EM∩IM: software for relating ion mobility mass spectrometry and electron microscopy data

Matteo T. Degiacomi and Justin L.P. Benesch

Electronic Supplementary Information

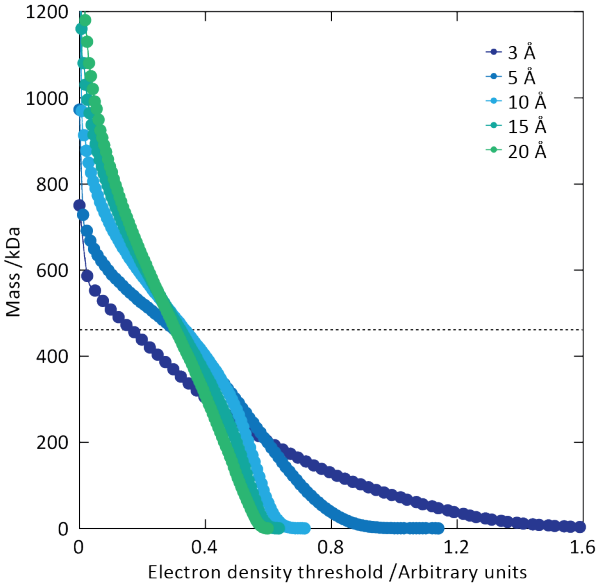

Figure S1: The trend in mass as a function of electron density does not reveal features that identify the correct mass. *We used the β-galactosidase crystal structure to simulate noise-free density maps at different resolutions, from 3 to 20 Å. Unlike CCS (Fig. 2), the trends in mass as a function of electron density threshold do not display features that reflect the known mass of the protein (dashed line).*

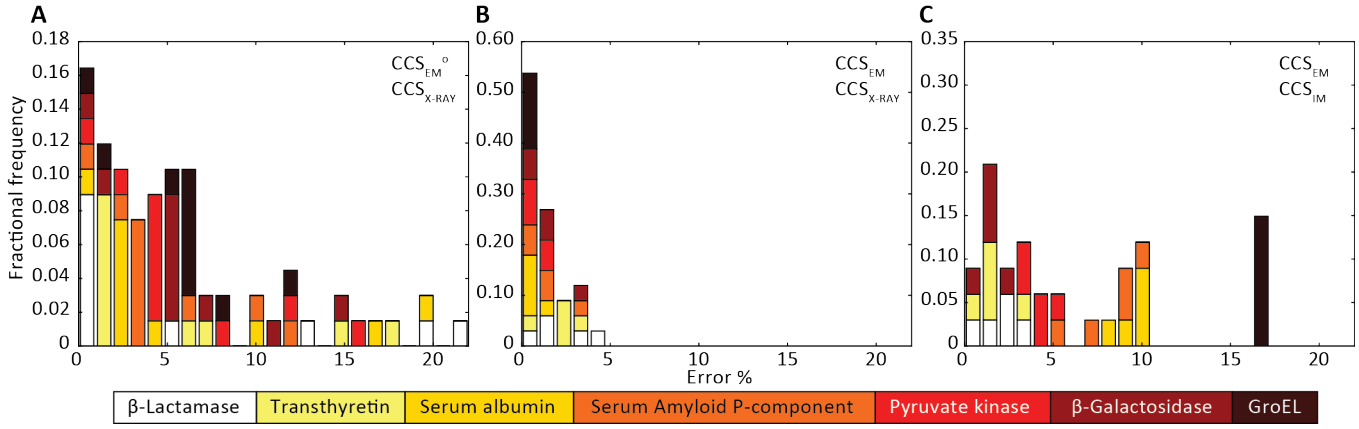

Figure S2: Histograms of error in CCS estimation before and after correction (Fig. 3). *The error distributions obtained when comparing* ***(A)*** *CCS_{EM}° with $CCS_{X\text{-}RAY}$,* ***(B)*** *CCS_{EM} (obtained after resolution-dependent correction of CCS_{EM}°) with $CCS_{X\text{-}RAY}$, and* ***(C)*** *CCS_{EM} with that measured by CCS_{IM}. The calibration reduces all CCS estimation errors significantly, with all errors <5% when compared with the crystal structure, and GroEL emerging as an expected outlier when considering IM-MS data.*

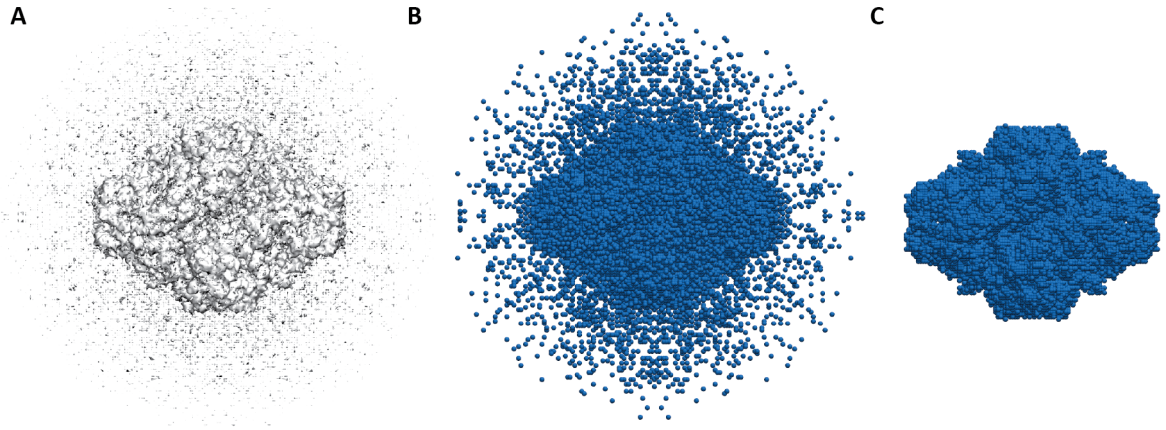

Figure S3: Illustration of de-noising algorithm in EM∩IM. ***(A)*** *Examining β-galactosidase (EMDB: 2824) at ρ*=0.07 the protein shape is clearly distinguishable, though surrounded by a large amount of noise.* ***(B)*** *The contoured map is converted into a bead model, by placing a sphere at the centres of every voxel having intensity larger than our selected ρ*.* ***(C)*** *Applying our de-noising protocol, contiguous voxels are identified. Only clusters representing more than 1% of total voxel count are retained, resulting in effective noise removal.*

Protein	PDB	Chains	$CCS_{X\text{-}RAY}$	$CCS_{IM\text{-}MS}$
β-Lactamase (1mer)	2Q2M	A	1646	1660
TTR	1F41	A,B (C,D)	3285	3410
Serum albumin	1E78	A	4470	4090
Serum Amyloid P-component (5mer)	1SAC	A-E	6438	7030
Pyruvate kinase	1F3W	A-D	9851	10220
β-Galactosidase	3IAP	A-D	15888	15520
GroEL	1SS8	A-G (H-O)	24519	20600

Table S1: Structures used to generate synthetic density maps and their CCSs (Figs. 2, 3). *$CCS_{X\text{-}RAY}$ contains the CCS calculated with IMPACT (Marklund, E.G. et al, Structure, 2015, 23, 791-9), rescaled as described previously (Benesch, J.L.P. & Ruotolo, B.T., Curr Op Struc Biol, 2011, 21, 641-9). CCS_{IM} are previously published values (Bush, M.F. et al, 2010, 82, 9557-65). All CCS values are given in $Å^2$. Chains in parentheses were symmetry-generated using BIOMATRIX operations provided in the PDB file.*

Protein	EMDB	Resolution
GroEL	5002	4.7
GroEL	1457	5.4
GroEL	1081	6.0
GroEL	5337	6.7
GroEL	1997	7.0
GroEL	1587	7.0
GroEL	5336	7.3
GroEL	2221	8.4
GroEL	1080	11.5
GroEL	5143	18.0
GroEL	5043	21.0
β-Galactosidase	5995	3.2
β-Galactosidase	2824	4.2

Table S2: Electron density maps used as experimental test cases (Fig. 4). *Resolution as reported, in Å.*