Supplementary information

1. Characterization techniques and catalyst testing

X-ray diffraction patterns (XRD) were recorded on a Bruker D8 Advance diffractometer, with a Ni detector side filtered Cu K α radiation (1.5406 Å) over a 2 θ range of 5-60° and a position sensitive detector using a step size of 0.02° and a step time of 2 s. Synchrotron powder diffraction data were collected on the Materials Science Beamline at the Swiss Light Source (SLS) in Villigen, Switzerland (Table S1) [14]. The indexing of the XRD pattern was performed using Topas [15], which was also used for the Rietveld refinement [16] (Table S2). The structure drawings were produced using CrystalMaker [17] and the profile plots with the program ppp14 [18].

Scanning Electron Microscopy (SEM) micrographs were either acquired on a JEOL FEG 6700F microscope working at 9 kV accelerating voltage, or on a Hitachi S4800 SEM with a cold field emission gun. The Si / Al ratios (SAR) of the zeolites were determined by EDX analysis coupled with the SEM chamber, and separately by X-ray fluorescence using a SPECTRO XEPOS spectrometer equipped with a 50-Watt end-window X-ray tube to excite the samples. The target changer, with up to 8 polarization and secondary targets, offers many different excitation conditions ensuring optimum determination of all elements from Na to U. The detection system consists of a 10 mm² Si-Drift Detector (SDD) with Peltier cooling and spectral resolution of less than 155 eV at Mn K α is achieved. Measurements were conducted in a He gas atmosphere.

For TEM investigations, the samples were dispersed in CHCl3, sonicated for few seconds and drop-deposited on a copper TEM grid with a holey carbon film. Bright field TEM and HAADF (high-angle annular dark-field) STEM images were recorded using an image Cs corrected FEI Titan instrument operated at 300 kV. Further HRTEM images and elemental composition of the materials were acquired over a Hitachi HF-3300kV instrument. Likewise, the sample preparation was done by ultrasonically dispersion of the sample in ethanol, placing a drop over a copper grid coated with an amorphous carbon film, and then drying under open air conditions.

²⁷Al (I = 5/2) magic angle spinning nuclear magnetic resonance (MAS NMR) was carried out with a Bruker Avance II 400 spectrometer operating at $B_0 = 9.4$ T (Larmor frequency $v_0 = 104.2$ MHz) equipped with a Bruker 2.5 mm double channel probe.

Samples were spun at 25 kHz, and free induction decays (FID) were collected with a $\pi/12$ rf pulse (0.5 µs) and a recycle delay of 1 s. $[Al(H_2O)_6]^{3+}$ was used as external standard reference. Liquid-phase NMR (500 MHz, Bruker) was also used to investigate the interactions between silicon source, aluminium source and biomass species during zeolite syntheses. D-xylose and D-glucose were added to alkaline medium (in D₂O) along with silicon or aluminium source individually to avoid any formation of solid aluminosilicate species in aqueous phase.

The basic texture characteristics involving the apparent BET surface area S_{BET} , the mesopore surface area S_{meso} and the micropore volume V_{μ} were evaluated from the nitrogen physical adsorption-desorption isotherms measured at 77 K by means of ASAP2020M and ASAP2050M instruments (Micromeritics, USA). The high precision of pressure measurements was achieved by the use of the low pressure transducer with the capacity of 0.1 Torr. The specific surface area, S_{BET} , was evaluated from the nitrogen adsorption isotherm in the range of relative pressure $p/p_0 = 0.05-0.25$ (*p* is the adsorbate pressure and p_0 is the adsorbate vapor pressure at the measuring temperature) using the standard Brunauer–Emmett–Teller (BET) procedure [20].

Pyridine FT-IR Analysis of the acidic properties of samples was performed by adsorption of pyridine (Py) followed by infrared spectroscopy. Before analysis, the samples were pressed at ~1 ton.cm⁻² into thin wafers (10 mg.cm⁻²) and placed inside the IR cell. Prior to Py adsorption / desorption experiments, the wafers were activated by calcination under static condition at 723 K for 3 h in O₂ (1.6 × 10⁴ Pa) and then outgassed under secondary vacuum at 573 K (10⁻³ Pa) for 1 h. Those wafers were contacted at room temperature with gaseous Py (133 Pa) *via* a separate cell containing liquid Py. The spectra were then recorded following desorption from 423 and 573 K with a Bruker Vector 22 spectrometer (resolution 2 cm⁻¹, 128 scans). The reported spectra were obtained after subtraction of the spectrum recorded before Py adsorption. The amount of Brønsted and Lewis acidic sites titrated was obtained using a molar extinction coefficient value of $\varepsilon = 1.67$ cm.µmol⁻¹ for the vibrations of protonated pyridine (Py-H⁺) at ~1540 cm⁻¹ and of $\varepsilon = 2.22$ cm.µmol⁻¹ for the vibrations of coordinated pyridine (Py-L) at ~1455 cm⁻¹ according to Emeis [48].

n-hexane cracking experiments were performed in a high throughput unit (Vinci Technologies). Eight tubular quartz reactors assembled in parallel to each other (187 mm in length and 6 mm internal diameter) were used. After cationic exchange (353 K for 4h with a 1 M NH₄NO₃ aqueous solution, repeated three-times) and calcination 5h at 773 K in air, H-ZSM-5 zeolite was activated under nitrogen at 773 K for 2 h prior to catalytic evaluation at the same temperature, then the flow was shifted to an 11 %v/v of n-hexane in nitrogen (60 mL/min). The products were injected on line after 3, 17 and 32 min on stream using a GC-2010 Shimadzu chromatograph. The set-up details as well as chromatographic conditions were already reported elsewhere [21]. The activity was presented as the average result obtained at 3, 17 and 32 min on stream (in the absence of significant deactivation).

Prior to use in methanol-to-hydrocarbons reaction, ZSM-5 catalysts were calcined at 823 K for 15h in static air. H-ZSM-5 zeolites were sieved and particles < 250 μ m were used for the tests; 60 mg was introduced in a tubular quartz reactor, packed between quartz wool plugs. A methanol-saturated nitrogen-flow was fed to the reactor at 673 K (WHSV = 1.2 g_{MeOH}.(g_{cat}.h)⁻¹). Samples withdrawn at the reactor outlet were analyzed by GC (HP5890 Series II, Pona column, 50 m). The activity of the catalysts was expressed in terms of methanol and dimethylether conversion, calculated from the difference between inlet and outlet concentrations of methanol. The selectivity was defined as the mole ratio of each product referred to the moles of converted methanol and dimethylether. Conversions and selectivities were expressed after 1 h on stream, once the quasi steady-state was reached.

Periodic Density Functional Theory (DFT) calculations were performed with the VASP code [22] with the PBE (Perdew, Burke, and Ernzerhof) exchange-correlation functional [23]. Van der Waals corrections as proposed within the Grimme formalism (D2) [24] was applied. The projected augmented wave (PAW) method [25] was used to describe the core-electrons interactions and the plane wave basis set was limited to a kinetic cut-off energy of 800 eV. The convergence criterion for the electronic self-consistent field relaxation was fixed to 10^{-5} eV. Starting from the bulk cell parameters of MFI (obtained from International Zeolite Association database) [26], substitution of silicon positions by aluminum and introduction of a Na⁺ compensation cation was performed at various SAR, from [∞] to 2. Some trials were also performed with H⁺ as compensation cations, leading to similar conclusions. The bulk cell parameters were optimized at each SAR by energy minimization (criterion: 10^{-4} eV), all atoms being mobile. Such bulk cell optimizations were systematically followed by a constant lattice

geometry optimizations of ionic positions, performed using a conjugate-gradient algorithm, with a convergence criterion on forces of 0.02 eV $Å^{-1}$.

2. Thermogravimetric analyses

Thermogravimetric and differential analysis (TG/DTG) was determined by Iris TG 209 F1 thermal analyzer (Netzsch, Selb, Germany). The samples were placed in an 85 μ L crucible and analyzed in a nitrogen environment. 10 mg of samples were loaded into a crucible, heated from 35 to 800 °C at heating rate of 10 °/min under N₂ (60 mL/min). The TG/DTG of pristine sugar cane bagasse, hydrolyzed, washed and after hydrolysis are presented in Figures S1, S2, S3 and S4, respectively.

Un-treated bagasse (Fig. S1) showed the presence of extractives (around 200°C), hemicellulose (270°C), cellulose (340°C) and lignin (broad band from 300 to 600°C). The TG/DTG of un-treated sugar cane bagasse residues (after alkaline hydrolysis) confirms that extractives were completely removed and a sharp reduction in the hemicellulose contribution could be evidenced (Fig. S2). After washing the bagasse, the presence of extractives could be discarded. The shoulder in the peak around 320°C corresponds to the hemicellulose contribution, cellulose (peak dislocated to 360°C) and lignin. After reaction of the washed sugar cane, the residue showed the latter three components. In addition, the deconvolution of DTG profile suggests the presence of remaining hemicellulose fraction (Fig. S4).

By comparison of Fig S1 and S2, it appears that all extractives (peak at 200°C in Fig S1) were completely removed. Likewise, hemicellulose was partially removed (peak at 270°C in Fig S2). From Fig. S3 and S4, it can be observed that hemicellulose was also partially removed. The solution of un-treated sugar cane contains sugars (mainly sucrose) and hemicellulose and probably cellulose derivatives. Besides, the solution of washed sugar cane biomass may contain hemicellulose and probably cellulose derivatives and almost no sugars. These assumptions based on TG/DTG are verified in LC/MS/MS and GPC characterizations, as presented in further sections.



Figure S1. TG/DTG of un-treated sugar cane bagasse.



Figure S2. TG/DTG of the residue after hydrolysis of un-treated sugar cane bagasse at pH 10 during 24h at 15°C.



Figure S3. TG/DTG of washed sugar cane bagasse at 60°C.



Figure S4. TG/DTG of the residue after hydrolysis washed sugar cane bagasse at pH 10 during 24h at 25°C.

3. Gel permeation chromatography (GPC)

The molecular weights of components present in the alkaline hydrolysate from sugar-cane biomass were determined by gel permeation chromatography (GPC). The sample was prepared at a concentration equal to 2 mg / mL of potassium phosphate solution in Milli-Q water (1 mg / L concentration, pH 7.2 adjusted with sodium hydroxide solution). The same solution of potassium phosphate in Milli-Q water was used as a mobile phase. The analyzes were conducted at a temperature of 40 °C, at a

flow rate of 0.5 mL / min, operating at a pressure of 250 psi. Before being injected, the samples were filtered using Teflon (PTFE) membranes with pores 0.45 μ m in diameter. The injection volume was 200 μ L. The analyses were carried out in Phenomenex Chromatograph, model TS-430, equipped with 3 SHODEX-OH PAK columns and Viscotec refractive detector (VE3580 model). The equipment was calibrated with sulphonated polystyrene standards with polar masses in the range of 6 to 2300 KDa. The molar masses in Dalton are presented in Figure S5.



Figure S5. Molar masses detected in Dalton for the solution containing biomassderivate compounds obtained by hydrolysis at pH 10-11 of sugarcane bagasse at room temperature.

4. Liquid chromatography / mass spectrometry

Liquid chromatography analysis was performed on an Agilent 1200 Series LC system (Agilent, Waldbronn, Germany) consisting of a quaternary pump, an autosampler and a column oven compartment. The LC system was interfaced to an API 2000 ESI-MS/MS triple quadrupole mass spectrometry instrument (Applied Biosystems, Foster City, USA) equipped with a Turbo Ion Spray source operating in positive ion mode. Quantitative analysis was performed using the multiple reaction monitoring (MRM) mode of precursor/product ion transitions. D-Xylose, D-Arabinose, D-Fructose and D-Glucose (due to their prevalence in hemicellulose and cellulose) and saccharose were the selected monitored m/z ions. The resin-neutralized (cationic and anionic) and lyophilized extract were analyzed by LC-MS / MS using a Waters Xbridge Amida column (150 x 4.6 mm, 3.5 µm) at 35 °C in isocratic mode (mobile phase NH₄OH (A) and CH₃CN (B) and the gradient selected was the following: 0 min B80%, 20 min B60%, 25 min B 60%, 26 min B 80 and equilibrium in the initial conditions (during 10 min, flow 0.4 mL min⁻¹). The sample was diluted in acetonitrile/water (50/50 in volume) to obtain concentration around 100 µg mL⁻¹. Detection was carried out in MRM mode by ESI (-) MS / MS (API 2000 triple-quadrupole). In general, under the low energy analytical conditions, the main ions observed are those corresponding to the deprotonation of the analytes of [M-H +] -. 149/89 [MH +] - (D-xylose), 179/89 [MH +] - (D-glucose and D-fructose), 341/59 [MH +] - sucrose) and 163/59 (L-rhamnosis used as internal standard). Additionally, the samples were subjected to ESI-MS analysis in positive ionization mode giving (m / z) signals. In order to test the effects of adducts and energy on ionization and subsequent fragmentation, the extract was dissolved in CH₃CN / H₂O (85:10 in volume).

Figure S6 shows the chromatogram of those sugars used in the standard solution and Figure S7 the ESI(+) MS/MS spectrum of the standards. Similarly, the chromatogram and ESI(+) MS/MS for the solution containing biomass-derivate compounds obtained after the hydrolysis of washed biomass are presented in Figs. S8 and S9, respectively. Among the standard sugars, only sucrose was observed in the solution obtained after the hydrolysis of washed sugar cane, and in minor amount on the basis of the intensity (517) compared to the chromatogram in Fig. S6 (2.8e4). The detection and quantification limits for sucrose are 0.6 and 1.8 µg mL⁻¹, respectively and the correlation between sucrose / internal standard (raminose 5 µg mL⁻¹) x concentration is y = 0.3602x + 0.0902 (r=0.9972) in the range of $0.2 - 20 \ \mu g \ mL^{-1}$. The observed intensity of sucrose in the solution is below the detection limit. Likewise, ESI(+) MS/MS (Fig. S9) reveals a completely different pattern compared to Figure S7. The ESI(-) MS/MS of the solution of biomass-derivatives obtained by the washed sugar cane (Fig. S10) reveals the presence of several m / z fragments. The presence of two lignin-derivatives, coumaric- and ferulic acid derivatives could be assessed, as presented in Fig. S11 and S12, respectively. The ESI(+) and ESI(-) MS/MS of the solution containing biomass-derivatives obtained from un-treated biomass are presented in Figs. S13 and S14. Clearly, the presence of hexoses and sucrose could be ascertained.



Figure S6. LC–MS/MS chromatogram of standards: 1 (D-xylose), 2 (D-arabinose), 3 (D-fructose), 4 (D-glucose) and 5 (sucrose).



Figure S7. ESI(+) MS/MS spectrum of standards : D-xylose / arabinose, m / z 149), (glucose / fructose m / z 179) and sucrose, m / z 341 (CH₃CN-H₂O with 1% NH₄OH)].



Figure S8. LC–MS/MS chromatogram of the alkaline hydrolysate from washed sugar cane bagasse, (1)-sucrose, m / z 341.



Figure S9. ESI (+) MS / MS spectrum of alkaline hydrolysate obtained from washed sugar-cane bagasse. The fragmentation (1) m/z 360 indicates the adduct of ammonium (18) $[M + NH_4]^+$ and sucrose (341).



Figure S10. ESI(-) MS/MS spectrum of the alkaline hydrolysate from washed sugarcane bagasse. The fragmentation signal at m/z 163 (1) indicates the presence of conjugated base of cumaric acid and the signal at m/z 193 (2) indicates de presence of conjugated base of ferulic acid.



Figure S11. ESI(-) MS/MS fragmentation spectrum of substance detected at m/z 163 produced through alkaline hydrolysis of washed sugarcane bagasse (alkaline hydrolyzate from washed sugar-cane bagasse)



Figure S12. ESI(-) MS/MS fragmentation spectrum of substance detected at m / z 193 produced through alkaline hydrolysis of washed sugarcane bagasse (alkaline hydrolyzate from washed sugar-cane bagasse).



Figure S13. ESI (+) MS/MS spectrum of the alkaline hydrolysate of un-treated sugar cane bagasse at pH: 10 during 24h at 15° C. The signal (1) at m/z 203 indicates the presence of a hexose (D-glucose or D-fructose) adduct with sodium [M + Na]⁺; the signals at *m/z* 343 (2) and *m/z* 365 (3) indicate the presence of two cationic adducts of sucrose: (2) corresponds to the presence of protonated sucrose [M + H]⁺ and (3) indicates the presence of a sucrose adduct with sodium [M + Na]⁺.



Figure S14. ESI (-) MS/MS spectrum of the alkaline hydrolysate of un-treated sugar cane bagasse at pH 10 during 24h at 15°C. The signal at m/z 179 (1) indicates the presence of a deprotonated hexose (D-glucose or D-fructose). The signals at m/z 215 (2) and m/z 217 (3) indicate the presence of two hexose adductors with two isotopic species of chlorine ([M + ³⁵Cl]⁻ or [M + ³⁷Cl]⁻). The signal at m/z 377 (4) corresponds to the presence of a sucrose adduct with chlorine ([M + ³⁵Cl]⁻.

5. XRD refinement studies and electronic microscopy

Firstly, the geometry of both frameworks was optimized using a distance-least-squares procedure. All the atoms in tetrahedral coordination were modelled as Si. A difference electron density map of each of the two phases showed an electron density cloud that resembled the shape of a TPA⁺ cation located at the intersection of the straight with the sinusoidal 10-ring channels, being the position usually found for TPA⁺ cations occluded in ZSM-5. As an approximation, organic species where modelled as TPA⁺ cations, so in both phases a TPA⁺ cation was placed in that position, with the N atom on the mirror plane perpendicular to the b *axis*. The C atoms of two of the propyl moieties refined to positions very close to that mirror plane so their y coordinate was fixed (y= 0.25). The other two propyl moieties are disordered

over two equivalent positions with a maximum occupancy of 0.5. The H atoms in the SDA were taken into account by increasing the population factors of the C atoms according to the number of H atoms bonded to them. The occupancy of the TPA⁺ cation in each of the two ZSM-5 phases refined to 1 (4 TPA⁺/ u.c.). Subsequent differences map was featureless, so no extra-framework species other than organic SDA were located. Refinement of this model with geometric restraints on the bond lengths and angles yielded a reasonable geometry for the organic species and the framework. The relative amounts of each phase refined to 71.8 % of the main phase, which had the slightly larger cell, and 28.8 % of the second phase, which has the smaller cell.



Figure S15. Lorentzian fit profile for refined of ZSM-5 SAR 8 zeolite. Observed (black), calculated (red) and difference (blue) profiles for the profile fit of the sample ZSM-5 SAR 8. The sample contains two phases of ZSM-5, the main one (71.8 wt. %) with unit cell parameters a= 20.1080(3), b=19.9718(3), c= 13.4232(2) A and the secondary phase (28.2 wt. %) with cell parameters a= 20.0391(2), b=19.9248(1), c= 13.3990(1) A.

Synchrotron	SLS
facility	
Beamline	Material
	Science
Diffraction	Debye-Scherrer
geometry	
Detector	MYTHEN II
Monochromator	Si 111
Wavelength	0.7086 Å
Sample	rotating 0.3 mm
	capillary
Nominal step	0.004 °2 <i>0</i>
size	
Detector	4
positions	
Exposure time	4 x 80 s
2θ range	3.2–29.9 °2 <i>θ</i>

 Table S1. Synchrotron powder diffraction data collection parameters for ZSM-5 SAR8

Chemical composition	(C ₁₂ H ₂₈ N) ₂ [Si ₃₆ O ₇₂]	
Unit cell main phase		
а	20.1082(2) Å	
b	19.9721(3) Å	
С	13.3990(1) Å	
Unit cell second phase		
a	20.0392(2) Å	
b	19.9248(1) Å	
С	13.4229(2) Å	
Space Group	Pnma	
Data points	6986	
Contributing reflections	1184	
Geometric restraints	436	
T-O	1.61(1) Å	96
O-T-O	109.5(20)Å	144
T-O-T	145(8)°	54
C-N	1.51(1) Å	8
C-C	1.55(1) Å	16
C-N-C	109.8(20)°	12
C-C-C	111.0(20)°	8
C-C-N	116(20)°	8
Parameters		
structural	293	
profile	16	
R_F (main/second phase)	0.046/0.058	
R_{wp}	0.091	
R _{exp}	0.114	

 Table S2. Crystallographic data from the Rietveld refinement of ZSM5 SAR8^a

^{*a*} The numbers given in parentheses are the esd's in the units of the least significant digit given. Each restraint was given a weight equivalent to the reciprocal of its esd.



Figure S16. SEM image of as-obtained sample after 24 h autoclaving at 170°C.



Figure S17. XRD pattern of as-obtained sample after 24 h autoclaving at 170°C.



Figure S18. XRD patterns of ZSM-5-a and ZSM-5-b



Figure S19. SEM images of ZSM-5-a and ZSM-5-b samples. The individual crystals exhibit a nano French fries morphology forming micron-sized assemblies.



Figure S20. TEM image of ZSM-5-b sample assessing the presence of inter-crystalline mesoporosity

6. DFT calculations

The formation energy $E_{(n_{Al})}^{f}$ of a given cell containing $n_{Al} = n_{Na}$ aluminum atoms, of energy $E_{(n_{Al})}$, from the purely siliceous cell (SAR = ∞ containing $n_{Si} = 96$ and $n_{Al} = 0$, energy: $E_{(0)}$) and from the highest substituted cell (SAR = 2, $n_{Si} = 64$ and $n_{Al} = 32$, energy: $E_{(32)}$) was quantified for any cell by the formation energy defined in equation 1.

$$E_{\binom{n}{n_{Al}}}^{f} = E_{\binom{n}{Al}} - \frac{n_{Al}}{32} \times E_{(32)} - \left(1 - \frac{n_{Al}}{32}\right) \times E_{(0)}$$
Equation 1

Figure S22 depicts the evolution of $E_{(n_{Al})}^{f}$ as a function of n_{Al} . Note that for each SAR, an arbitrary configuration in terms of Al and Na location was chosen as starting point of the optimizations, which does not guarantee that the most stable one was found.



Figure S21. Evolution of the formation energy of the cells, as calculated by DFT, as a function of the SAR or the number of aluminum atoms within the cell.



Figure S22. Water crystallization in the presence of BSST: a) reference at pH=11; b) bagasse hydrolysate at pH 14 ; c) bagasse hydrolysate at pH=11; d) wood lignin solution at pH=11. Amphiphilic biomass hydrolysate seems to inhibit or direct the crystalline growth of ice in a peculiar manner. Furthermore, a structuring role of both bagasse hydrolysate and wood lignin can be evidenced during ice crystallization. Similarly to AFP solutions, a growth in the form of needles was observed.



Figure S23. ²⁹Si NMR ; a) sodium silicate, b) $Na_2SiO_3 : D$ -xylose, 1:2.75, c) $Na_2SiO_3 : D$ -glucose, 1:2.75. The presence of silicic acid (Si(OH)₄) and its oligomers can be detected in the blank spectrum (a). After xylose addition, the appearance of peaks at - 102 ppm (spectrum b) confirms the presence of penta-coordinated Si species [36a]. In stark contrast, the addition of glucose did not lead to the appearance of those signals, thus indicating that C₆ sugar barely interact with silicate species.



Figure S24. ²⁷Al NMR ; sodium aluminate (bottom); NaAlO₂ : D-xylose 1:4.4 (middle); NaAlO₂ : D-glucose 1:4.4 (top). The signal at 81 ppm in the blank spectrum corresponds to Al(OH)₄⁻. After either C₅ or C₆ sugar addition, the latter signal vanished. This has been explained by a broadening until complete disappearence for Al(III) - gluconate species [45].



Figure S25. FT-IR of pyridine adsorption on ZSM-5 SAR8