Protein Structures

DOI: 10.1002/anie.200((will be filled in by the editorial staff))

Intrinsically Disordered p53 and its Complexes Populate Compact Conformations in the Gas Phase**

*Kevin Pagel, Eviatar Natan, Zoe Hall, Alan R. Fersht, Carol V. Robinson**

The use of ion mobility mass spectrometry (IM-MS) to study the topology of proteins and their complexes is attracting considerable attention.[1] Under the appropriate conditions results have shown that unique insight can be gained for intractable systems such as protein aggregates [2] viral capsids [3] and membrane complexes.[4] From fundamental studies of peptides and proteins it is clear that folded structure, present in bulk solution, upon transfer into the gas phase, can retain native-like compact states but may unfold when activated or stored.[1, 5] Similarly assemblies of globular proteins studied by IM-MS have been found to retain defined topological arrangements,[6] with unfolding of subunits being linked to activation in the gas phase. High charge states of proteins and their complexes in the gas phase are associated with extended unfolded structures, analogous to unfolding processes induced by protonation in solution.[1, 7]

While the link between charge state in solution and unfolding in the gas phase appears to hold for individual proteins and globular complexes it is not clear how the relationship between charge and structure would hold for complexes containing intrinsically disordered domains. Given that such domains are prevalent in nature (~30% of the eukaryotic genome[8]) and have a high proportion of charged residues and a low percentage of hydrophobic ones,[9] these regions could remain elongated in the gas phase, analogous to the extended conformations of highly charged individual proteins. Alternatively in the absence of solvent these unstructured regions could collapse onto structured motifs. To examine these possible scenarios we applied IM-MS to one of the most studied natively disordered tetrameric complexes, that of the tumour suppressor protein p53. This protein complex, discovered more than 30 years ago is a major chemotherapeutic target,[10] and has continued to thwart structural biologists largely because disordered regions comprise ~40% of the protein.[11] For our study we systematically introduced intrinsically disordered regions (flexible linker, N and C terminii) to the two folded domains (Fig. 1). This enabled us to probe the effects of disordered regions on the various constructs as well as full-length tetrameric p53 bound to cognate DNA.

The basic concept of IM-MS starts with injection of an ion packet into a cell filled with inert neutral gas. Aided by a weak electric field, the ions travel through the cell, colliding with neutral gas molecules in their path. Ions with a compact globular shape undergo fewer collisions and thus possess higher mobility than ions with more extended conformations, enabling their separation. Drift times measured in IM are usually converted into collision cross sections (CCS) - values that are independent of specific instrument parameters.[12] Theoretical CCS of candidate structures can be calculated either via a projection approximation (PA)[13] or an exact hard sphere scattering (EHSS)[14] algorithm. When high-resolution X-ray or NMR structures are available, a comparison between experimental and theoretical CCS can reveal different conformational states.[15] For many large protein assemblies, structural information is often only available for structured motifs, subunits or sub-complexes as is the case for p53.

In the absence of DNA, p53 is thought to exist as an ensemble of homo-tetramers with different coexisting conformations. Weakly associated dimers are formed by the core domain (C), which binds to DNA, and the tetramerization domain (T), which controls oligomerization. Upon DNA binding, p53 wraps around the double helix and a compact, rigid structure has been proposed.[11b, 16] Starting with the simplest 4.1 kDa tetramerization domain (T) we applied IM-MS to the intact $(T)_4$ tetramers, the most abundant species observed from an ammonium acetate (AA) buffered solution (Fig. 1a and Fig. S1). The average CCS obtained was in close agreement with theoretical values (Table 1). We next considered the isolated core domain (C), which was found to exist exclusively as 26.2 kDa monomers (Fig. 1b), in accord with previous studies that show that C does not form higher-order oligomers in the absence of DNA.[17] Comparison of the experimental CCS with the theoretical value (Table 1) indicates a slightly collapsed native-like conformation of the p53 core, in line with previous IM data for this construct.[18] MS of the core-DNA complex $(C)_4$DNA shows three species populated simultaneously: C, $(C)_2$DNA and $(C)_4$DNA (Fig. 1c). Monomeric C does not bind to DNA, and $(C)_2$DNA and C are likely present in solution and not formed via CID due to their charge state signatures.[19] These observations are consistent with each dimer binding separately to a DNA repeat element in the dimer of dimers which constitute the p53 core.[20] Close agreement was found between theoretical and experimental values for this $(C)_4$DNA complex (Table 1).

Having obtained close correlation between theoretical and experimental CCS of $(T)_4$, C and $(C)_4$DNA we considered the effects of the 14 amino acid disordered linker between C and T domains in the core-linker-tet tetramer $(CT)_4$ construct. The 120kDa

[*] Dr. K. Pagel, [‡+] Z. Hall, Prof. Dr. C.V. Robinson
Physical & Theoretical Chemistry Laboratory,
Department of Chemistry, University of Oxford,
South Parks Road, OX1 3QZ,
Oxford, UK
E-mail: carol.robinson@chem.ox.ac.uk

Dr. E. Natan,[+] Prof. Dr. A.R. Fersht
MRC Laboratory of Molecular Biology,
Hills Road, CB2 0QH
Cambridge, UK

[‡] Current address: Department of Molecular Physics
Fritz Haber Institute of the Max Planck Society
Faradayweg 4-6
14195 Berlin Germany

[+] Equal contribution

[**] We thank J.L.P. Benesch and G. von Helden for critical reading, the German Academy of Sciences Leopoldina (KP), The Royal Society (ARF, CVR) for support and an MRC Programme Grant G0901534 to ARF.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

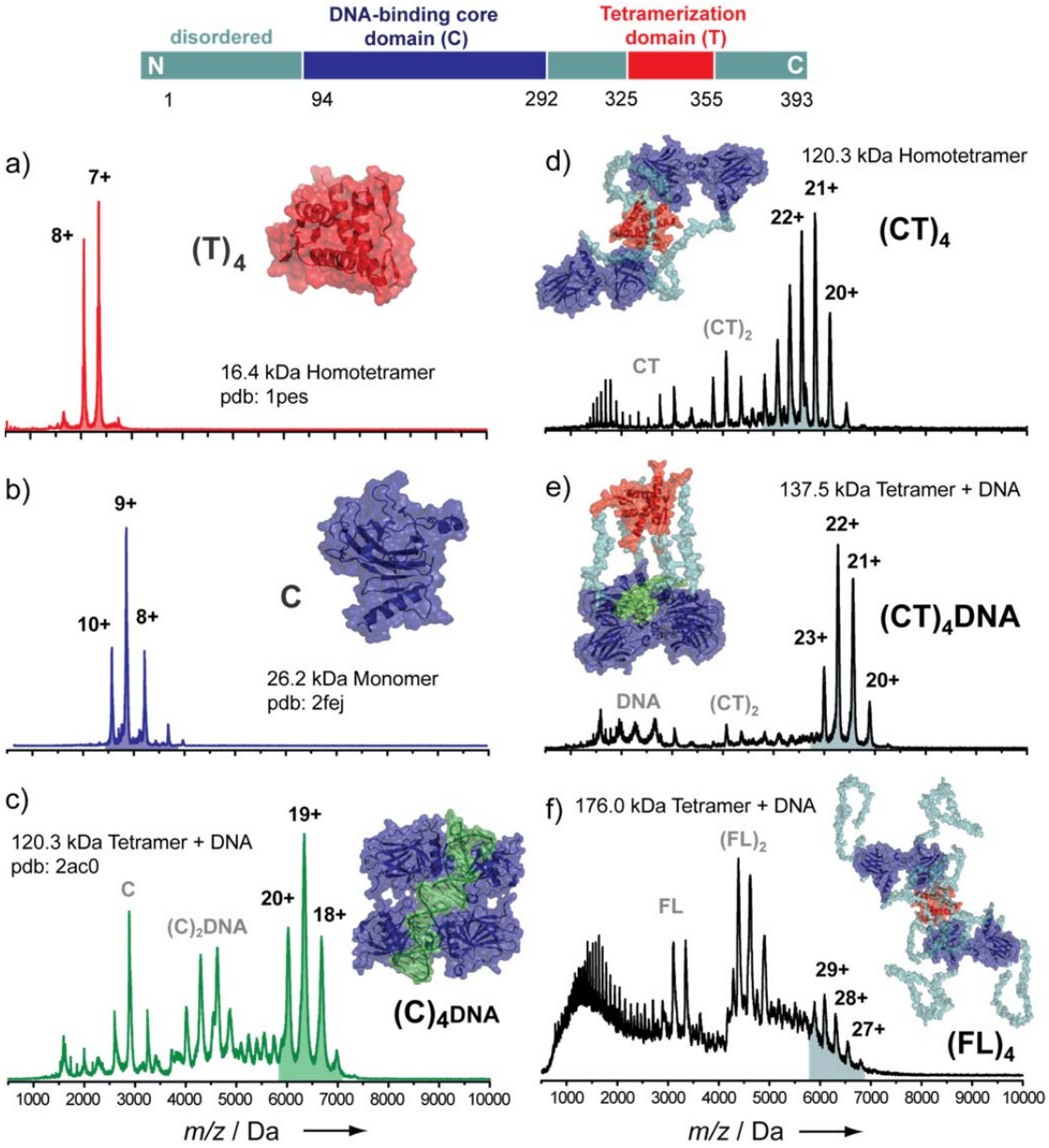


Figure 1. Schematic representation of p53 and mass spectra of individual domains and constructs. All spectra were recorded from ammonium acetate buffered solution and at similar instrument conditions. Structural models for $(T)_4$, C and $(C)_4$DNA correspond to X-ray or NMR structures with the pdb codes (a-c). Model structures for $(CT)_4$, $(CT)_4$DNA and FL_4 were obtained from a previous study.[11b]

tetrameric, non-DNA containing form of the complex is accompanied by a small fraction of dimeric and monomeric CT (Fig. 1d).[8b, 14b] Surprisingly the experimental CCS of $(CT)_4$ corresponds to ~60% of the theoretical values (Table 1). Interestingly there is little difference in the experimental CCS between $(CT)_4$ and $(C)_4$DNA. This indicates a significant structural collapse of $(CT)_4$ in the gas phase, giving rise to a value close to the compact $(C)_4$DNA structure. Adding DNA to $(CT)_4$ to form $(CT)_4$DNA results in an increase in stability, as evidenced by the reduction in the population $(CT)_2$ (Fig. 1e). The average CCS is much lower than theoretical values calculated for the proposed model structure of $(CT)_4$DNA with extended disordered regions (Table 1). Comparing the experimental CCS of $(CT)_4$ and $(CT)_4$DNA we find them to be closely similar. Therefore it is likely that the structural collapse observed in both complexes is not due to the subdomains T_4 and C_4, but rather to the connecting disordered linker.

We next considered the effects on the CCS of adding further intrinsically disordered regions. Full-length tetrameric p53 complex contains extensive, intrinsically disordered domains at the N and C termini and, similar to $(CT)_4$, forms tetramers also in the absence of DNA. In mass spectra three major species are populated: FL monomer with a bimodal charge state distribution, $(FL)_2$ and $(FL)_4$ (Fig. 1f). The coexistence of $(FL)_2$ and $(FL)_4$ is in accord with p53 being a weakly associated dimer of dimers. The presence of FL monomers with high charge states is indicative of unfolding implying either gas phase activation and/or solution dissociation, in agreement with the low thermal stability of full-length p53.[17] The experimental CCS of $(FL)_4$ ~50% lower than the theoretical value.

To investigate the origin of this dramatic difference between experimental and theoretical CCS we performed molecular dynamics (MD) simulations *in vacuo* using models of $(CT)_4$ and the $(CT)_4$DNA as starting points.[11b] Simulations, three replicates (1ns) showed that both complexes equilibrate towards compact conformations after ~0.1 ns. Excellent agreement was found between experimental CCS and the average values calculated for the most populated conformational states of $(CT)_4$ and $(CT)_4$DNA during MD (Table 1, Fig. 2). The backbone RMSD from corresponding starting structures was calculated over the simulation period. The RMSD of the individual core and tetramerization domains within the $(CT)_4$ and $(CT)_4$DNA complexes was found to

increase by ~3-4 Å. By contrast the full complexes exhibited a significant RMSD from their initial conformations (16-20 Å) (Fig. S3). This value is considerably larger than the 2-3 Å RMSD observed *in silico* for other proteins and complexes in solvent free environments.[21] This implies that rearrangement of p53 constructs arises from convergence of the essentially unaltered core and tetramerization domains, *via* collapse of the flexible linkers.

Table 1. Experimental and theoretical CCS for p53 constructs and their DNA complexes. (*) $^{13}C/^{15}N$ labelled variant, (**) MW_{DNA} 15946 Da, for details see SI.

Complex	Res.	MW [Da]	pdb	z	CCS exp. [Å^2]	CCS calc. [Å^2] PA	EHSS
$(T)_4$	325 - 355	16400*	1pes	8 7	1453 1397	1457	1862
C	94 - 312	26150*	2fej	10 9 8	1746 1851 1907	1944	2463
$(C)_4$DNA	94 - 312 + DNA**	120100*	2ac0	20 19 18	5327 5392 5434	5397	7098
$(CT)_4$	94 - 360	120300	ref 11b	22 21 20 19	5449 5536 5554 5521	8184	10319
$(CT)_4$DNA	94 - 360 + DNA**	137500	ref 11b	23 22 21 20	6012 5978 6092 6105	8022	10352
$(FL)_4$	1 - 393	176000	ref 11b	29 28 27 26	7160 7145 7139 7107	15942	19288

In principle, all proteins or protein complexes might undergo a compaction when transferred into a solvent-free environment. This collapse, is characterized by minor rearrangements of the charged residues at the protein exterior, which fold back to the protein surface to form strong electrostatic interactions upon removal of water.[22] For large protein complexes such a compaction would produce only a subtle effect. More pronounced reorganizations have been reported in which complexes undergo spontaneous compaction when transferred into the gas phase [3, 23, 25] or collapse in response to collisional activation.[15, 21a, 23, 24, 26] In both scenarios, the extent of collapse is related to the topology of the complex and the major difference between the two was shown to be the charge state dependence.[21a] For the p53 constructs used here, no correlation between charge state and collapse was observed (see also SI) implying that compaction occurs spontaneously.

The average gas phase density of native-like, globular proteins is ~ 0.61 g cm^{-3} enabling the CCS of an idealized spherical protein to be calculated.[12] Comparing CCSs for p53 constructs as well as the constant-density estimate (Fig. 2) reveals a clear correlation between measured values and CCSs estimated for spherical particles. Thus despite extensive collapse, p53 constructs exhibit typical packing densities of native-like globular proteins. Collapsed native-like structure is further indicated by *in vacuo* MD simulations of $(CT)_4$ and $(CT)_4$DNA (Fig. 2). CCSs calculated for the compact MD-derived structures agree with values obtained experimentally. Interestingly backbone RMSD analysis, during collapse, is consistent with the majority of residues in the folded C and T domains remaining unaffected. By contrast considerable rearrangements are observed for residues in the intrinsically disordered linker (Fig. S3).

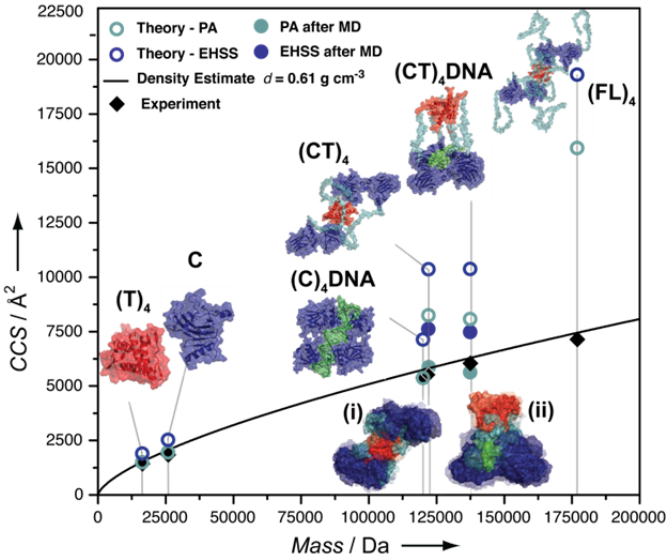


Figure 2. Plot of mass against CCS for p53 domains and constructs together with representative structures from MD simulations. Experimental values (black diamonds) are compared with theoretical CCS obtained *via* the projection approximation (PA, light blue open circles) and exact hard sphere scattering (dark blue open circles, EHSS) algorithms (Table 1) using the MOBCAL software.[13b] CCSs are calculated. using atomic information from pdb files (Table 1) or proposed models [11b] Analogous values, are calculated after MD simulation (filled circles). The estimated mass to CCS-correlation of idealized spherical particles, with the average density of a native-like gas-phase protein, is shown (line).[12b] Three of the most populated conformers formed during MD simulations of $(CT)_4$ and $(CT)_4$DNA are overlaid (i) and (ii) respectively.

Given this dramatic collapse it is interesting to consider at what stage this occurs. If we examine the average charge state of all complexes we find that $(T)_4$, C and $(C)_4$DNA show good agreement to predictions based on the correlation with surface area[27] but $(CT)_4$ and $(CT)_4$DNA reveal significantly lower charge states than predicted (Table S2). This is in contrast to isolated, monomeric intrinsically disordered proteins, for which the charge states were shown to correlate well with surface area estimates.[28] It is noteworthy that the surface areas calculated for the collapsed structures, formed during MD simulations, recapitulate the experimental average charge state remarkably closely (Table S2) providing further evidence for collapse. Interestingly loss of surface area of the native-like p53 constructs implies that collapse occurs concomitantly with charging. This is in accord with recent proposals for the electrospray mechanism in which complexes reside in the interior of droplet and solvent and electrolyte ions expel the excess surface charge of droplets. Charging of the protein takes place at the very last stages of desolvation, via transfer of charges onto the gas phase complex.[29] Consequently, fewer charges than expected from the surface area estimate are obtained for the p53 constructs with large intrinsically disordered regions, which must therefore collapse very late in the electrospray process.

The magnitude of the collapse observed here is unprecedented among protein complexes considered to date. This may be attributed in part to overestimation of the intrinsically disordered domains in the CCS calculations. However such an effect is unlikely to account for the differences in CCS and average charge of the magnitude observed here. We conclude that the vast majority of the gas phase rearrangement of p53 occurs at the flexible linker and the disordered termini, which exhibit similar properties to unfolded protein chains that collapse in the gas phase. Thus the CT and FL p53 complexes do not retain the same conformational state in the gas phase as observed in solution. The inner core, by contrast, which governs the

intermolecular organization of the complex, appears to retain its structural integrity.

Experimental Section

Protein purification and expression is described in SI. All proteins and complexes were buffer-exchanged into 500 mM AA, pH 6.9 except FL (250 mM AA, pH 6.9) using Micro Bio-Spin 6 columns (Bio-Rad, UK) or Slide-A-Lyzer dialysis cassettes (Fisher Scientific, UK). As charge reduction can stabilize complexes without noticeably affecting their folding behaviour in solution, AA buffer was in some cases replaced with triethylammonium acetate (TEAA) buffer.[30] In these conditions lower charged ions were formed (Fig. S2). The CCSs remained similar (Table S1). DNA, added after buffer exchange, is the p53 response element.[11b] Absolute CCS measurements were performed on a modified Synapt HDMS[31] described in detail previously.[12b]

Received: ((will be filled in by the editorial staff))
Published online on ((will be filled in by the editorial staff))

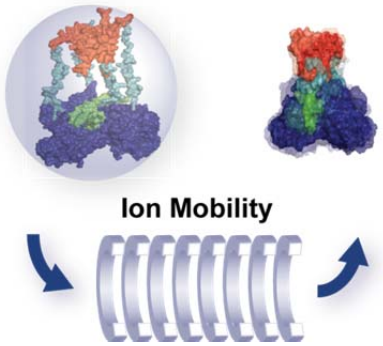
Keywords: Mass spectrometry · Ion mobility · Protein folding · Intrinsically Disordered Proteins · Analytical methods

[1] B. C. Bohrer, S. I. Merenbloom, S. L. Koeniger, A. E. Hilderbrand, D. E. Clemmer, *Annu. Rev. Anal. Chem.* **2008**, *1*, 293-327.
[2] S. L. Bernstein, N. F. Dupuis, N. D. Lazo, T. Wyttenbach, M. M. Condron, G. Bitan, D. B. Teplow, J. E. Shea, B. T. Ruotolo, C. V. Robinson, M. T. Bowers, *Nat. Chem.* **2009**, *1*, 326-331.
[3] C. Uetrecht, I. M. Barbu, G. K. Shoemaker, E. van Duijn, A. J. R. Heck, *Nat. Chem.* **2011**, *3*, 126-132.
[4] M. Zhou, N. Morgner, N. P. Barrera, A. Politis, S. C. Isaacson, D. Matak-Vinkovic, T. Murata, R. A. Bernal, D. Stock, C. V. Robinson, *Science* **2011**, *334*, 380-385.
[5] L. W. Zilch, D. T. Kaleta, M. Kohtani, R. Krishnan, M. F. Jarrold, *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1239-1248.
[6] B. T. Ruotolo, K. Giles, I. Campuzano, A. M. Sandercock, R. H. Bateman, C. V. Robinson, *Science* **2005**, *310*, 1658-1661.
[7] H. Shi, N. A. Pierson, S. J. Valentine, D. E. Clemmer, *J. Phy.s Chem. B* **2012**, *116*, 3344-3352.
[8] R. Pancsa, P. Tompa, *PLoS One* **2012**, *7*, e34687.
[9] V. N. Uversky, *Int. J. Biochem. Cell Biol.* **2011**, *43*, 1090-1103.
[10] A. J. Levine, M. Oren, *Nat. Rev. Cancer* **2009**, *9*, 749-758.
[11] a) E. N. Nikolova, E. Kim, A. A. Wise, P. J. O'Brien, I. Andricioaei, H. M. Al-Hashimi, *Nature* **2011**, *470*, 498-502; b) H. Tidow, R. Melero, E. Mylonas, S. M. V. Freund, J. G. Grossmann, J. M. Carazo, D. I. Svergun, M. Valle, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12324-12329.
[12] a) B. T. Ruotolo, J. L. P. Benesch, A. M. Sandercock, S.-J. Hyung, C. V. Robinson, *Nat. Protocols* **2008**, *3*, 1139-1152; b) M. F. Bush, Z. Hall, K. Giles, J. Hoyes, C. V. Robinson, B. T. Ruotolo, *Anal. Chem.* **2010**, *82*, 9557–9565.
[13] a) G. von Helden, M. T. Hsu, N. Gotts, M. T. Bowers, *J. Phys. Chem.* **1993**, *97*, 8182-8192; b) M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz, M. F. Jarrold, *J. Phys. Chem.* **1996**, *100*, 16082-16086.
[14] A. A. Shvartsburg, M. F. Jarrold, *Chem. Phys. Lett.* **1996**, *261*, 86-91.
[15] B. T. Ruotolo, S.-J. Hyung, P. M. Robinson, K. Giles, R. H. Bateman, C. V. Robinson, *Angew. Chem. Intl. Ed.* **2007**, *119*, 8147-8150.
[16] E. Natan, D. Hirschberg, N. Morgner, C. V. Robinson, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 14327-14332.
[17] A. C. Joerger, A. R. Fersht, *Annu. Rev. Biochem.* **2008**, *77*, 557-582.
[18] P. A. Faull, H. V. Florance, C. Q. Schmidt, N. Tomczyk, P. N. Barlow, T. R. Hupp, P. V. Nikolova, P. E. Barran, *Int. J. Mass Spectrom.* **2010**, *298*, 99-110.
[19] J. L. P. Benesch, *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 341-348.
[20] E. Arbely, E. Natan, T. Brandt, M. D. Allen, D. B. Veprintsev, C. V. Robinson, J. W. Chin, A. C. Joerger, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8251-8256.
[21] a) Z. Hall, A. Politis, M. F. Bush, L. J. Smith, C. V. Robinson, *J. Am. Chem. Soc.* **2012**; b) R. Friemann, D. S. D. Larsson, Y. Wang, D. van der Spoel, *J. Am. Chem. Soc.* **2009**, *131*, 16606-16607.
[22] K. Breuker, F. W. McLafferty, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 18145-18152.
[23] C. J. Hogan, B. T. Ruotolo, C. V. Robinson, J. Fernandez de la Mora, *J. Phys. Chem. B* **2011**, *115*, 3614-3621.
[24] T. W. Knapman, V. L. Morton, N. J. Stonehouse, P. G. Stockley, A. E. Ashcroft, *Rapid Comm. Mass Spectrom.* **2010**, *24*, 3033-3042.
[25] C. Kaddis, S. Lomeli, S. Yin, B. Berhane, M. Apostol, V. Kickhoefer, L. Rome, J. Loo, *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1206-1216.
[26] J. Freeke, C. V. Robinson, B. T. Ruotolo, *Int. J. Mass Spectrom.* **2010**, *298*, 91-98.
[27] Z. Hall, C. V. Robinson, *J. Am. Soc. Mass. Spectrom.* **2012**, *in press.*
[28] L. Testa, S. Brocca, R. Grandori, *Anal. Chem.* **2011**, *83*, 6459-6463.
[29] C. J. Hogan, J. A. Carroll, H. W. Rohrs, P. Biswas, M. L. Gross, *Anal. Chem.* **2008**, *81*, 369-377.
[30] a) D. Lemaire, G. Marie, L. Serani, O. Laprevote, *Anal. Chem.* **2001**, *73*, 1699-1706; b) K. Pagel, S.-J. Hyung, B. T. Ruotolo, C. V. Robinson, *Anal. Chem.* **2010**, *82*, 5363-5372.
[31] S. D. Pringle, K. Giles, J. L. Wildgoose, J. P. Williams, S. E. Slade, K. Thalassinos, R. H. Bateman, M. T. Bowers, J. H. Scrivens, *Int. J. Mass Spectrom.* **2007**, *261*, 1-12.

Kevin Pagel, Eviatar Natan, Zoe Hall, Alan R. Fersht and Carol V. Robinson*

_____ **Page – Page**

Intrinsically Disordered p53 and its Complexes Populate Compact Conformations in the Gas Phase



Spontaneous Shrinking: The intrinsically disordered tumor suppressor protein complex p53 was analyzed *in vacuo* using a combination of ion mobility mass spectrometry and molecular dynamics simulations. Structured p53 subdomains retain their overall topology upon transfer into the gas phase. When intrinsically disordered segments are introduced into the protein sequence however the complex spontaneously collapses in the gas phase to a compact conformation.