Supportive information

Evolution of Nucleophilic High-molecular-weight Organic Compounds in Ambient Aerosols

Chen He[1], Hanxiong Che[2], Zier Bao[2], Yiliang Liu[2], Qing Li[2], Miao Hu[3], Jiawei Zhou[2], Shumin Zhang[4], Xiaojiang Yao[2], Quan Shi[1], Chunmao Chen[1], Yan Han[2], Lingshuo Meng[2], Xin Long[2], Fumo Yang[5]and Yang Chen[2]*

[1]State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing 102249, China.

[2]Research Center for Atmospheric Environment, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China.

[3] CNOOC Institute of Chemicals & Advanced Materials, Beijing 102200, China.

[4] Institute of Basic Medicine and Forensic Medicine, North Sichuan Medical College, Nanchong 637000, Sichuan, China.

[5] Department of Environmental Science and Engineering, College of Architecture and Environment, Sichuan University, Chengdu 610065, China

Correspondence to: Yang Chen (chenyang@cigit.ac.cn)

1. Methods

1.1 FT-ICR MS analysis

FT-ICR MS was calibrated using a reference list formed by the manually assigned known formula in Data Analysis. Mass peaks with a signal-to-noise (S/N) ratio ≥ 4 were exported to a data sheet. Data analysis was performed using in-house software. The assigned formulae were limited in the following elemental composition: $^{12}C_{0-100}$, $^{1}H_{0-200}$, $^{14}N_{0-10}$, $^{16}O_{0-20}$, and $^{32}S_{0-2}$. The formula assignment section set the mass accuracy window was set to 1.0 ppm in formula assignment section. All elemental formulae should meet basic chemical criteria: (1) the number of H atoms should be at least 1/3 that of C atoms and cannot be greater than that of $2C + N + 2$; (2) the sum number of N and H atoms should be even; and (3) the H/C and O/C value should be restricted to be less than 3 and 1.5, respectively.

These compound groups were: ; lipids (O/ C= 0−0.2, H/C = 1.7−2.2); Aliphatics (O/C = 0.2−0.6, H/C = 1.5−2.2); lignin (O/C = 0.1−0.6, H/C = 0.6−1.7, $AI_{mod} < 0.67$); carbohydrates (O/C = 0.6−1.2, H/C = 1.5−2.2); tannins (O/C = 0.6−1.2, H/C = 0.5−1.5, $AI_{mod} < 0.67$); unsaturated hydrocarbons (O/C = 0−0.1, H/C = 0.7−1.5); saturated compounds (H/C>2.0); polycyclic aromatics (PCAs, $AI_{mod} \geq 0.67$); polyphenols (0.67 $\geq AI_{mod} \geq 0.50$); and highly unsaturated compounds ($AI_{mod} < 0.50$, H/C < 1.5)

The double bond equivalent (DBE) and modified aromaticity index (AI_{mod}) value are calculated as Eqs. S1 and S2 (Koch and Dittmar, 2006; Koch and Dittmar, 2016):

$$DBE = C - 0.5H + 0.5N + 1 \qquad (S1)$$
$$AI_{mod} = (1 + C - 0.5O - S - 0.5H)/(C - 0.5O - S - N - P) \qquad (S2)$$

40   A Kendrick mass defect (KMD) analysis is used to provide an overview of the

41   composition of the identified formulae in NFC samples. The KMD calculations are

42   shown in Eqs. (S3 and S4) (Sleighter and Hatcher, 2007):

43   $$\text{Kendrick mass (F)} = (\text{observed mass}) \times \frac{\text{nominal mass(F)}}{\text{exact mass(F)}} \qquad \text{(Eq. S3)}$$

44   $$\text{KMD} = \text{nominal mass} - \text{Kendrick mass (F)} \qquad \text{(Eq. S4)}$$

45   where F stands for the functional groups or repeated groups, including COO, OO,

46   CH_2O, and H_2O.

47

48   The value of X_c is used to characterize aromatic and polyaromatic compounds in highly

49   complex compound mixtures. X_c normally ranges from 0 to 3.0 and is calculated as Eqs.

50   (S5) (Yassine et al., 2014):

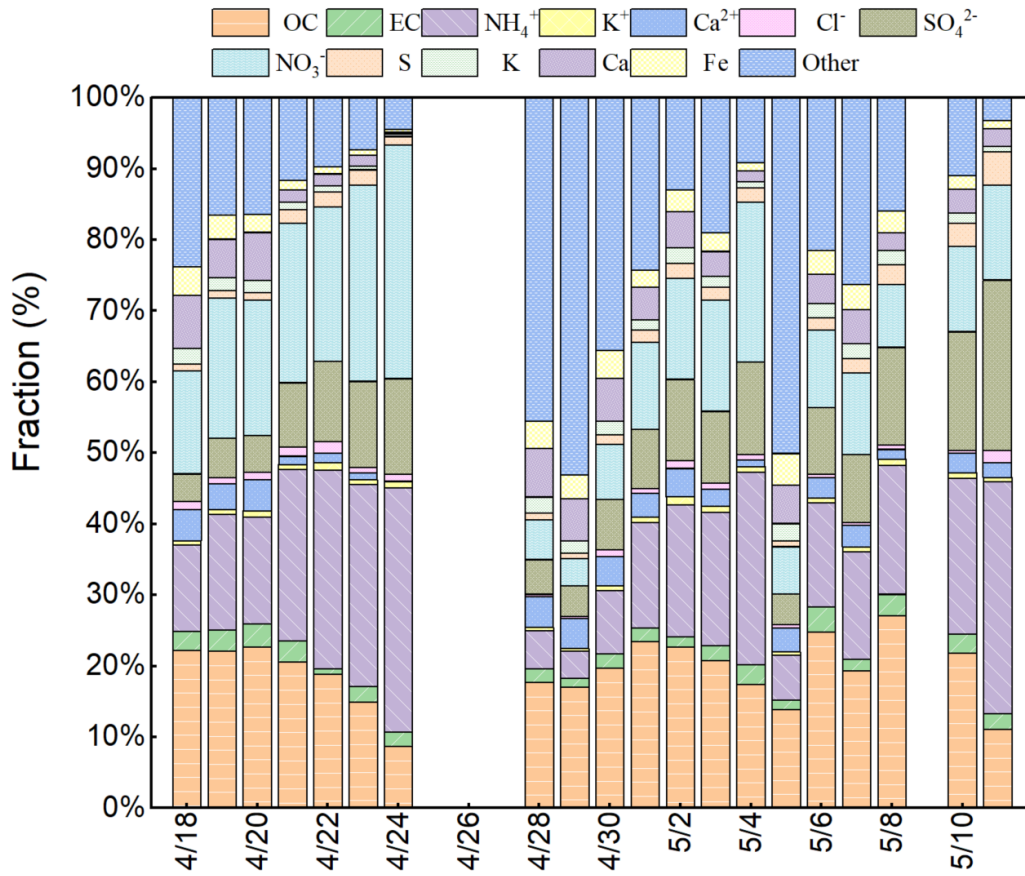
51   $$X_c = \frac{3(\text{DBE} - mN_O - nN_S) - 2}{\text{DBE} - mN_O - nN_S} \qquad \text{(Eq. S5)}$$

52   If DBE $\leq mN_O + nN_S$, then $X_c = 0$. For (-)ESI and (+)ESI mode, we used m=n=0.5 and

53   m=n=1.0 for the X_c in this study.

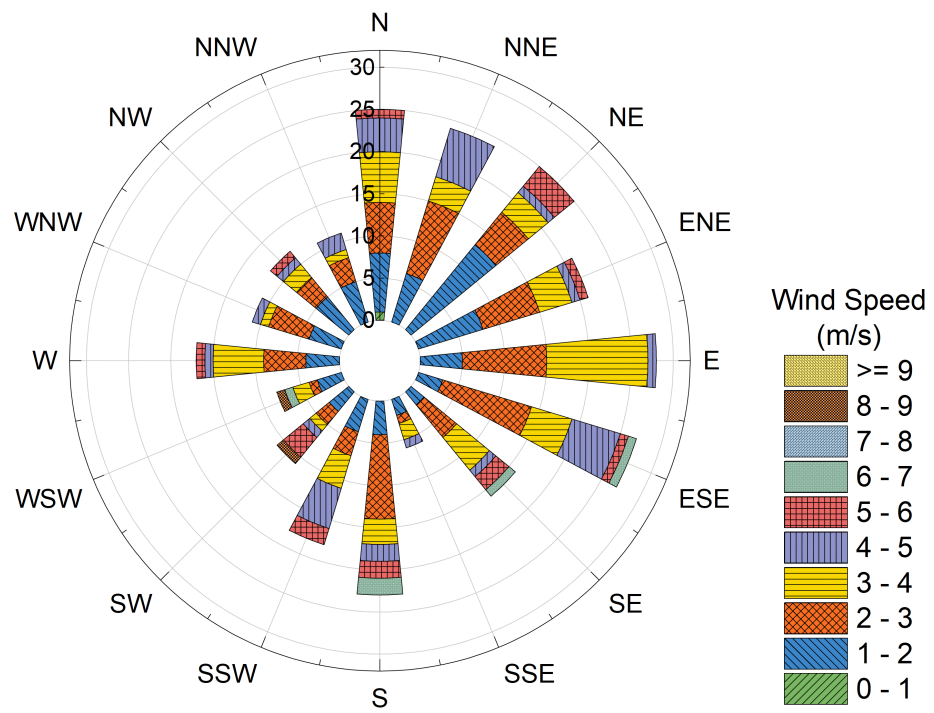
54

55   Figures

56

57

58   Figure S1. Fraction of OC, EC, water-soluble ions, and crustal elements during the

59   sampling period.

60

61

62

63   Figure S2. Windrose plot during the sampling period.

64

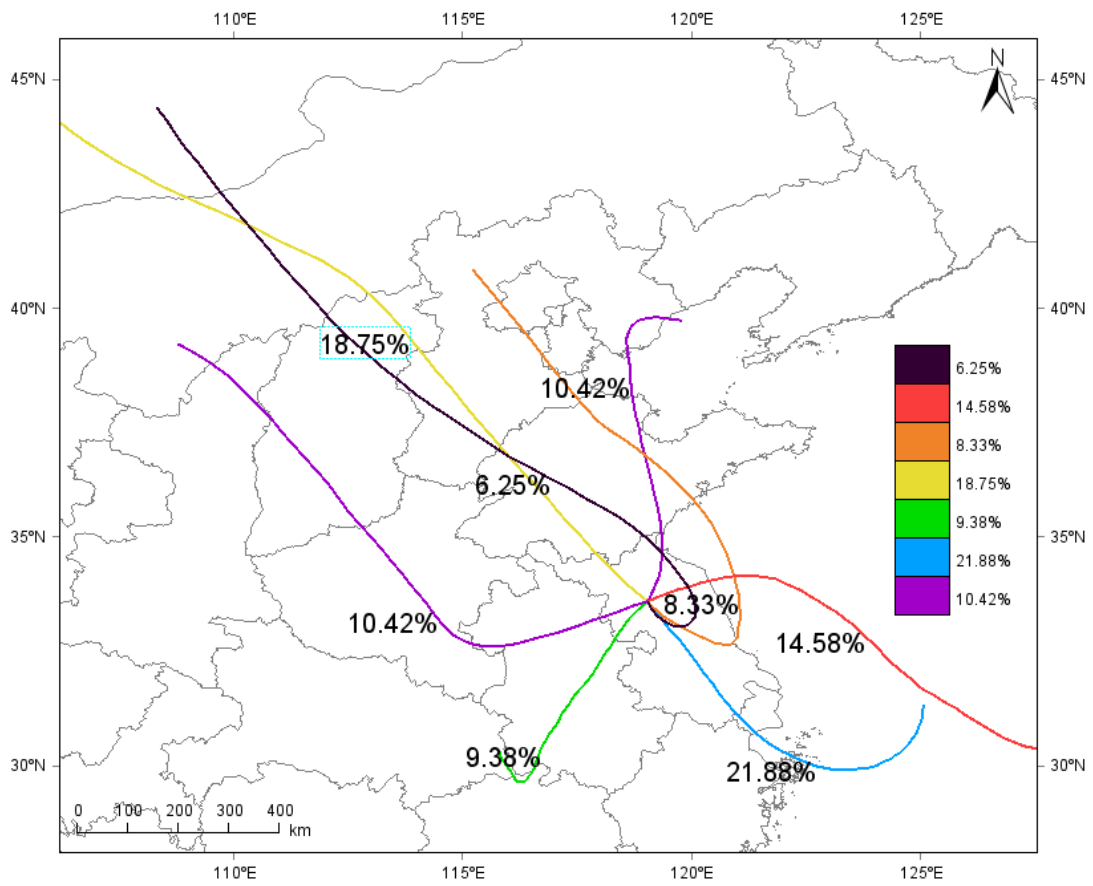
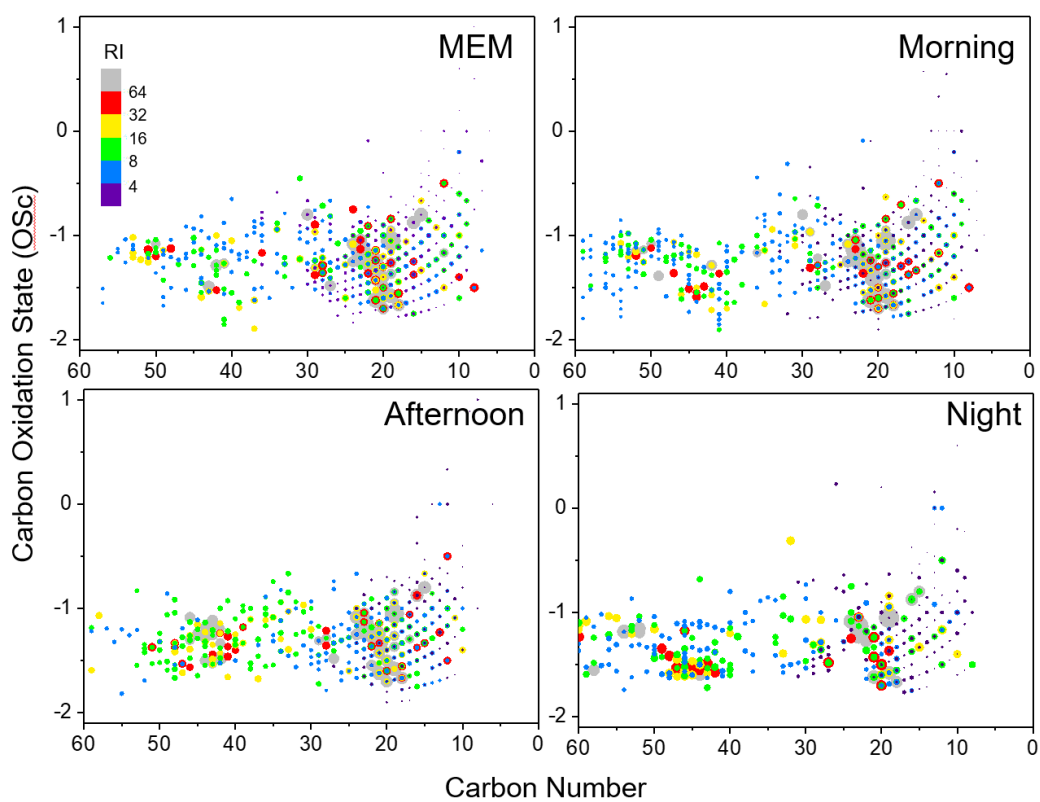



Figure S3. Backward trajectory of air mass passing through the sampling site.

68

69

70

71 Figure S4. Overlaid carbon oxidation state (OSc) symbols for CHO species. The size

72 and color bar of the markers reflect the relative peak intensities of molecular formulae

73 on a logarithmic scale. RI: relative intensity.

74

75

76

77

78

79

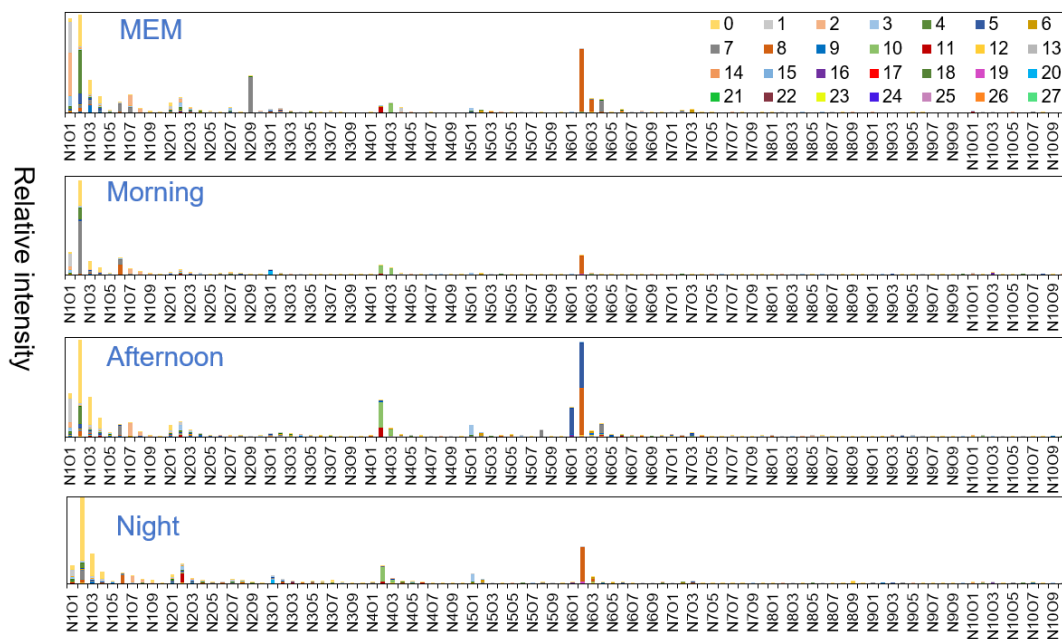



Figure S5. Relative abundance of CHON species in aerosol samples obtained by +ESI FT-ICR MS as well as the double bond equivalent (DBE) distribution. The color bar denotes the DBE value. The height of the symbols reflects the relative peak intensity of molecular formulae.

87   References

88   Koch BP, Dittmar T. From mass to structure: an aromaticity index for high-resolution

89   mass data of natural organic matter. Rapid Communications in Mass

90   Spectrometry 2006; 20: 926-932.

91   Koch BP, Dittmar T. From mass to structure: an aromaticity index for high-resolution

92   mass data of natural organic matter. Rapid Communications in Mass

93   Spectrometry 2016; 30: 250-250.

94   Sleighter RL, Hatcher PG. The application of electrospray ionization coupled to

95   ultrahigh resolution mass spectrometry for the molecular characterization of

96   natural organic matter. Journal of Mass Spectrometry 2007; 42: 559-574.

97   Yassine MM, Harir M, Dabek-Zlotorzynska E, Schmitt-Kopplin P. Structural

98   characterization of organic aerosol using Fourier transform ion cyclotron

99   resonance mass spectrometry: aromaticity equivalent approach. Rapid Commun

100   Mass Spectrom 2014; 28: 2445-54.

101