spectrum observed but especially the more aromatic fraction, which is easier to identify. Please also note that the appearance of structural motifs in AFM depends on the tip–molecule distance and therefore on the adsorption geometry of the molecule [2]. The structural assignments are mostly based on several AFM images recorded at different scan heights (but only one image per molecule is shown in the manuscript).

In Figure S2 additional AFM measurements of sample **A1** are shown. As stated in the manuscript a variety of different structures can be found among highly aromatic (**A1.10**, **A1.14**, **A1.16**, **A1.22**, **A1.26**, **A1.43**, **A1.51**), dominantly aliphatic (**A1.11**, **A1.13**, **A1.20**, **A1.23**, **A1.31**, **A1.35**, **A1.38**, **A1.44**, **A1.49**, **A1.54**, **A1.59**, **A1.60**) and three-dimensional 'bulky' molecules (**A1.7**, **A1.17-A1.19**, **A1.24**, **A1.25**, **A1.27**, **A133**, **A1.34**, **A1.40**, **A1.46**, **A1.48**, **A1.55**, **A1.56**), which are more difficult to measure and identify. But even for such 'bulky' molecules, often a planar PAH core that is obscured by the bulkier side groups (that stick out) can be identified as a homogeneous dark (attractive) background (e.g. **A1.27**, **A1.48**). In other cases the tip and adsorbed molecule are so stable that one can scan very close to the molecule to unveil the PAH irrespective of bulkier side groups (**A1.8**, **A1.28**, **A1.39**, **A1.47**, **A1.53**, **A1.56**), the latter however then often appears unstable.

In Figure S3 the structural assignments of some molecules in Figure S2 are given. Additional AFM measurements of **A2** along with their proposed chemical structures can be found in Figure S4 and Figure S5, respectively. We find that samples **A1** and **A2** are very similar in terms of the size of molecules and the ratio of PAH : aliphatic chains : 'bulky'

structures. In tendency, there is less 'bulky' molecules in **A2**.

Based on the proposed chemical structures of samples **A**, the C vs DBE relation was evaluated, shown in Figure S18 and Figure S19. In these plots the (labeled) black squares indicate the positions of the molecules without considering unknown side groups. The size of such unknown side-groups can, however, be estimated based on the extent of their AFM contrast. Because these side-groups are typically non-aromatic, the actual molecule coordinate is horizontally right shifted with respect to the black squares in the C vs DBE diagram. The estimated actual molecule position in Figure S18, Figure S19 is plotted in red and was used for Figure 4 as well. The error bars were estimated based on the uncertainty of the extent of the side groups, possible presence of aromatic rings in the non-planar molecule parts and heteroatom content.

For samples **B1** and **B2**, complementary AFM measurements and proposed structure assignments can be found in Figure S6-S9. The sample **B1** differentiates from samples **A** and **B2** by the fewer occurrence of 'bulky' molecules. We mainly find aromatics, among a completely unsubstituted 17-ring PAH (**B1.1**), and isolated alkane chains. The hydroprocessed sample **B2** exhibits on average larger (e.g. **B2.6**, **B2.9**, **B2.13**, **B2.19**, **B2.22**, **B2.28**, **B2.31**) and more bulkier molecules (**B2.6**, **B2.8**, **B2.11-B2.14**, **B2.19-B2.24**, **B2.28**, **B2.30-B2.33**).

For samples **C1** and **C2**, additional AFM measurements and proposed structure assignments are shown in Figure S10-S13. The steam cracked tar asphaltene sample **C1** was best suited for AFM analysis because it is highly aromatic with mostly only methyl side groups. The molecules are smaller and have a narrower size distribution as compared to samples **A**, **B** and **C2**. Nevertheless its molecule sizes range from 2-3 rings (**C1.1**, **C1.50**) to 14 rings (**C1.27**), up to more than 60 rings (**C1.5**) in extreme cases. Importantly, the observed rings are not all hexagonal but we frequently find pentagonal rings (**C1.1-C1.4**, **C1.8**, **C1.11**, **C1.17**, **C1.19**, **C1.26-C1.29**, **C1.31**, **C1.32**, **C1.34**, **C1.37**, **C1.40**, **C1.43**, **C1.45**, **C1.49**, **C1.53**. **C1.58**, **C1.59**, **C1.62**, **C1.67**) mainly in the form of fluorene-type and fluoranthene motifs, in a few cases as alicyclic rings (see Figure S11). The fluorene-type moieties were often substituted by heteroatoms that we assign to S, CO or NH. Fluorene

moieties (e.g. **C1.1**) are identified by a faint repulsive feature at the position of its sp^3 carbon (c.f. Ref. [2, 3]). The sulfur in dibenzothiophene (e.g. **C1.2**) appears as a larger repulsive (bright) spot. The carbonyl group in fluorenone (e.g. **C1.20**) is imaged as an attractive (dark) spot in CO-tip AFM imaging (c.f. Ref. [2–5]) and an elongated bright feature at the pentagon apex that we frequently observe (e.g. **C1.19**) we tentatively assign to the amino group of carbazole. For the primarily aromatic molecules (where we attempted structure identification by AFM, accounting for 60% of sample **C**) we observe on average about one pentagonal ring per molecule. In rare cases also heptagonal (**B2.34**, **C1.5**) and octagonal rings (**C1.5**) were detected.

Characteristic in **C2** is the comparably low abundance of PAH moieties (**C2.1**, **C2.2**, **C2.8** center, **C2.12** top, **C2.14** center, **C2.23**, **C2.27** top). Mainly we find 'bulky' molecule parts attached to alkyl chains (**C2.4**, **C2.10**, **C2.11**, **C2.13**, **C2.15**, **C2.16**, **C2.19**, **C2.21**, **C2.22**, **C2.24-C2.26**). The size distribution (see Figure 1) is considerably broader than for **C1**.

For samples **D1** and **D2**, complementary AFM measurements and proposed structure assignments can be found in Figure S14-S17. In the shale oil bitumen asphaltenes **D1** and **D2** aromatic units and linear aliphatic chains were much less frequent than in any other sample. This might be related to their low maturity. On average the molecules are also smaller ranging down to a few atoms. Such small molecules give often rise to manipulation patterns (**D1.30-C1.36**) as discussed in the manuscript. The small molecules might be ascribed to enclosed gas. Between sample **D1** and **D2** no qualitative difference could be observed.

III. ADDITIONAL CHARACTERIZATION

A. Extraction procedures

*a. Sample **A1**.* Asphaltene from the crude oil were obtained using a modification of ASTM D6560 test [**?**]. In this modification, a sample/heptane ratio of 1/20 is used instead of the 1/30 used in the ASTM D6560. The blend sample/heptane is filtered at 80°C. The

precipitated material is washed using hot heptane prior to drying and weighing. Results using this modified test for crude oils with contents larger than 1 wt.% are comparable with results obtained by the regular ASTM D6560.

*b. Sample **A2**.* Asphaltenes from the deposit were separated as follows: A sample of the material is weighed (mass around 5.0 g) and dissolved in 50 mL of Methylene Chloride. 50 g of PTFE are added to the solution and stirred during 1 h at room temperature. The solvent was removed by heating at 60 oC under nitrogen. The PTFE supported sample is placed into a 100 mL stainless steel cell and extracted with heptane at room temperature with 60 min of soaking time. This produces the first extracted fraction, the maltenes (heptane solubles). The cell is then extracted with 15/85 CH_2Cl_2/n-heptane (Fraction #1), 30/70 CH_2Cl_2/n-heptane (Fraction #2), 100% CH_2Cl_2 (Fraction #3) and 90/10 CH_2Cl_2/Methanol(Fraction #4), respectively, for 60 min at room temperature each step. Finally, the cell is 'washed' three times with 90/10 CH_2Cl_2/MeOH (Fraction #5) at 120°C for 15 min. Fraction #3 represents 67 wt.% of the organics in the deposit. In this work, we labeled them as **A2**.

*c. Curde oil properties (source of **A1** and **A2**).* Additional characterization of the crude oil that served as the source material for **A1** and **A2** is provided in Table S1.

Table S1. Crude oil properties

Gravity, °API	33.0
Asphaltene, wt.%	1.7
Sulfur, wt.%	0.5
Nitrogen, ppm	829
Nickel, ppm	3.4
Vanadium, ppm	5.21
Viscosity at 50°C, cSt	3.44
MCR, wt.%	2.17

B. Elemental analysis

*a. Samples **A1** and **A2**.* Carbon, hydrogen, and nitrogen (CHN) analysis was carried out with a Carlo Erba model 1108 analyzer. Metal and sulfur were determined using a Thermo Intrepid ICP.

Table S2. **A1**: Asphaltenes from crude oil with precipitation behavior. **A2**: Deposit recovered from same oil field as **A1**. **B1**: Asphaltenes from a virgin heavy oil (from oil sands). **B2**: **B1** after residue hydroconversion unit. **C1**: Steam Cracked Tar Asphaltenes. **C2**: Aromatic Fraction isolated from Vacuum Residue. **D1**: Asphaltenes from bitumen from shale oil from Green River. **D2**: Like **D1** but from Eagle Ford. c: Asphaltene content. mw: Measured molecular weight. ρ: Density. δ: Hildebrand solubility parameter.

			elemental analysis (wt %)						metal (ppmw)			
sample	c (%w)	mw (g/mol)	C	H	S	N	O	H/C	V	Ni	ρ (g/ml)	δ ($\sqrt{MPa}$)
A1	1.7	676 $(200-1500)^a$	88.0	6.8		0.6		0.9 [6]			1.23[b]	21.3
A2	67	703 $(200-1500)^a$	88.3	6.3	3.1	0.7		0.85 [6]	41	89	1.26	21.7
B1	10.5[c]	442 $(200-700)^a$	80.5	7.7	7.5	1.4	2.6	1.1	1085	406	1.12	20.2
B2	11.7[c]	462 $(200-800)^a$	87.6	5.6	1.7	1.2	0.8	0.8	577	504	1.27	21.8
C1	24.1	457 $(155-896)^a$	89.0[d]	5.9[d]	3.2[d]	0.2[d]	0.2[e]	0.79				
C2	13.7	696 $(309-1154)^a$	82.2[d]	8.8[d]	5.6[d]	0.7[d]	1.6[e]	1.28				
D1		392[f]	66.6	7.6	3.4	1.8		1.36				
D2		528[f]	73.3	8.5	9.7	1.7	6.5	1.4 [7]			1.15 [7]	19.9 [7]

[a] APPI FT-ICR MS: Molecular weight calculated as weighted average based on responses

[b] Calculated based on correlation density/hydrogen content

[c] Concentration measured in the vacuum residue

[d] ASTM D5291

[e] Oxygen by reductive pyrolysis method

[f] L2MS

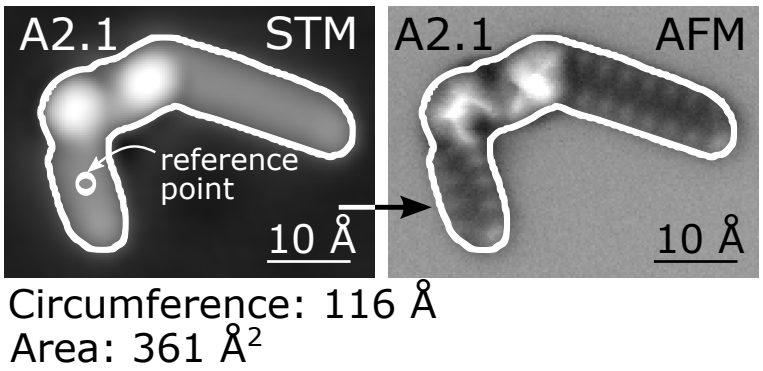


Figure S1. **STM area analysis example.** Measured area and circumference of molecule **A2.1** based on CO-tip STM data. The border was defined as the contour line at half the height of a reference point on the molecule. The border is shown as a white line, which is also overlaid on the corresponding CO-tip AFM measurement for comparison.

---

[1] Schuler, B., Zhang, Y., Collazos, S., Fatayer, S., Meyer, G., Pérez, D., Guitián, E., Harper, M. R., Kushnerick, J. D., Peña, D., and Gross, L., *Chem. Sci.*, **2017**, *8*, 2315–2320.

[2] Schuler, B., Liu, W., Tkatchenko, A., Moll, N., Meyer, G., Mistry, A., Fox, D., and Gross, L., Sep , **2013**, *111*, 106103.

[3] Pavliček, N., Mistry, A., Majzik, Z., Moll, N., Meyer, G., Fox, D. J., and Gross, L., *Nat. Nanotechnol.*, **2017**, *12*(4), 308–311.

[4] Mohn, F., Gross, L., and Meyer, G., *Appl. Phys. Lett.*, **2011**, *99*(5), 053106.

[5] van der Heijden, N. J., Hapala, P., Rombouts, J. A., van der Lit, J., Smith, D., Mutombo, P., Švec, M., Jelinek, P., and Swart, I., *ACS Nano*, **2016**, *10*(9), 8517–8525.

[6] Rogel, E., Miao, T., Vien, J., and Roye, M., *Fuel*, **2015**, *147*, 155–160.

[7] Rogel, E., Ovalles, C., Bake, K. D., Zuo, J. Y., Dumont, H., Pomerantz, A. E., and Mullins, O. C., *Energy Fuels*, **2016**, *30*(11), 9132–9140.

[8] Rogel, E. and Witt, M., *Energy Fuels*, **2016**, *30*(2), 915–923.

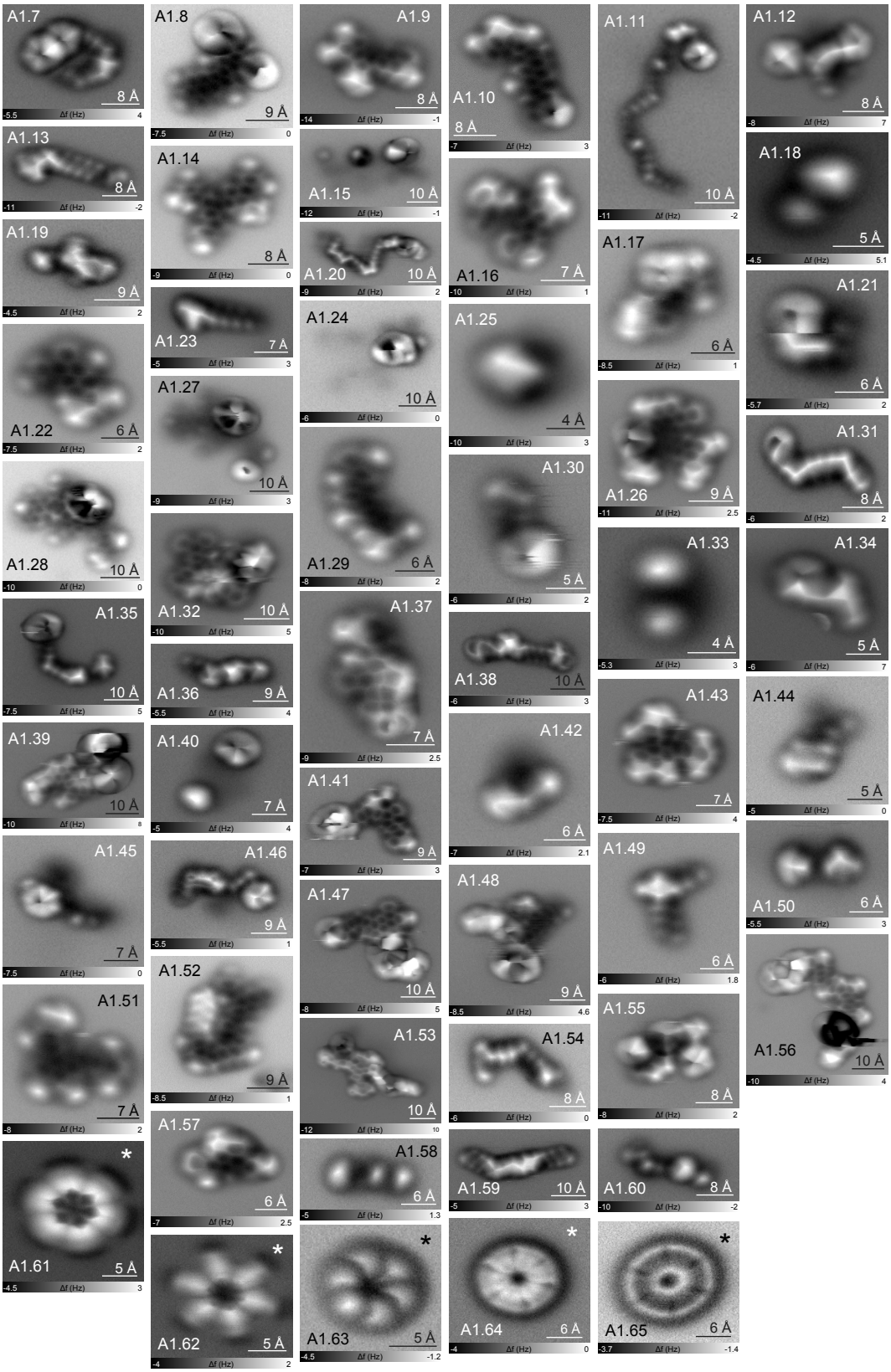


Figure S2. **Additional AFM measurements of sample A1.** The asterisk indicates a manipulation image where the molecule rotates around a pinning site.

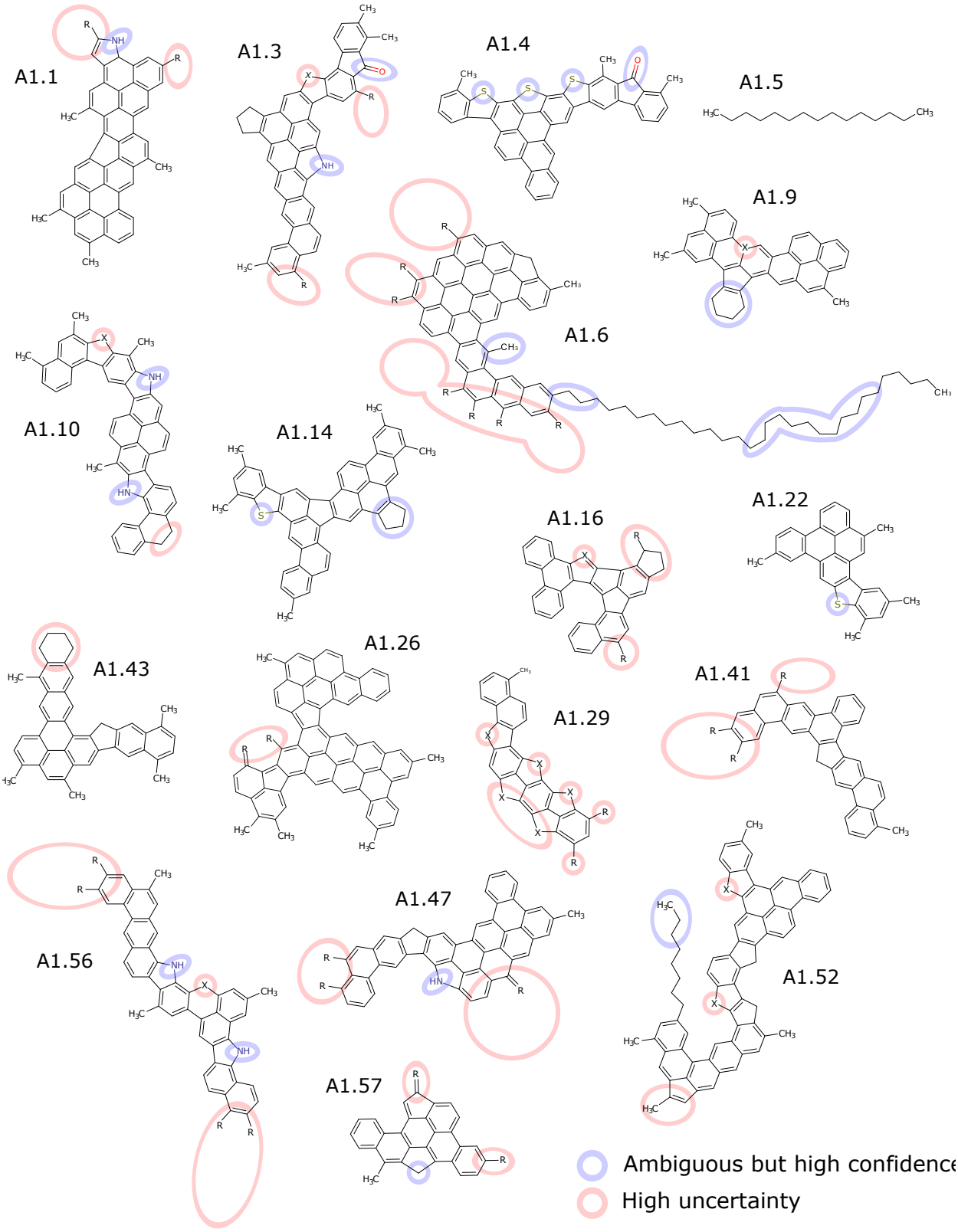


Figure S3. **Structure assignments in A1.** Structure proposals based on AFM measurements.

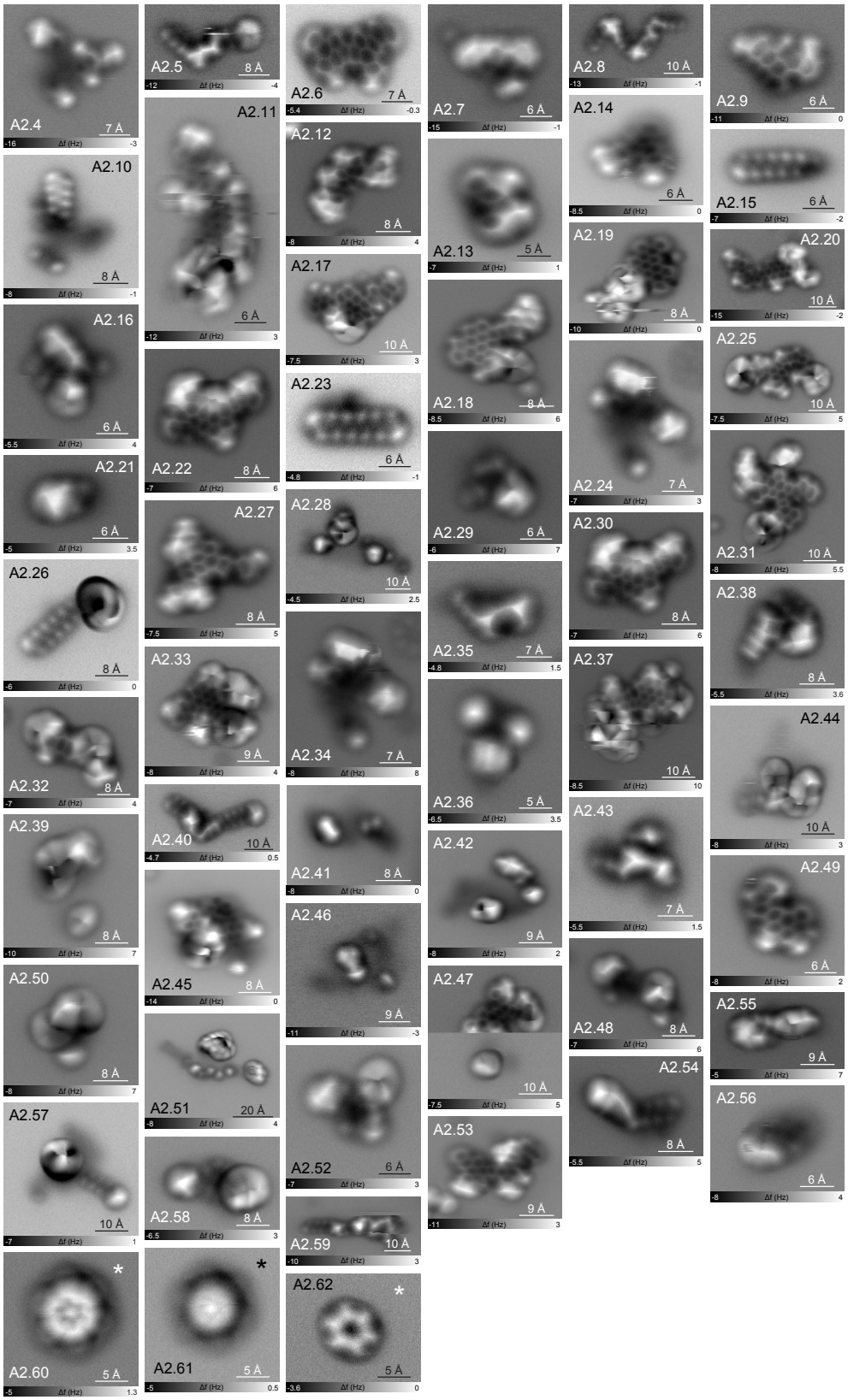


Figure S4. **Additional AFM measurements of sample A2.** The asterisk indicates a manipulation image where the molecule rotates around a pinning site.

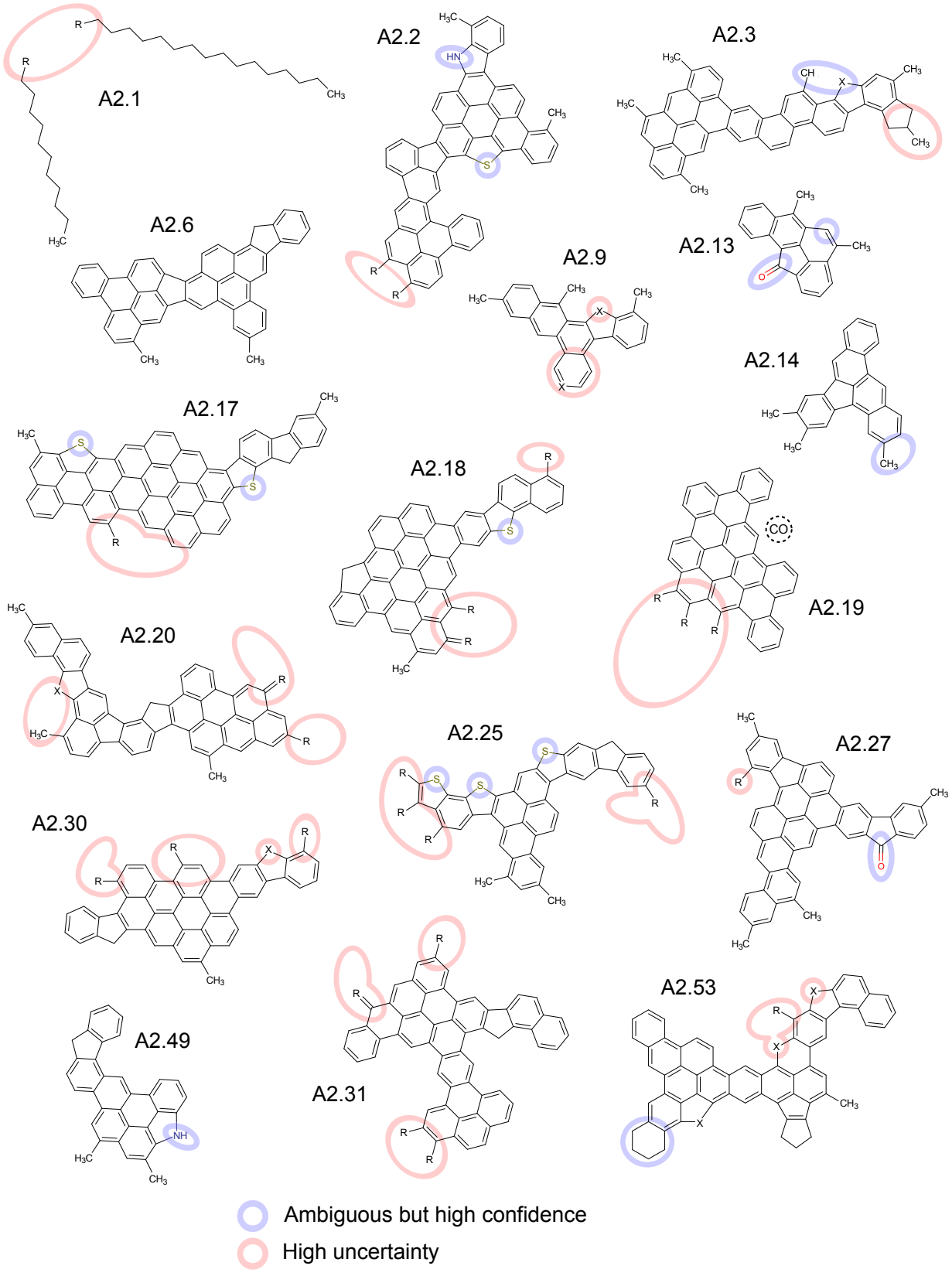


Figure S5. **Structure assignments in A2.** Structure proposals based on AFM measurements.

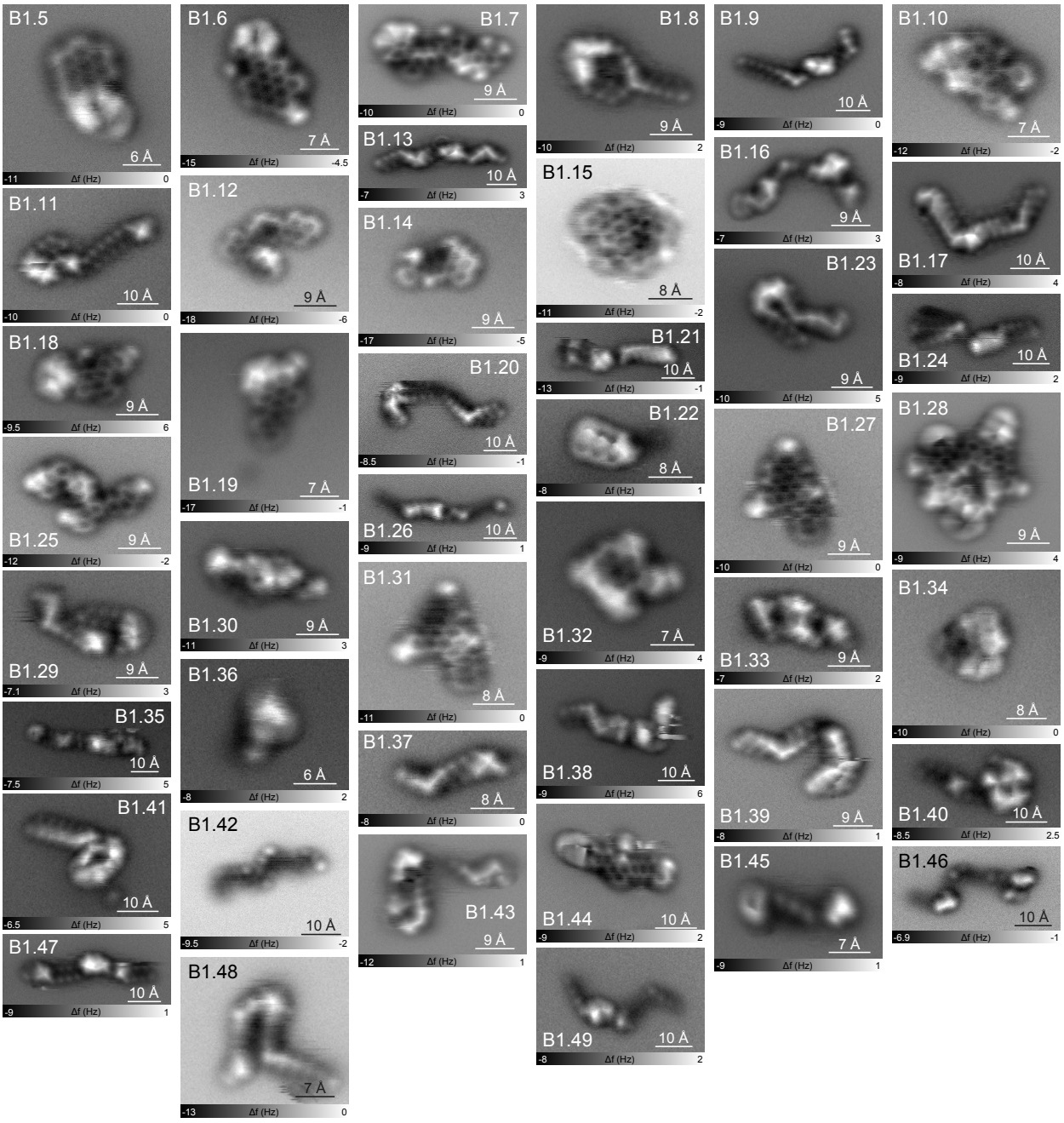


Figure S6. **Additional AFM measurements of sample B1.**

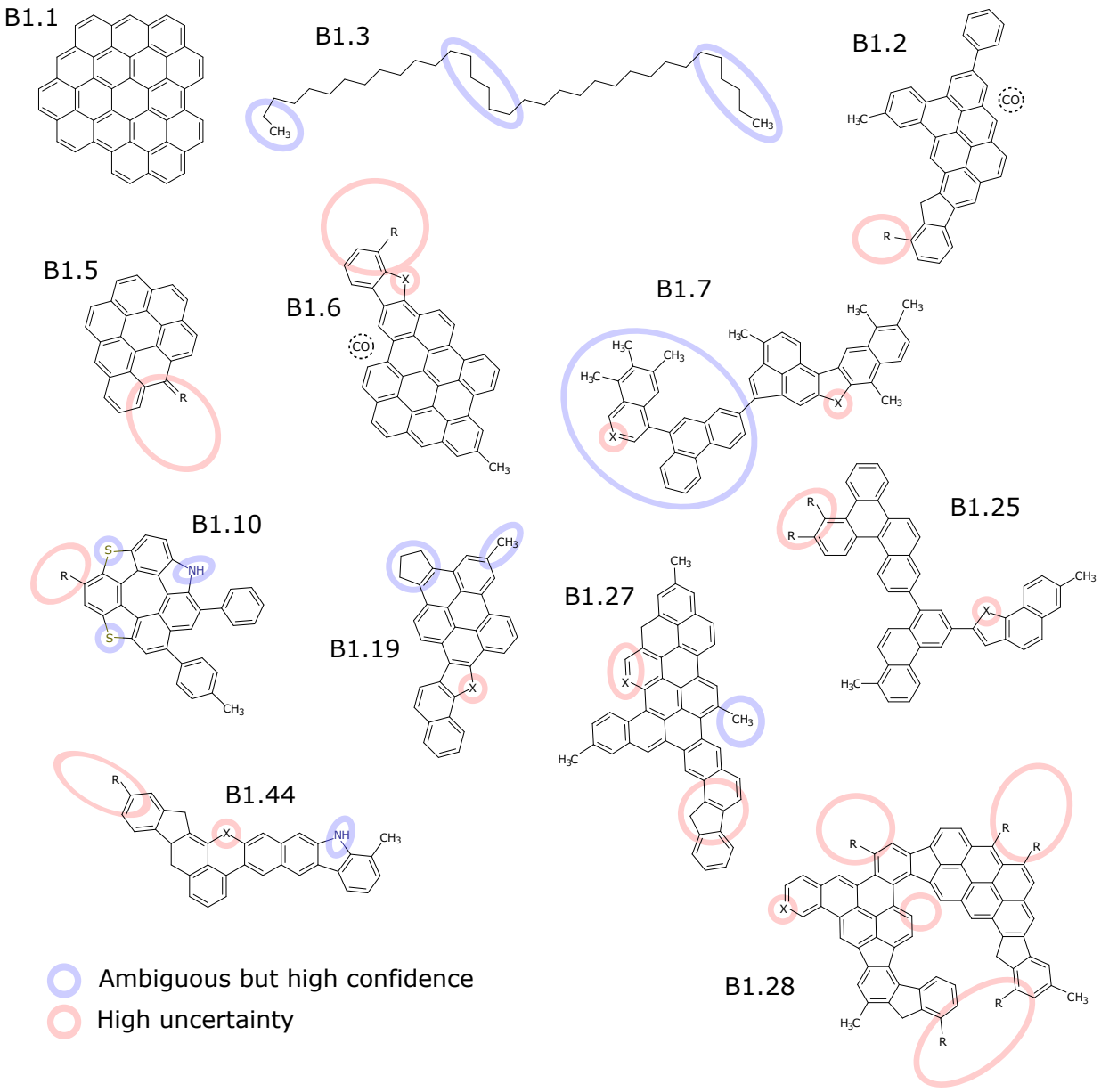


Figure S7. **Structure assignments in B1.** Structure proposals based on AFM measurements.

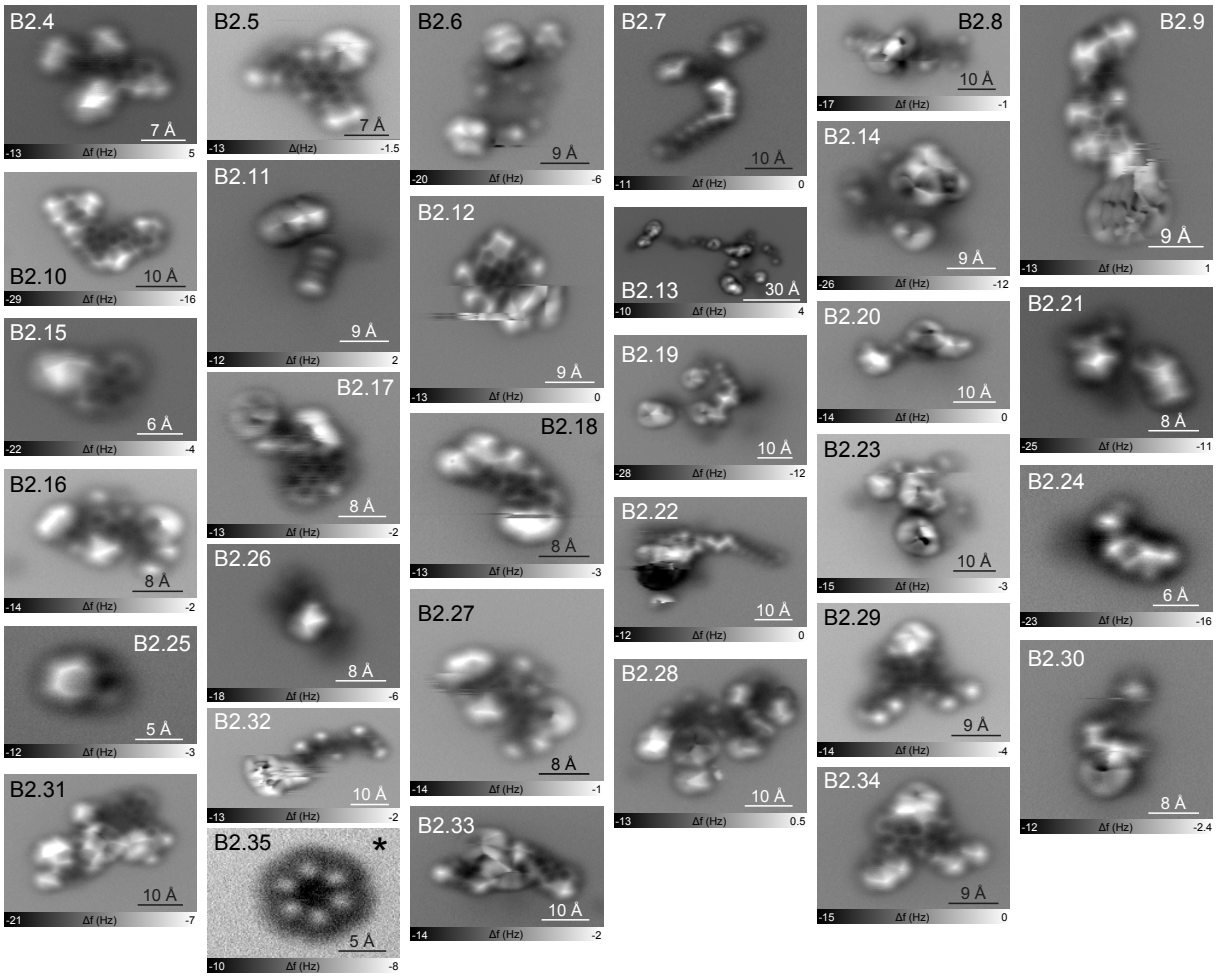


Figure S8. **Additional AFM measurements of sample B2.**

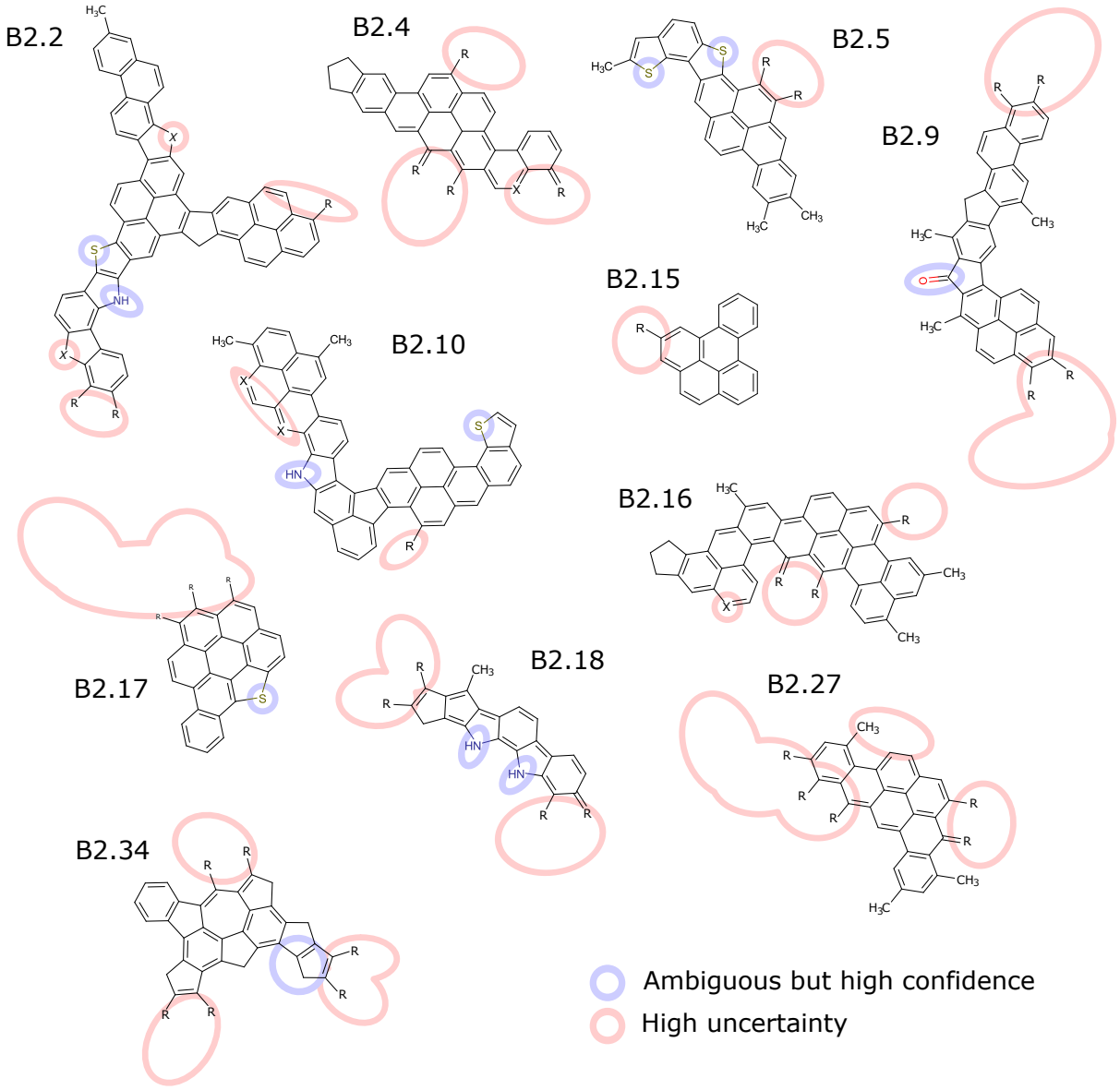


Figure S9. **Structure assignments in B2.** Structure proposals based on AFM measurements. The asterisk indicates a manipulation image where the molecule rotates around a pinning site.

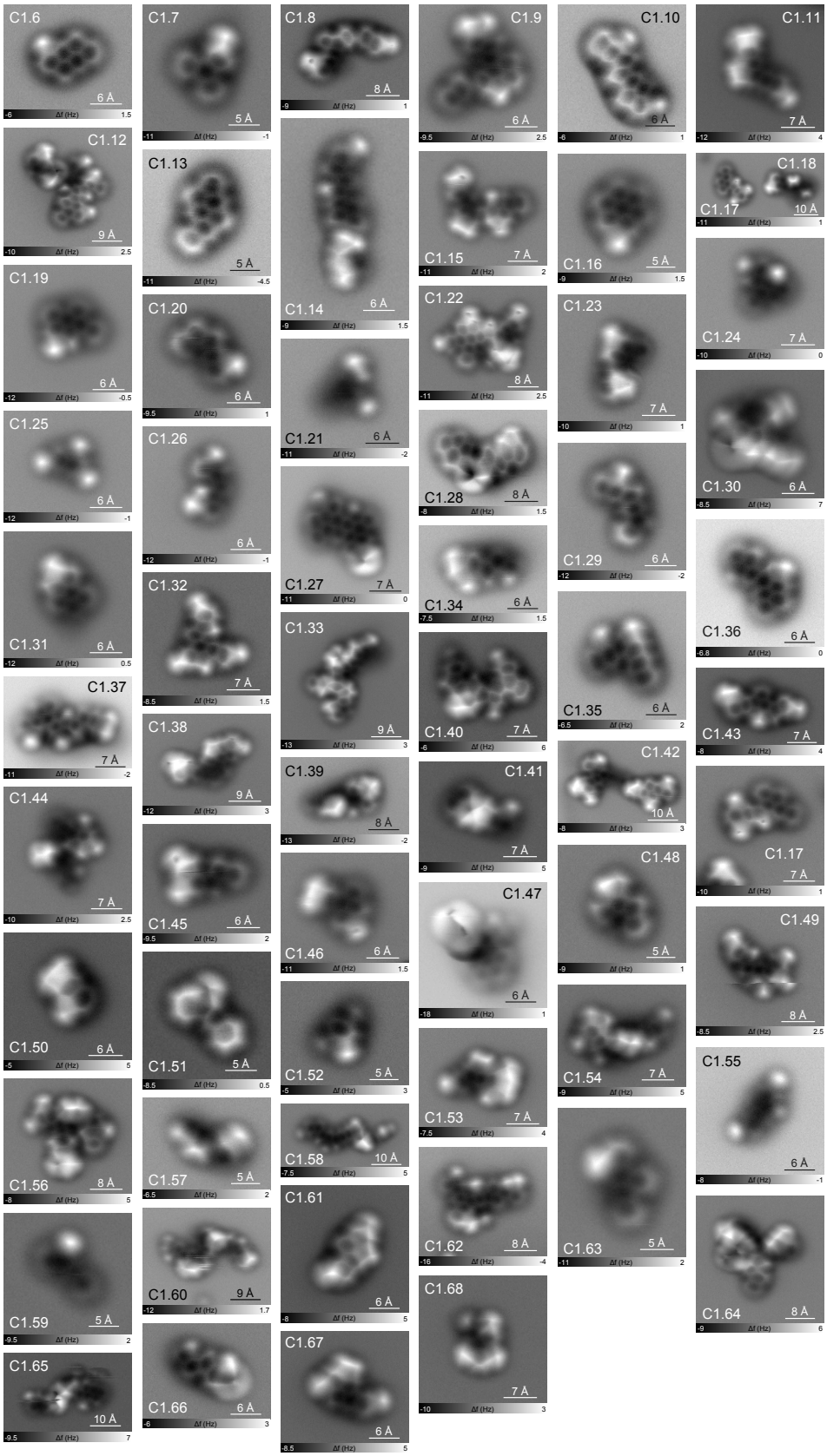


Figure S10. **Additional AFM measurements of sample C1.**

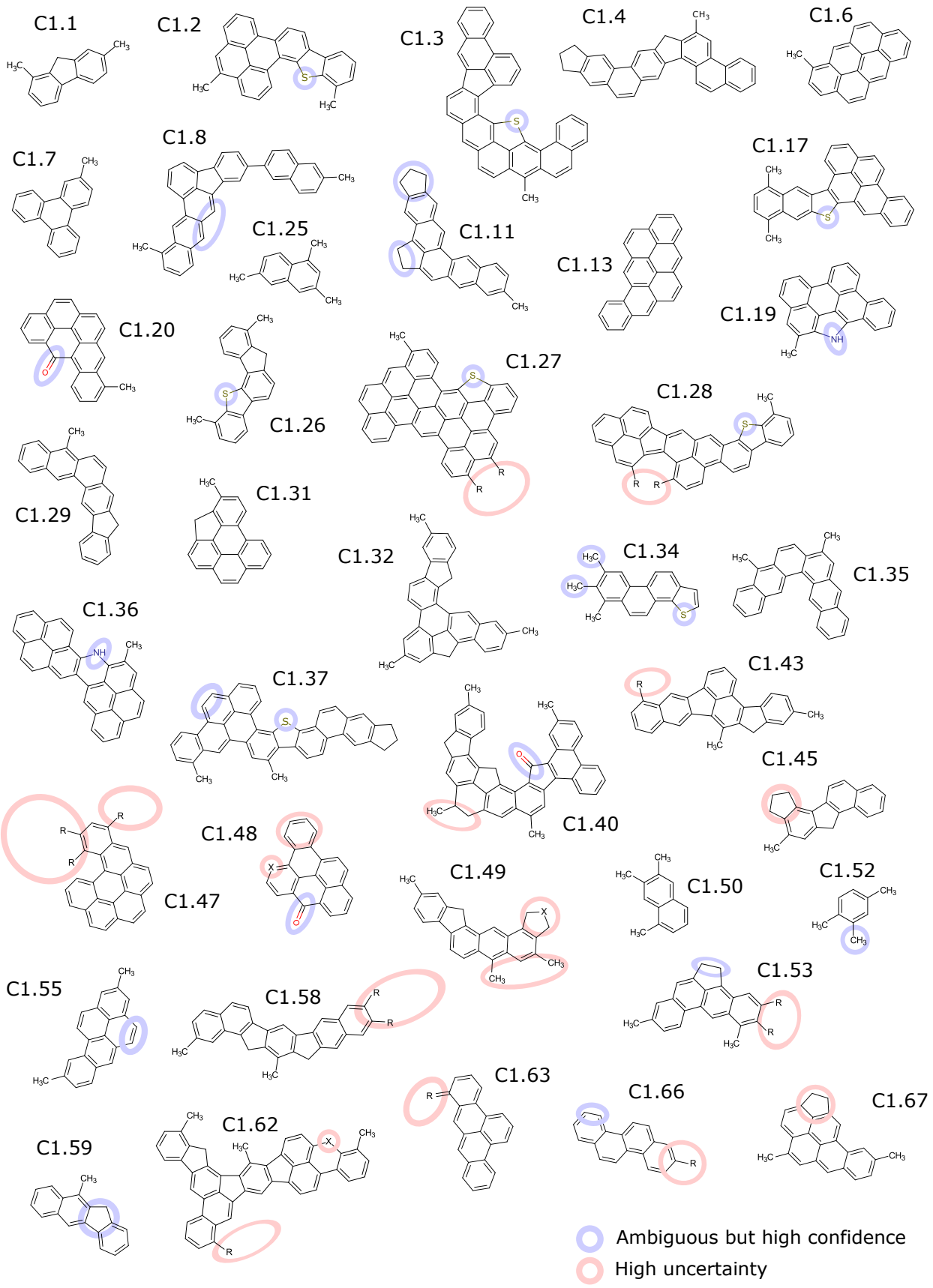


Figure S11. **Structure assignments in C1.** Structure proposals based on AFM measurements.

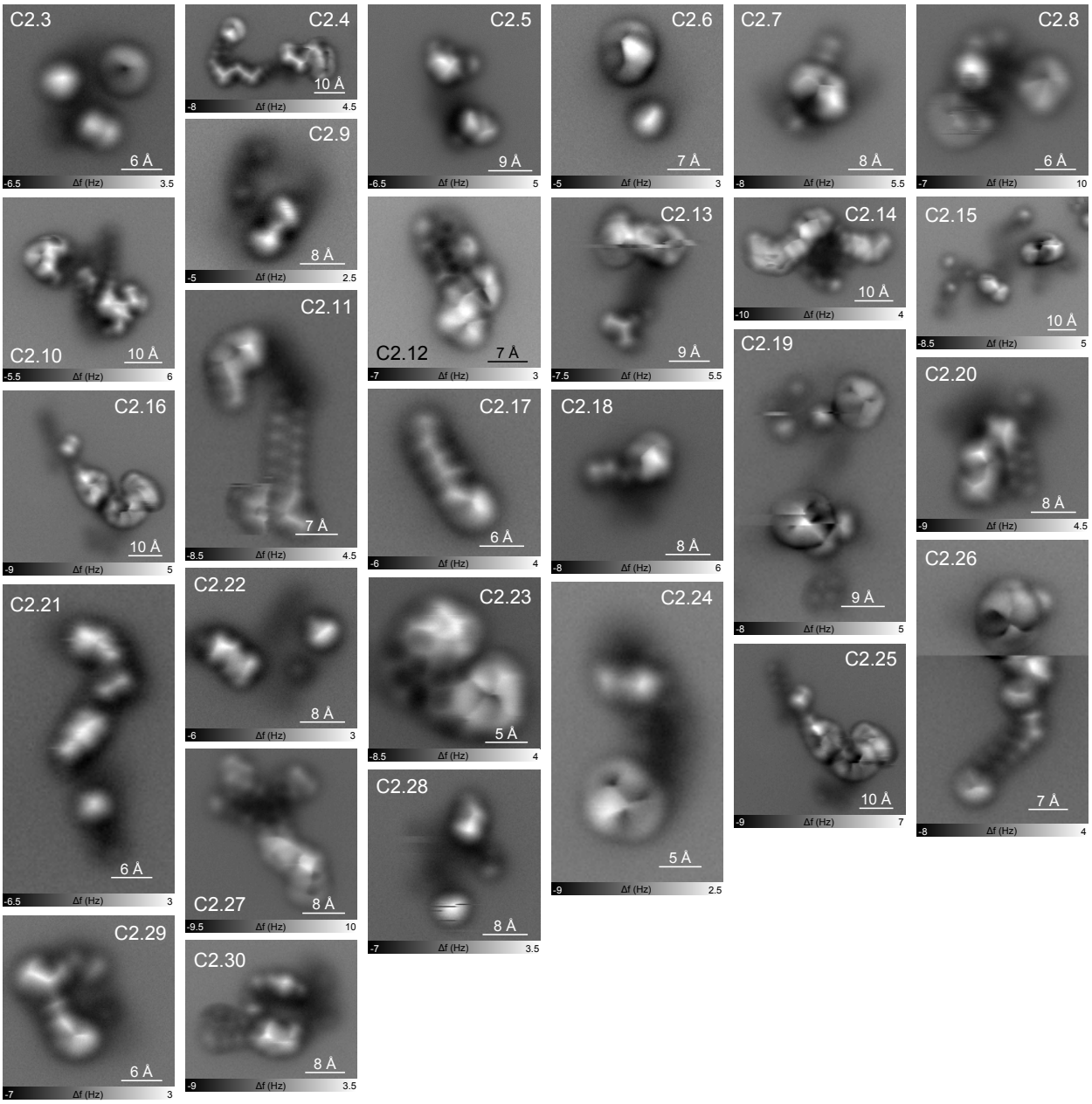


Figure S12. **Additional AFM measurements of sample C2.**

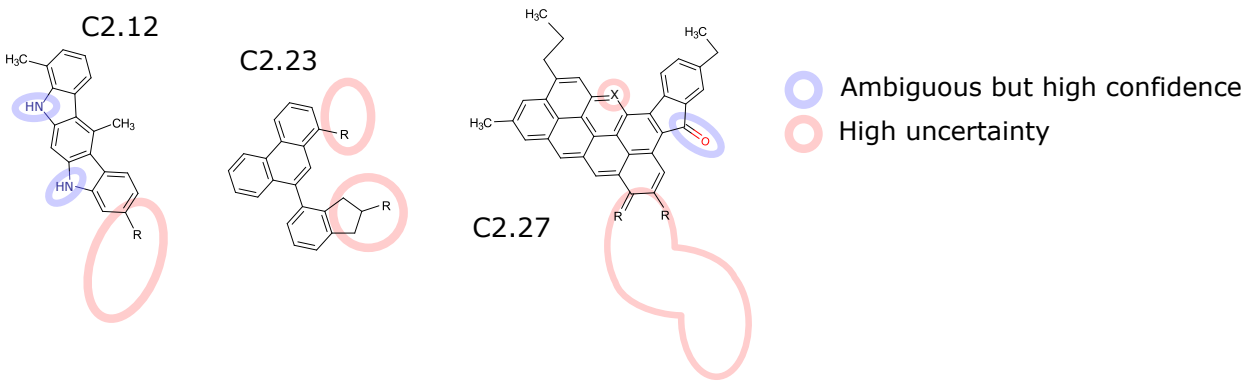


Figure S13. **Structure assignments in C2.** Structure proposals based on AFM measurements.

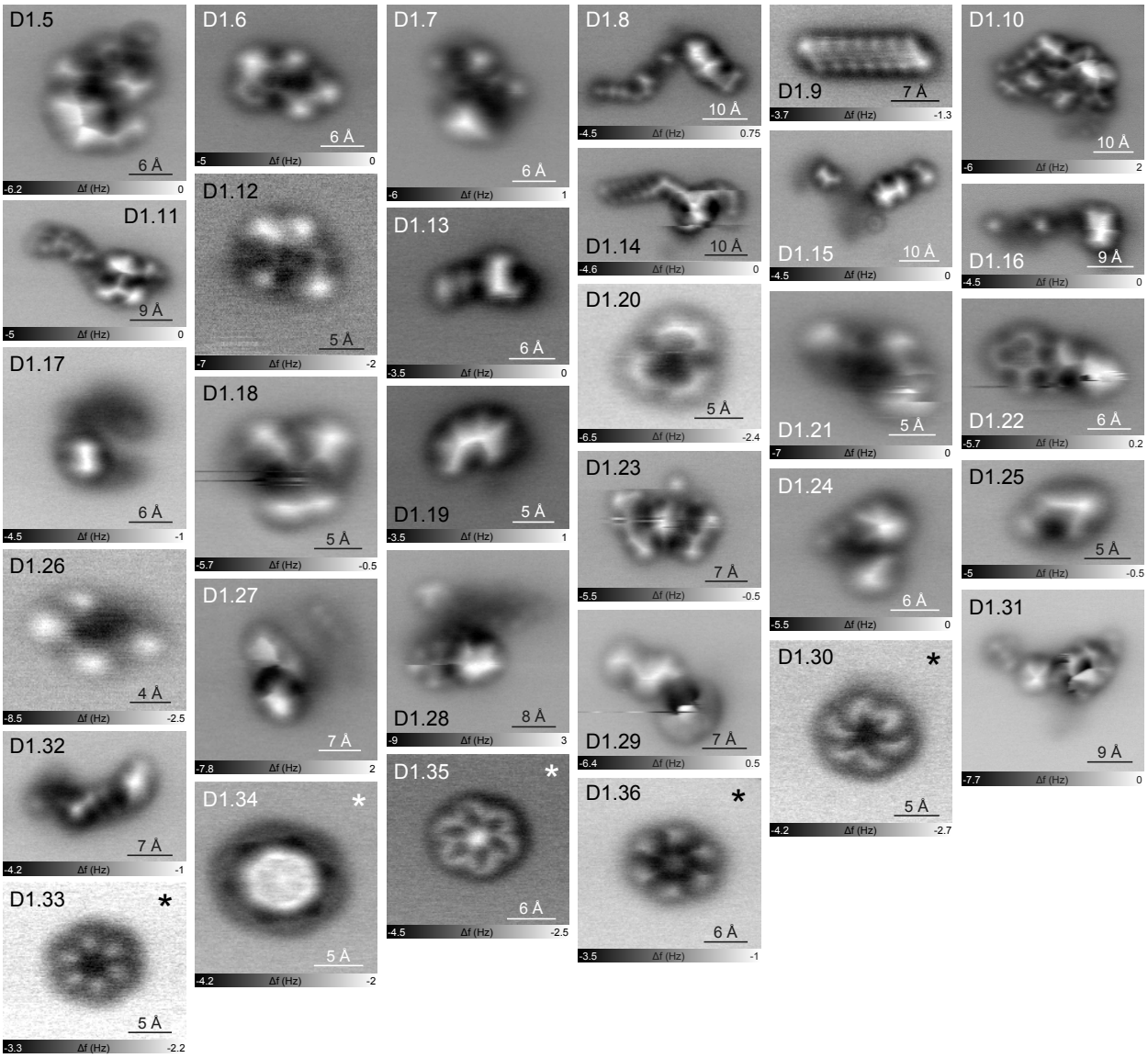


Figure S14. **Additional AFM measurements of sample D1.** The asterisk indicates a manipulation image where the molecule rotates around a pinning site.

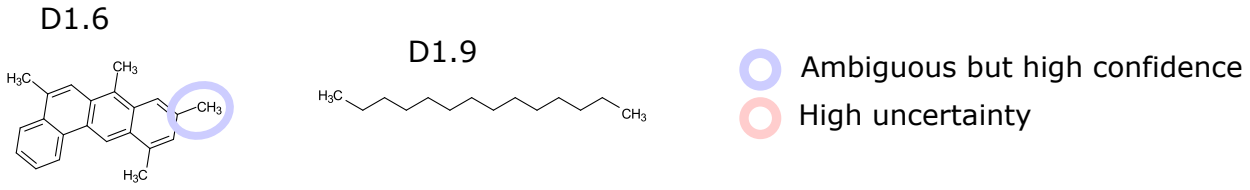


Figure S15. **Structure assignments in D1.** Structure proposals based on AFM measurements.

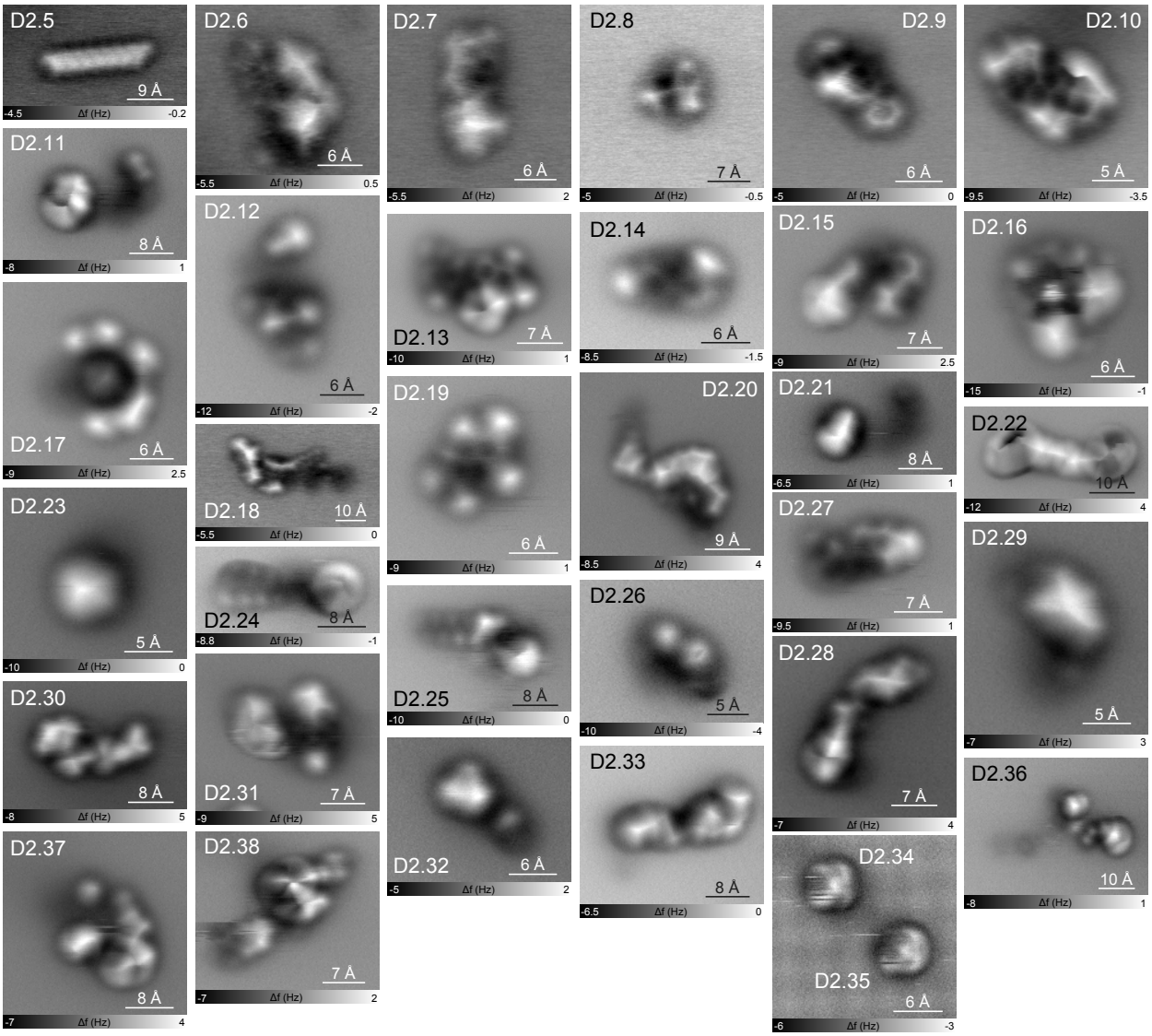


Figure S16. **Additional AFM measurements of sample D2.**

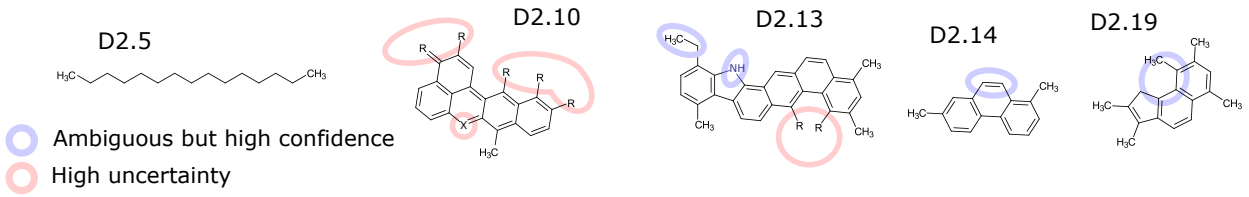


Figure S17. **Structure assignments in D2.** Structure proposals based on AFM measurements.

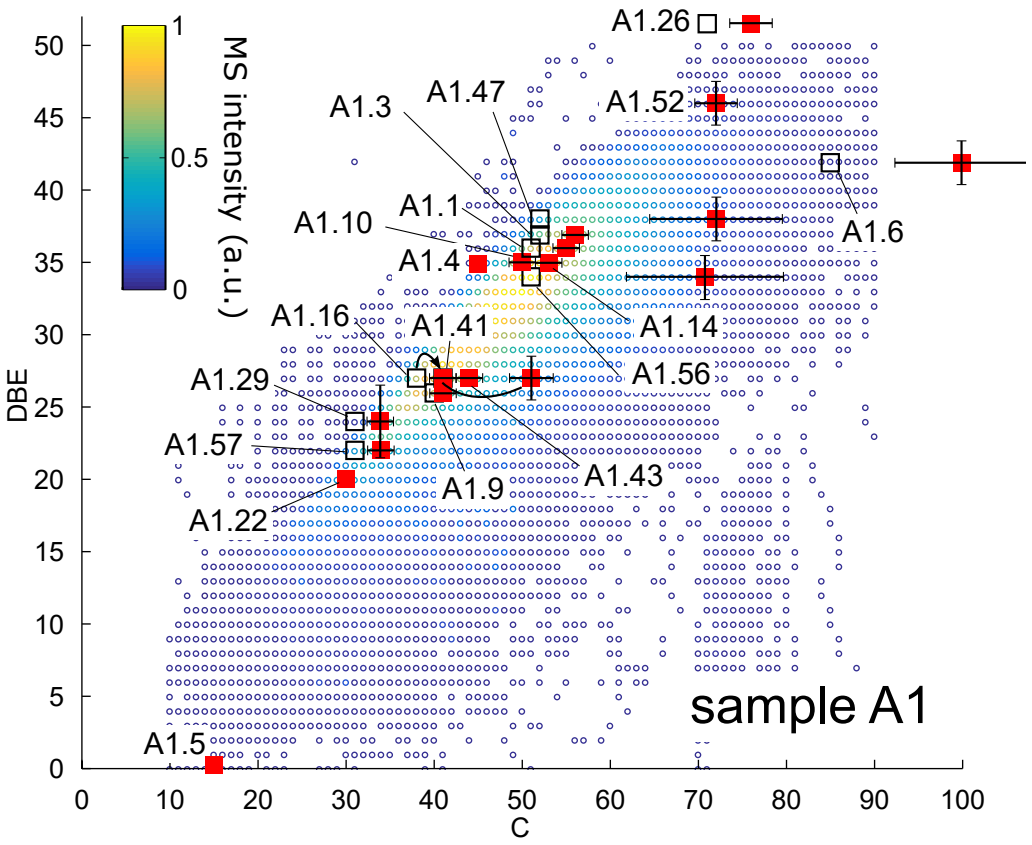


Figure S18. **Carbon number vs double bond equivalent (DBE) for sample A1.** Comparison of mass spectrometry data (colored scatter plot) [8] with structure proposals deduced by AFM. The black squares mark the position of the structure proposals in Figure **??** without considering unknown side groups. The red squares indicate the actual coordinate of the molecule (used in Figure 4a) after estimating the size of unknown side groups (based on AFM data), which leads to a horizontal shift in the C vs DBE plot.

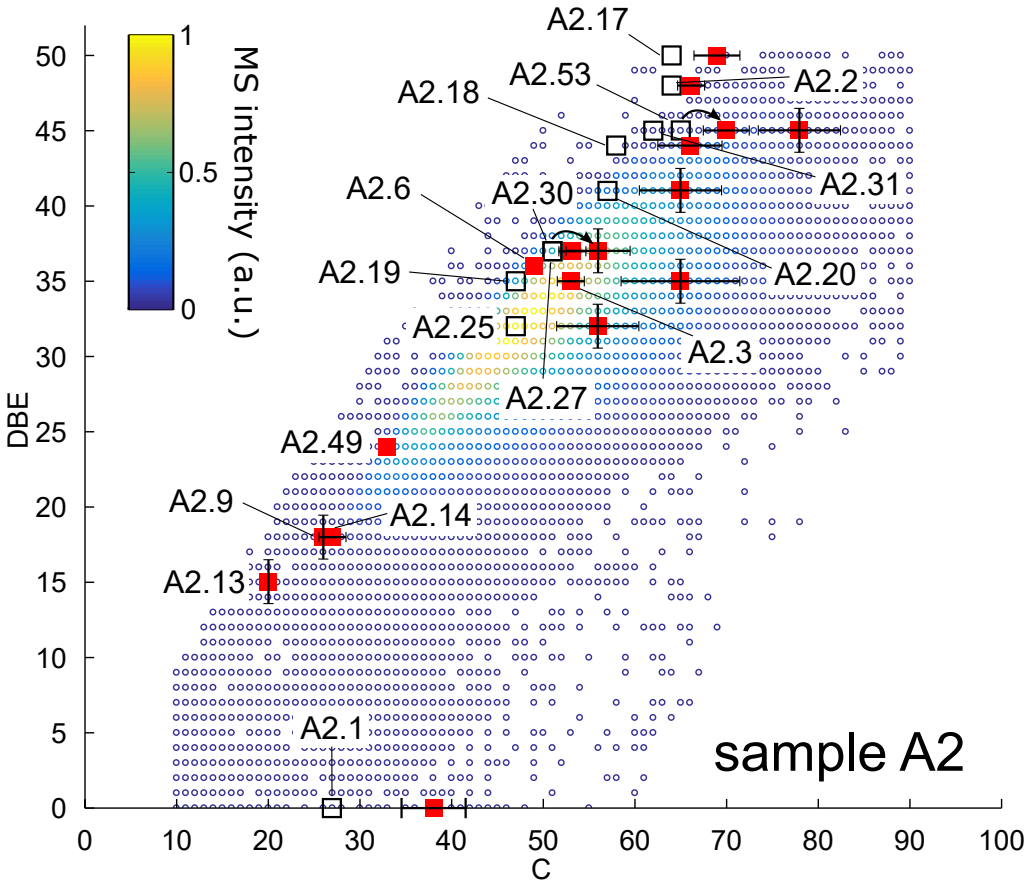


Figure S19. **Carbon number vs double bond equivalent (DBE) for sample A2.** Comparison of mass spectrometry data (colored scatter plot) [8] with structure proposals deduced by AFM. The black squares mark the position of the structure proposals in Figure **??** without considering unknown side groups. The red squares indicate the actual coordinate of the molecule (used in Figure 4b) after estimating the size of unknown side groups (based on AFM data), which leads to a horizontal shift in the C vs DBE plot.