

## Electronic Supplementary Information

### Identification of ion pairs in solution by IR spectroscopy: crucial contributions of gas phase data and simulations

Sana Habka, Thibaut Very, Jeremy Donon, Vanesa Vaquero-Vara, Benjamin Tardivel, Florence Charnay-Pouget, Michel Mons, David J. Aitken, Valérie Brenner and Eric Gloaguen

#### S1. Gas phase DFT-D calculations

All RI-B97-D3/dhf-TZVPP calculations,<sup>1-4</sup> have been performed with the following non default TURBOMOLE 7.0<sup>5</sup> options when using the program define:

```
$scfconv 8
$dft
  gridsize m4
  weight derivatives
```

Convergence criteria were set by calling the option -gcart 4 of the jobex program, leading to the following thresholds:

```
threchange 1.0d-6
thrmaxdispl 1.0d-3
thrmaxgrad 1.0d-4
thrmsdispl 5.0d-4
thrmsgrad 5.0d-4
```

#### S2. AMOEBA parameters

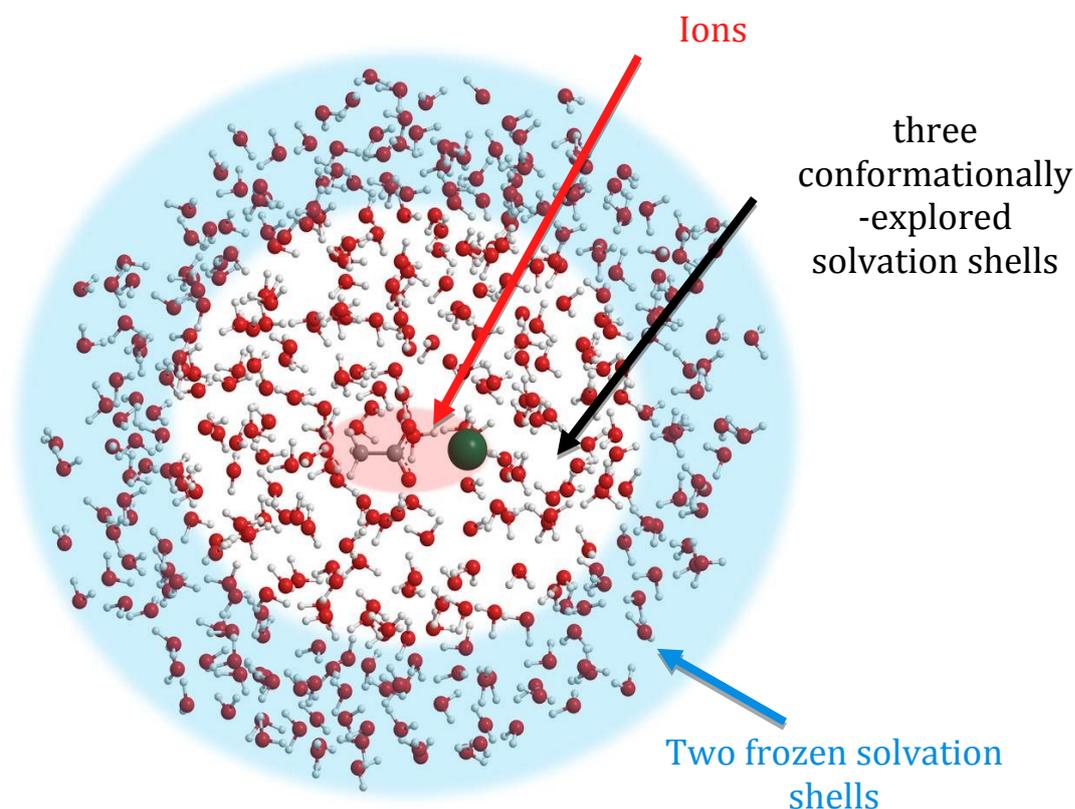
The parameters used for water and sodium are those developed in Tinker by M.L. Laury *et al* (AMOEBA14 model and associated Na model).<sup>6</sup> However parameters for the acetate anion were not available in the literature. They were built using the amoebapro13 parameters already available for the following atoms: 8 Methyl or Methylene Carbon, 9 Methyl or Methylene Hydrogen, 30 Carboxylate Carbon, 31 Carboxylate Oxygen.<sup>7</sup> Using the GDMA 2.2.1 program,<sup>8</sup> the multipolar distribution up to the quadrupole of the acetate anion has been obtained from the Distributed Multipole Analysis (DMA) of the Full-MP2/TZVPPD wavefunction expressed in terms of Gaussian atomic orbitals, calculated using the Gaussian system of programs.<sup>9</sup> This analysis has been done on the geometry of acetate optimized at the Full-MP2/TZVPPD level. Geometry optimisations of small clusters (2 to 4 molecules or ions) combining sodium acetate and water molecules, were carried out and compared to RI-B97-D3/dhf-TZVPP structures to check the performance of the set of AMOEBA parameters. All geometry optimizations at the AMOEBA force field level were carried out using the newton program included in the TINKER software package.<sup>10</sup>

atom	1	1	C	"Methyl Carbon"	6	12.011	4
atom	2	2	H	"Methyl Hydrogen"	1	1.008	1
atom	3	3	C	"C-Terminal COO-"	6	12.011	3
atom	4	4	O	"C-Terminal COO-"	8	15.999	1
atom	5	5	O	"AMOEBA water O"	8	15.995	2
atom	6	6	H	"AMOEBA water H"	1	1.008	1
atom	7	7	Na+	"Sodium Ion Na+"	11	22.990	0
vdw	1			3.8200	0.1010		
vdw	2			2.9800	0.0240	0.920	
vdw	3			3.8200	0.1060		
vdw	4			3.7000	0.1290		
vdw	5			3.5791	0.1512		
vdw	6			2.1176	0.0105	0.8028	
vdw	7			2.8500	0.1500		



### S3. Conformational search at the AMOEBA force field level

A cubic box of  $3634.2 \text{ pm}^3$  filled with 1626 water molecules and a sodium acetate SIP was first considered, setting the density at 1.016. A geometry optimization was then performed, before extracting a cluster made of the ions and all the water molecules lying entirely at less than 1400 pm of either the sodium or the acetate ions. Then, explorations were carried out after freezing all the water molecules located entirely at more than 600 pm from any ion (Figure S1). These explorations were conducted following a global method based on a biased Monte-Carlo minimisation algorithm developed by Scheraga<sup>11</sup> included in TINKER. The parameters used are the following: cartesian moves; maximum steps: 100 pm; temperature 1000 K; RMS Gradient Criterion: 0.01 a.u.; threshold used for sorting out the minima: 0.01 a.u.. Clusters containing only acetate and water molecules have been investigated following the same procedure. For both systems, acetate ions and sodium acetate pairs, more than 50 simulations were conducted, generating tens of thousands of structures. These structures were sorted according to their AMOEBA energies, and analysed in order to select structures for quantum chemistry calculation. As many structures displayed very similar arrangements of the carboxylate group and its first neighbours, thus potentially leading to similar vibrational frequencies of the carboxylate stretches, many structures of low-energy were discarded in order to avoid duplicate calculations, and to end with a set of structures roughly representative of the conformational diversity found in the exploration.



**Figure S1**

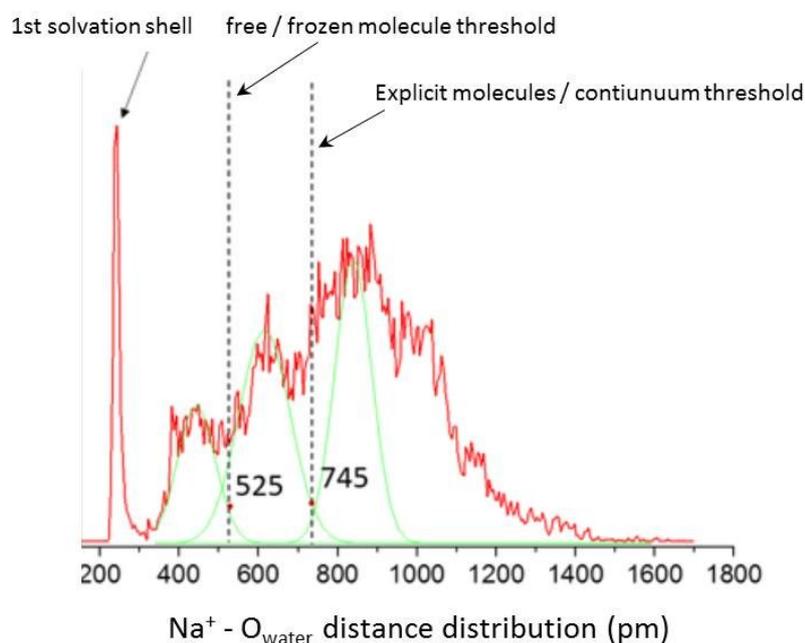
Schematic cut of a typical cluster investigated at the AMOEBA force field level

#### S4. DFT-D calculations in solution

Clusters selected after explorations were too large to be investigated at the RI-B97-D3/dhf-TZVPP level. Smaller clusters were thus defined, according to the following procedure.

Solvation was considered differently around (i) the sodium (ii) the oxygen atoms of the carboxylate and (iii) the carbon atom of the methyl group. For each of these groups, typical distances were defined by analyzing the pair distribution function resulting from the explorations (see Fig. S2). The number of solvation shells to be considered around the ionic groups was chosen after testing 3 cases: 1 optimized, 1 frozen; 2 optimized, 1 frozen; 3 optimized, 1 frozen. Since the last two cases give similar  $\text{CO}_2^-$  stretch frequencies (within  $1 \text{ cm}^{-1}$ ) the intermediate case was chosen. Furthermore, four solvation shells around the methyl group correspond to the minimum number of shells to consider so that the solvent continuum does not fill the cavity created by the hydrophobicity of the methyl group, leading to the definition of the thresholds reported in Table S1. Smaller clusters were thus obtained by removing any water molecules lying beyond all the explicit/continuum thresholds (Fig. S3). Then, water molecules lying beyond the free/frozen limit were frozen, and a geometry optimization was carried out using the solvent continuum model COSMO<sup>12</sup> implemented in TURBOMOLE. In addition to the options presented above for gas phase calculations, the following convergence thresholds were set:

```
$statpt  
threchange 1.0d-5  
thrmaxdispl 1.0d-2  
thrmaxgrad 1.0d-3  
thrrmsdispl 5.0d-3  
thrrmsgrad 5.0d-3
```

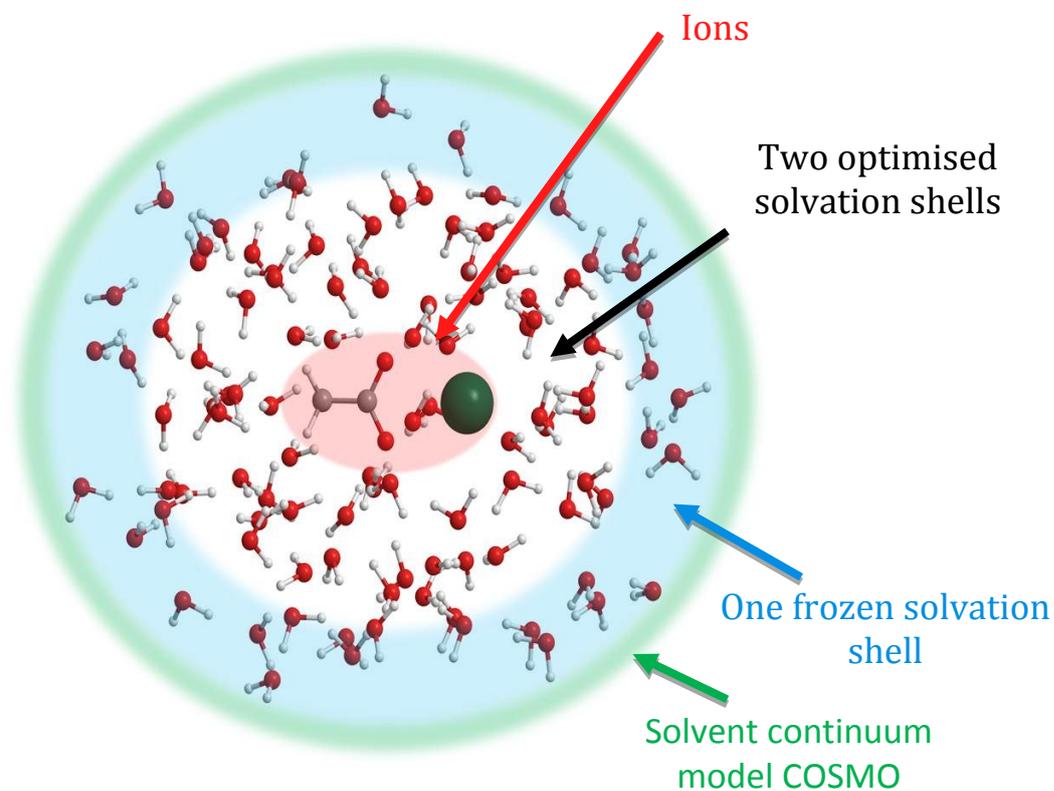


**Figure S2**

$\text{Na}^+$  -  $\text{O}_{\text{water}}$  distance distribution (red) and multiple gaussian fit (green) used to determine the thresholds needed to define clusters before optimization at the RI-B97-D3/dhf-TZVPP level.

Thresholds (pm)	free/frozen molecules	explicit/continuum
Na <sup>+</sup> - O <sub>water</sub>	525	745
O <sup>-</sup> <sub>carboxylate</sub> - O <sub>water</sub>	565	780
H <sub>methylene</sub> - O <sub>water</sub>	495	750

**Table S1**

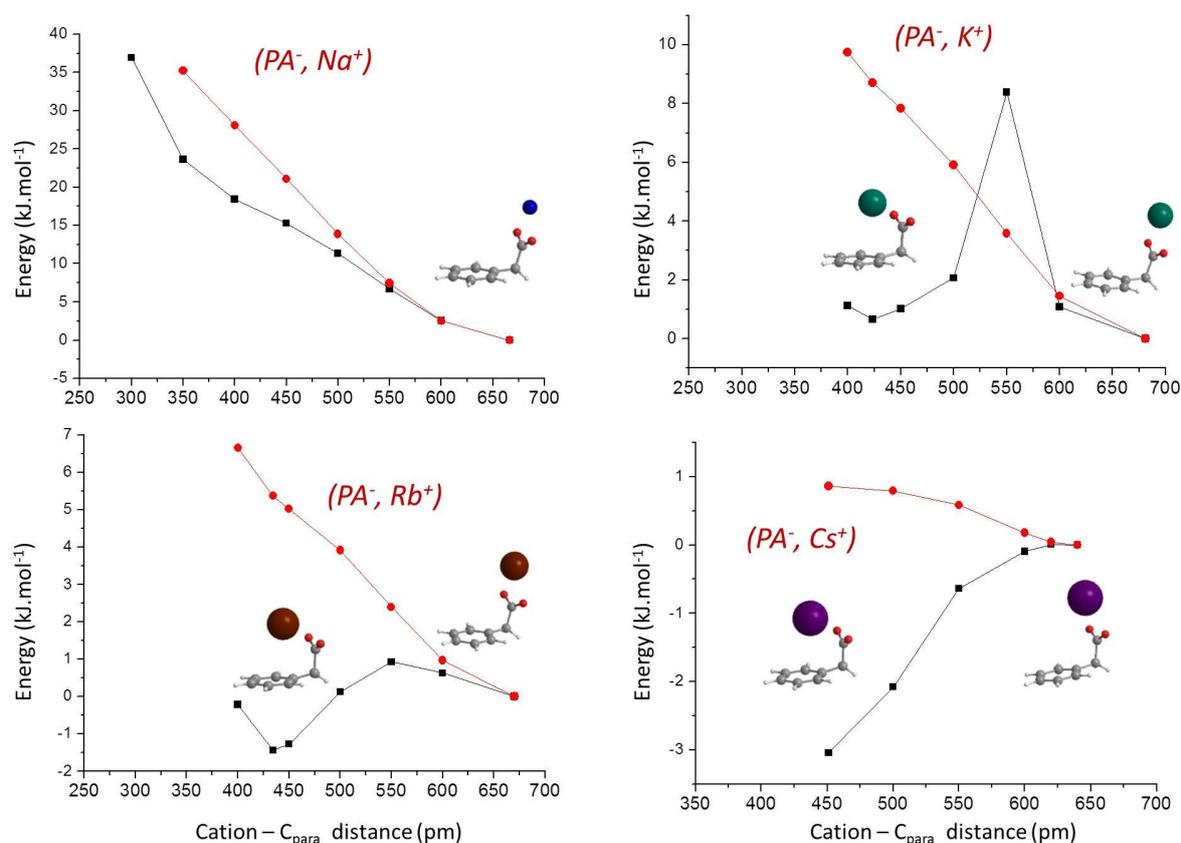


**Figure S3**

Schematic cut of a typical cluster investigated at the RI-B97-D3/dhf-TZVPP level

Finally, frequency calculations at the RI-B97-D3/dhf-TZVPP level were limited to the ions and the water molecules H-bonded to the carboxylate by freezing all the other atoms. Frequency calculations by releasing water molecules of the second shell gave similar results (within  $\sim 1 \text{ cm}^{-1}$ ). All these frequency calculations were carried out without including the solvent continuum model, and the absence of negative frequency was systematically checked.

## S5. Energy profile between O-O and O-O- $\pi$ conformers of ( $M^+$ , $PA^-$ ) systems



**Figure S4**

RI-B97-D3/dhf-TZVPP (black) and BSSE-corrected-full-CCSD(T)/dhf-TZVPP//RI-B97-D3/dhf-TZVPP (red) energies at several fixed distances between the cation and the carbon atom of the phenyl ring in *para* position, resulting in energy profiles connecting the conformation O-O (largest distances, energy arbitrarily set at 0) and O-O- $\pi$  structures where a minimum is found only at the RI-B97-D3/dhf-TZVPP level (short distances), except for Na where no O-O- $\pi$  minimum is found.

This figure shows that the RI-B97-D3/dhf-TZVPP level of theory may fail to describe properly the relative energies and the molecular structures of simple flexible systems where shaping interactions of different nature control the geometry (i.e. cation-anion and cation- $\pi$  interactions where the contribution of dispersion is different).

For the sodium acetate solutions investigated in this manuscript, such a situation is not expected to occur, but this effect may become an issue for more polarisable ions.

## S6. Results used to determine mode-dependent scaling functions

System	Harmonic RI-B97-D3/dhf-TZVPP CO <sub>2</sub> <sup>-</sup> stretch frequencies (cm <sup>-1</sup> ) sym-anti	Experimental CO <sub>2</sub> <sup>-</sup> stretch frequencies (cm <sup>-1</sup> ) sym-anti
Pro - Na <sup>+</sup>	1353.9 - 1641.3 <sup>a</sup>	1395 - n/a <sup>13</sup>
Trp - Ba <sup>2+</sup>	1464.7 - 1573.7 <sup>a</sup>	n/a - 1605 <sup>14</sup>
Pro - Ba <sup>2+</sup>	1397.8 - 1524.9 <sup>a</sup>	1419 - 1553 <sup>15</sup>
Val - Ba <sup>2+</sup>	1396.7 - 1524.1 <sup>a</sup>	n/a - 1549 <sup>15</sup>
Ser - Ba <sup>2+</sup>	1383.0 - 1540.1 <sup>a</sup>	n/a - 1572 <sup>15</sup>
His - Ba <sup>2+</sup>	1410.7 - 1473.5 <sup>a</sup>	1422 - 1501 <sup>16</sup>
[Cys - H] <sup>-</sup>	1294.7 - 1647.2 <sup>a</sup>	1317 - n/a <sup>17</sup>
[Phe - H] <sup>-</sup>	1307.5 - 1617.8 <sup>a</sup>	1328 - 1638 <sup>17</sup>
[Trp - H] <sup>-</sup>	1314.9 - 1612.2 <sup>a</sup>	1337 - 1630 <sup>17</sup>
[Tyr - H] <sup>-</sup>	1310.0 - 1616.1 <sup>a</sup>	1323 - 1635 <sup>17</sup>
Me-Pro - K <sup>+</sup>	1344.1 - 1646.0 <sup>a</sup>	1371 - 1659 <sup>18</sup>
Me <sub>2</sub> -Gly - K <sup>+</sup>	1329.3 - 1639.9 <sup>a</sup>	1365 - 1672 <sup>18</sup>
Phe - Ba <sup>2+</sup>	1404.9 - 1567.9 <sup>a</sup>	n/a - 1593 <sup>19</sup>
Pro - Li <sup>+</sup>	1383.0 - 1623.8 <sup>a</sup>	1411 - 1640 <sup>20</sup>
Pro - K <sup>+</sup>	1349.2 - 1649.0 <sup>a</sup>	1378 - 1667 <sup>20</sup>
Pro - Cs <sup>+</sup>	1348.0 - 1651.4 <sup>a</sup>	1381 - n/a <sup>20</sup>
Me-Ala - Cs <sup>+</sup>	1372.2 - 1658.7 <sup>a</sup>	n/a - 1673 <sup>20</sup>
PA·Li <sup>+</sup>	1372.6 - 1521.7 <sup>a</sup>	n/a - 1549 <sup>21</sup>
Ph-(CH <sub>2</sub> ) <sub>2</sub> -COO <sup>-</sup> Li <sup>+</sup> Conformer A	1397.5 - 1514.7 <sup>a</sup>	n/a - 1545 <sup>21</sup>
Ph-(CH <sub>2</sub> ) <sub>2</sub> -COO <sup>-</sup> Li <sup>+</sup> Conformer B	1389.4 - 1509.3 <sup>a</sup>	1449 - 1541 <sup>21</sup>
Ph-(CH <sub>2</sub> ) <sub>3</sub> -COO <sup>-</sup> Li <sup>+</sup> Conformer A	1397.0 - 1528.0 <sup>a</sup>	1453 - 1557 <sup>21</sup>
PA·Na <sup>+</sup>	1342.2 - 1538.4 <sup>a</sup>	n/a - 1563 <sup>a</sup>
PA·K <sup>+</sup>	1331.9 - 1551.5 <sup>a</sup>	n/a - 1570 <sup>a</sup>
PA·Rb <sup>+</sup>	1331.2 - 1548.9 <sup>a</sup>	n/a - 1571 <sup>a</sup>
PA·Cs <sup>+</sup>	1331.2 - 1550.3 <sup>a</sup>	n/a - 1570 <sup>a</sup>

a) this work

**Table S2**

## S7. Mode-dependent scaled harmonic frequencies and splitting compared to gas phase experimental values

	Conformer	$\nu(\text{CO}_2^-)^{\text{sym}}$ $\text{cm}^{-1}$	$\nu(\text{CO}_2^-)^{\text{anti}}$ $\text{cm}^{-1}$	$\Delta\nu(\text{CO}_2^-)$ $\text{cm}^{-1}$
<b>Li<sup>+</sup>, PA<sup>-</sup></b>	<b>exp</b>	<b>1417/1432/1453</b>	<b>1549</b>	<b>132/117/96</b>
	O-O	1405	1549	144
	O- $\pi$	1169	1716	547
<b>Na<sup>+</sup>, PA<sup>-</sup></b>	<b>exp</b>	<b>1383/1403</b>	<b>1563</b>	<b>180/160</b>
	O-O	1370	1564	194
	O- $\pi$	1171	1701	531
<b>K<sup>+</sup>, PA<sup>-</sup></b>	<b>exp</b>	<b>1377/1394</b>	<b>1570</b>	<b>193/176</b>
	O-O	1358	1577	218
	O-O- $\pi$	1339	1603	264
	O- $\pi$	1174	1694	520
<b>Rb<sup>+</sup>, PA<sup>-</sup></b>	<b>exp</b>	<b>1375/1391</b>	<b>1571</b>	<b>196/180</b>
	O-O	1358	1574	217
	O-O- $\pi$	1340	1599	259
	O- $\pi$	1176	1688	512
<b>Cs<sup>+</sup>, PA<sup>-</sup></b>	<b>exp</b>	<b>1373/1388</b>	<b>1570</b>	<b>197/182</b>
	O-O	1357	1575	218
	O-O- $\pi$	1341	1598	256
	O- $\pi$	1179	1684	504

Table S3

## S8. Structural changes from phenylacetic acid to (Cs<sup>+</sup>, PA<sup>-</sup>)

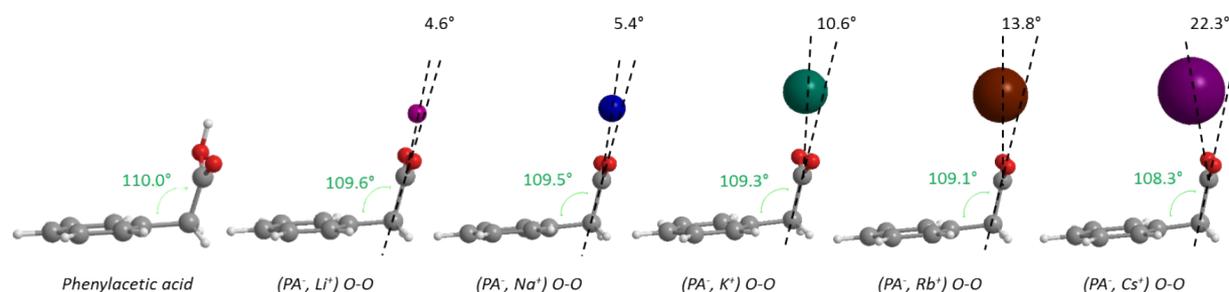


Figure S5

RI-B97-D3/dhf-TZVPP  $C_{\text{ipso}}C_{\text{methylene}}\hat{C}_{\text{carboxylate}}$  angles, together with the angle between the  $C_{\text{methylene}}C_{\text{carboxylate}}$  and  $M^+C_{\text{carboxylate}}$  axes, showing the out-of-carboxylate plane evolution along the Li-Cs series. Phenylacetic acid is also shown for comparison purposes.

**S9. Results of mode-dependent scaled harmonic frequency calculations in solution at the RI-B97-D3/dhf-TZVPP level**

	type	$\nu(\text{CO}_2^-)^{\text{sym}}$ ( $\sigma$ ) $\text{cm}^{-1}$	$\nu(\text{CO}_2^-)^{\text{anti}}$ ( $\sigma$ ) $\text{cm}^{-1}$	Relative Intensities sym/anti
AcO <sup>-</sup> (100 %)	(3 3) (47 %)	1432.2	1559.5	10/28
		1433.8	1557.8	10/23
		1427.3	1555.9	10/24
		1428.5	1555.7	10/21
		1433.0	1563.3	10/31
		1431.5	1558.9	10/26
		1434.0	1557.6	10/26
		1437.3	1562.4	10/24
		1434.4	1558.0	10/23
	(3 2) (45 %)	1420.4	1565.4	10/19
		1421.9	1559.3	10/27
		1420.4	1566.3	10/17
		1427.5	1569.1	10/38
		1399.5	1576.4	10/29
		1414.8	1560.3	10/16
		1433.5	1564.5	10/29
	(2 2) (6 %)	-	-	-
	(4 3) (1 %)			
	(4 2) (1 %)	1408.3	1593.6	10/34
(3 1) (<1 %)	1405.1	1588.4	10/27	
other (<1 %)				
all	1424.6 (11)	1565.1 (11)	10/26	
*****				
(AcO <sup>-</sup> ,Na <sup>+</sup> ) SIP (62 %)	<sup>1</sup> (3 2) <sup>0</sup> (29 %)	1431.3	1558.9	10/33
		1425.2	1566.2	10/28
		1407.5	1568.4	10/23
		1424.0	1566.1	10/42
		1404.9	1563.3	10/23
		1428.9	1563.7	10/33
		1423.5	1568.5	10/31
		1429.4	1546.0	10/28
	<sup>1</sup> (3 3) <sup>0</sup> (17 %)	1430.8	1548.2	10/23
		1427.8	1563.5	10/33
		1420.6	1551.2	10/43
		1426.8	1554.4	10/20
		1426.5	1570.6	10/32
		1416.8	1547.6	10/20
		1425.4	1549.5	10/33
		1416.5	1552.9	10/33
	<sup>1</sup> (3 3) <sup>1</sup> (4 %)	1429.4	1552.6	10/35
		1443.7	1555.5	10/32

		1431.5	1550.3	10/28
		1429.9	1548.2	10/25
		1428.6	1558.9	10/31
		1416.7	1554.3	10/27
		1432.7	1553.7	10/28
		1429.4	1560.9	10/31
	<sup>0</sup> (3 2) <sup>1</sup> (2 %)	1414.6	1575.4	10/30
		1409.5	1549.4	10/27
		1427.8	1560.7	10/26
		1415.1	1562.8	10/39
		1410.9	1561.4	10/43
	<sup>2</sup> (3 3) <sup>0</sup> (2 %)	1413.0	1571.9	10/38
		1438.7	1549.6	10/50
	<sup>0</sup> (3 2) <sup>1</sup> (1 %)	-	-	-
	<sup>2</sup> (3 3) <sup>1</sup> (1 %)	1413.6	1551.7	10/20
		1426.2	1544.6	10/23
		1421.4	1550.4	10/29
		1423.8	1544.9	10/25
		1435.7	1548.0	10/26
	<sup>2</sup> (3 2) <sup>0</sup> (1 %)	1440.6	1551.3	10/32
1432.4		1566.3	10/32	
<sup>2</sup> (3 2) <sup>1</sup> (1 %)	-	-	-	
<sup>0</sup> (3 2) <sup>2</sup> ( $<1$ %)	-	-	-	
<sup>3</sup> (3 3) <sup>0</sup> ( $<1$ %)	1441.1	1546.6	10/19	
<sup>2</sup> (4 2) <sup>0</sup> ( $<1$ %)	1411.8	1562.5	10/26	
	1407.3	1569.3	10/25	
other (<4%)	-	-	-	
all	1424.2 (10)	1557.1 (9)	10/30	
(AcO <sup>-</sup> ,Na <sup>+</sup> ) CIP (38 %)	(3 3) (14 %)	1427.2	1548.9	10/27
		1409.4	1563.6	10/24
		1435.4	1554.3	10/37
		1439.3	1552.3	10/26
		1429.4	1544.9	10/34
		1434.9	1545.4	10/35
		1431.7	1557.1	10/34
		1413.6	1556.0	10/28
		1421.6	1555.9	10/21
	(3 2) (10 %)	1430.7	1570.6	10/35
		1438.7	1570.7	10/28
		1431.8	1565.8	10/29
		1430.3	1571.2	10/29
		1421.8	1548.3	10/27
		1431.6	1566.9	10/30
	(3 2) (8 %)	1416.7	1583.1	10/29
		1408.3	1569.3	10/22

		1406.1	1594.3	10/32
		1419.4	1576.3	10/22
		1429.0	1574.1	10/26
		1416.7	1552.2	10/24
		1415.7	1573.3	10/17
	(3[1]3) (2 %)	1433.2	1550.4	10/35
		1426.3	1545.9	10/32
		1431.1	1561.6	10/35
		1405.7	1555.9	10/33
		1430.3	1554.5	10/40
	(2 2) (2 %)	-	-	-
	(3[1]2) (1 %)	1413.4	1562.4	10/23
		1412.8	1562.8	10/28
		1412.9	1582.7	10/46
	other (<2%)	-	-	-
	all	1423.5 (10)	1562.4 (12)	10/30

**Table S4**

Mode-dependent scaled harmonic frequency calculations and relative intensities of the CO<sub>2</sub><sup>-</sup> stretching modes of several structures optimised at the RI-B97-D3/dhf-TZVPP level. Structure distributions from the conformational search at the AMOEBA level are also reported.

## S10. References

1. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 19.
2. A. Schäfer, C. Huber and R. Ahlrichs, *J. Chem. Phys.*, 1994, **100**, 5829.
3. F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Lett.*, 1998, **294**, 143.
4. F. Weigend and A. Baldes, *J. Chem. Phys.*, 2010, **133**.
5. TURBOMOLE V7.0 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>
6. M. L. Laury, L. P. Wang, V. S. Pande, T. Head-Gordon and J. W. Ponder, *J. Phys. Chem. B*, 2015, **119**, 9423.
7. Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder and P. Ren, *J. Chem. Theory Comput.*, 2013, **9**, 4046.
8. A. J. Stone, *J. Chem. Theory Comput.*, 2005, **1**, 1128.
9. Gaussian 09, Revision B.01
10. TINKER V7.1-2015, Software tools for molecular design, Jay Ponder Lab, Department of Chemistry, Washington University, Saint Louis, Missouri 63130 U.S.A.; available from <http://dasher.wustl.edu/tinker/>
11. L. Piela, J. Kostrowicki and H. A. Scheraga, *J. Phys. Chem.*, 1989, **93**, 3339.
12. A. Klamt and G. Schuurmann, *Journal of the Chemical Society-Perkin Transactions 2*, 1993, DOI: 10.1039/p29930000799, 799.
13. C. Kapota, J. Lemaire, P. Maitre and G. Ohanessian, *J. Am. Chem. Soc.*, 2004, **126**, 1836.
14. R. C. Dunbar, N. C. Polfer and J. Oomens, *J. Am. Chem. Soc.*, 2007, **129**, 14562.
15. M. F. Bush, J. Oomens, R. J. Saykally and E. R. Williams, *J. Am. Chem. Soc.*, 2008, **130**, 6463.
16. R. C. Dunbar, A. C. Hopkinson, J. Oomens, C. K. Siu, K. W. M. Siu, J. D. Steill, U. H. Verkerk and J. F. Zhao, *J. Phys. Chem. B*, 2009, **113**, 10403.
17. J. Oomens, J. D. Steill and B. Redlich, *J. Am. Chem. Soc.*, 2009, **131**, 4310.
18. M. K. Drayss, D. Blunk, J. Oomens, B. Gao, T. Wyttenbach, M. T. Bowers and M. Schaefer, *J. Phys. Chem. A*, 2009, **113**, 9543.
19. R. C. Dunbar, J. D. Steill and J. Oomens, *Phys. Chem. Chem. Phys.*, 2010, **12**, 13383.
20. M. K. Drayss, P. B. Armentrout, J. Oomens and M. Schaefer, *Int. J. Mass Spectrom.*, 2010, **297**, 18.
21. S. Habka, V. Brenner, M. Mons and E. Gloaguen, *J. Phys. Chem. Lett.*, 2016, **7**, 1192.