USE OF CALIX[4]ARENES TO RECOVER THE SELF-ASSEMBLY ABILITY OF MUTATED p53 TETRAMERIZATION DOMAINS

Susana Gordo Villoslada

2008

Memòria presentada per

Susana Gordo Villoslada

per optar al grau de doctor per la Universitat de Barcelona

Revisada per:

Prof. Ernest Giralt i Lledó

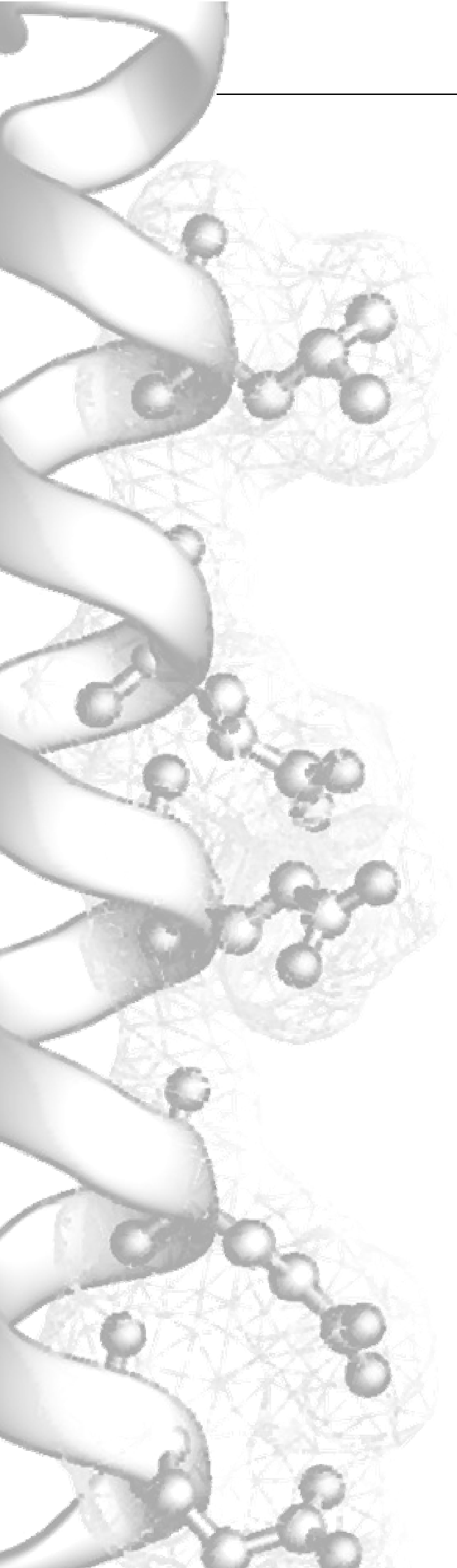
Universitat de Barcelona

Director

Programa de Química Orgànica

Bienni 2003-2005

Barcelona, abril de 2008



RESULTS

RESULTS 1

The p53 tetramerization domain and its mutants

The tetramerization domain of p53 and its mutants with compromised oligomerization properties represent an excellent case for designing ligands able to recognize and stabilize the native state of oligomeric proteins.

This chapter introduces p53TD and the mutant proteins selected for the present dissertation. Some of their biophysical properties are also analyzed.

1.1. The tetramerization domain of p53 and its mutants: who they are

The structure of the isolated tetramerization domain of p53 (p53TD) was solved long ago by both NMR spectroscopy[1,2] and X-ray crystallography.[3] The domain comprises no more than 30 residues (**Figure 1.1**), structured in a β-strand (E326-R333) and an α-helix (R335-A355 or K357, depending on the source) linked by a sharp turn (G334). The highly symmetric tetramer is better described as a dimer of dimers, in which each primary dimer is formed by an antiparallel β-sheet and two antiparallel α-helices. The dimers are arranged orthogonally, burying a hydrophobic core formed by the α-helices bundle, and leaving the two β-sheets on opposite faces of the surface (**Figure 1.1**).

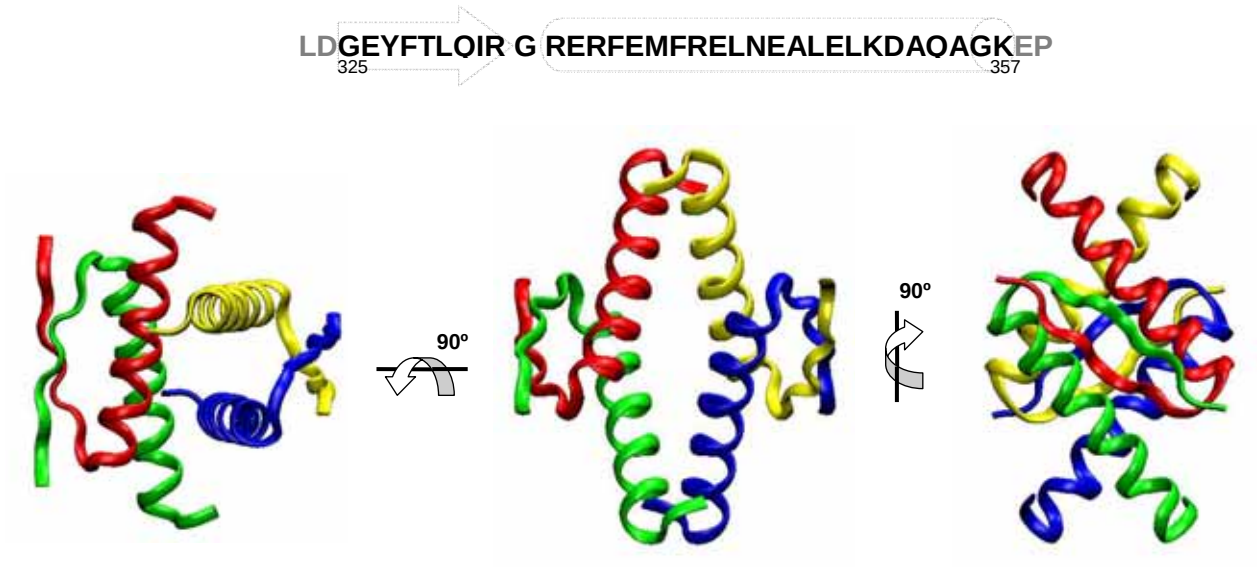


Figure 1.1. The p53 tetramerization domain sequence and ribbon models of the backbone structure (1AIE at PDB)

The hydrophobic effect is the major stabilizing force of p53TD structure.[4-6] First, the monomer adopts a sharp V-shape that forms a small hydrophobic core comprising the side-chains of residues I334, G334, F338 and F341 (**Figure 1.2A**). Then, the interaction of two monomers in the primary dimer leads to a hydrophobic interface encompassing F328, L330, I332, F338, F341 and N345 (**Figure 1.2B**). At this level, the hydrogen-bonds established in the antiparallel β-sheet, and an intermonomer salt bridge between R337-D352 side-chains also are important for the packing. Finally, the interaction of two primary dimers results in an interface between α-helices that buries the side-chains of residues M340, L344, A347, L348 and L350 from the solvent, thus generating a highly hydrophobic core (**Figure 1.2C**). The resulting tetramer is a highly stable structure and, indeed, monomeric or dimeric units are not thermodynamically stable for the wild-type domain.[4,7]

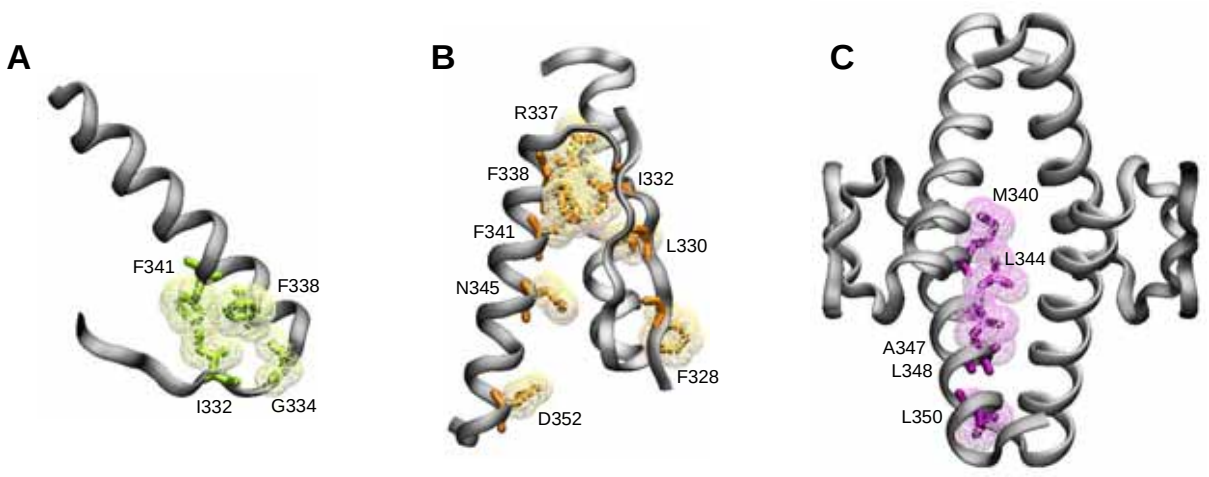


Figure 1.2. Residues involved in the stabilization of **(A)** the monomer, **(B)** the primary dimer and **(C)** the tetramer. For simplicity, only the residues of one of the monomers are represented.

Since it was discovered, p53TD has been the matter of thorough biophysical studies,[5-15] to the point that the role of nearly every single amino acid in the structure is understood. Residues within the hydrophobic core are know to be crucial for the tetramer assembly,[5,6,16] as reflected in their conservation throughout evolution.[17,18] In contrast, solvent exposed residues tune the overall stability of the domain.[11,19,20]

From a functional point of view, the tetramerization domain has attracted little attention, mainly because p53 research tends to focus on the DNA binding core. p53TD not only assembles the tetrameric active structure of the protein;[21-26] directly or indirectly, it is also responsible for the subcellular localization,[27-31] degradation[32,33] and numerous protein-protein interactions of p53.[34-45]

Furthermore, several cases of human cancer have been associated to mutations within the sequence of the p53TD.[46] A single mutation that compromises the ability of p53 to tetramerize may result in the loss of protein function,[47] which in turn causes the organism to be prone to neoplastic transformations. The greater the destabilization of the tetramerization domain, the higher the risk of tumor development.[19,20,47-52] However, only a few cases of genetic alterations in p53TD have been reported. It has been suggested that the frequency of mutations within this domain may be underestimated because only the DNA binding region is usually considered for sequencing.[24] In fact, the structure of the tetramerization domain is unusually sensitive to mutations since, for each mutation in a monomer, a total of four residues are changed in the tetramer.[5]

In the present thesis, the wild-type protein (**p53wt**) and three natural single-point mutations from the p53 tetramerization domain have been studied. Namely, they are: **R337H**, **G334V** and **L344P**. These three natural mutants have been imputed in cases of cancer, and their defective tetramerization abilities are described in the literature.

1.1.1. Mutant R337H

Mutation of arginine 337 to histidine (R337H) is the most important mutation reported to date for the tetramerization domain. Indeed, R337H is the most frequent *germline mutation*[a] for the *p53* gene (**Figure 1.3**), although whether it can be classified as a cause of Li-Fraumen syndrome[b] remains a subject of controversy.[54-56]

This inherited mutation is mainly associated with the atypically frequent cases of pediatric adrenocortical carcinoma (ACC) in southern Brazil.[56-62] Since first being reported, mutant R337H has been extensively studied for understanding the unusual tissue-selective deactivation mechanism of p53 function. DiGiammarino and co-workers[50] examined in detail the biophysical properties of the mutated tetramerization domain and found that the pH-dependent instability of R337H might be the origin of the tissue-specificity in tumor development.

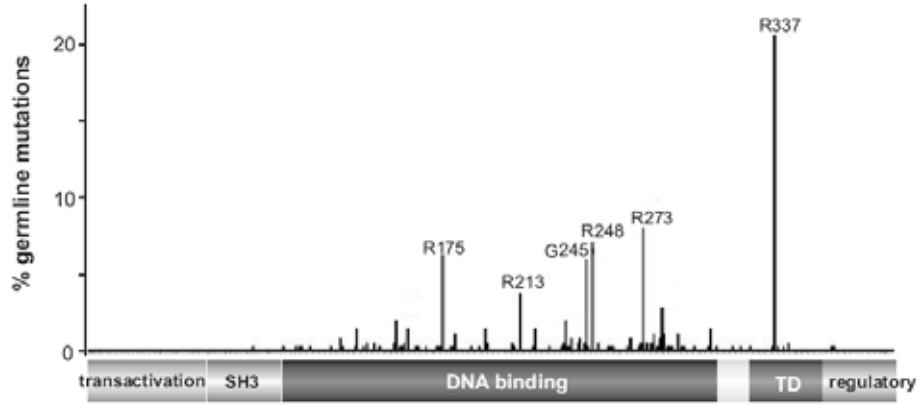


Figure 1.3. Relative frequency of germline mutations along p53 sequence. Codon 337, within the tetramerization domain, is the most frequently affected position in tumors associated to p53 germline mutations. There are two predominant mutations: R337H, associated with ACC in southern Brazil, and, to a lesser extend, R337C, which has been associated with Li-Fraumeni-like syndrome.[63] From the IARC TP53 Mutation Database, release *R12*, November 2007.[46]

---

[a] Most cancers arise from several genetic mutations that accumulate in cells over the lifespan. These are called **somatic mutations**, are tissue specific and the genes involved are usually located on non-sex chromosomes. Cancer may also have a **germline mutation** component, meaning that mutations occur in germ cells (*i.e.* the ovum or the sperm), and thus, they can be passed to offspring. Germline mutations may occur *de novo* (*i.e.* for the first time) or be inherited from parents' germ cells. When germline mutations occur in cancer susceptible genes (such as *p53*), the risk in the individual for developing the disease increases.[53,54]

[b] **Li-Fraumeni syndrome** is a familial inherited multi-cancer syndrome; it is described further in Section 1.1.3.

The tetramerization domain of mutant R337H can adopt a native-like structure, but is less stable than the wild-type domain. The tetrameric stability is highly sensitive to pH in the physiological range (pH 6.5 to 8),[64] and this sensitivity can be correlated with the protonation state of H337.

In the wild-type sequence, the guanidinium moiety from R337 side-chain participates in a salt bridge (plus hydrogen-bond) with the carboxylate group from the D352 side-chain (**Figure 1.4**).[3,65] This ionic pairing, together with some hydrophobic interactions established between the methylene groups from the arginine and other residues (*i.e.* I332, M340, F341 and F348), markedly contribute to the stability of the whole tetrameric structure.[5]

Histidine replacement alters both non-polar and charge characteristics of the side-chain. Firstly, the histidine side-chain is shorter and less hydrophobic than that of arginine; and secondly, its imidazole ring can only be protonated under relatively acidic conditions (*i.e.* when pH < pK_a, **Figure 1.4**). The pK_a for the H337 side-chain was found to be 7.7, which is unusually high for histidines; said pK_a strongly supports the premise that the protonated state is stabilized by a favorable electrostatic interaction with the D352 side-chain.[66]

Therefore, if H337 is protonated, the mutant protein can form a stable tetramer, and roughly function like the wild-type p53. However, the neutralization of the H337-D352 salt bridge at the high end of the physiological range results in an unstable tetramerization domain that is hardly folded under physiological conditions. As such, intracellular pH modulates the function of mutant R337H. The tissue specificity observed in ACC may be associated with the high pH of adrenal cells (~7.9) under certain situations.[67]

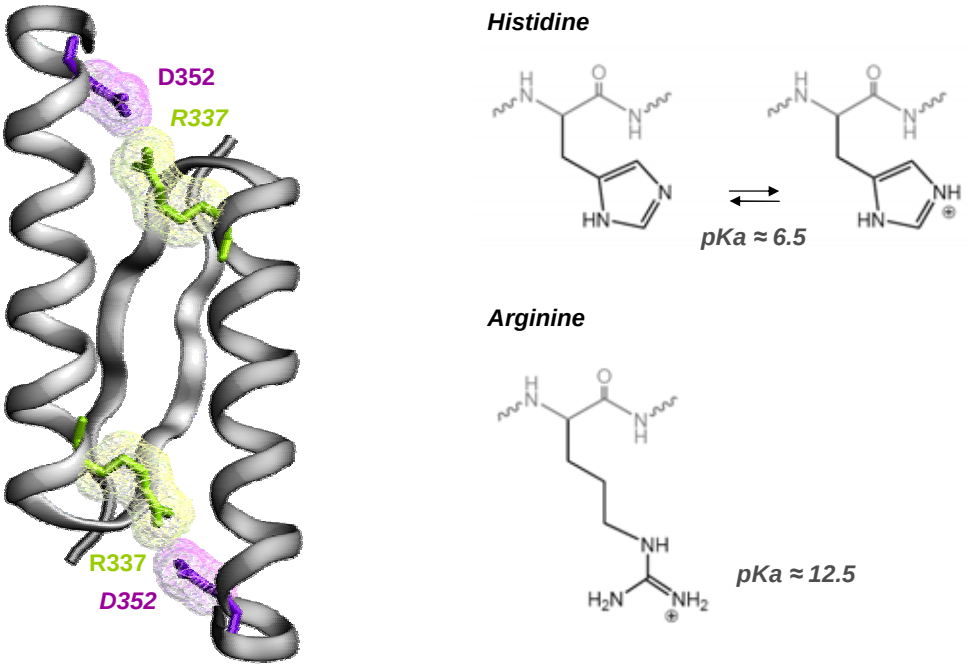


Figure 1.4. Salt bridge between the R337 (green) and D352 (violet) side-chains in a primary dimer of p53TD, and comparison of arginine and histidine side-chains (the positive charges are actually delocalized). Standard pKa values.[68]

1.1.2. Mutant G334V

The *missense point mutation*[c] in codon 334 of *p53* that changes glycine to valine (G334V) has been described in a few cases of human lung cancer.[69-71]

Glycine at position 334 is an irreplaceable residue within p53TD structure, as reflected in its conservation among many species.[18] As glycine lacks a side-chain, it is the only natural amino acid capable of adopting the tight hairpin turnover between the α-helix and the β-strand in the folded p53TD monomer (**Figure 1.5**), which is required for compact packing of the tetramer. Mutation of G334 to valine may not drastically affect the secondary structure elements (*i.e.* the α-helix and the β-strand), but the bulky side-chain of the latter is likely to prevent proper tight packing of the tetramer. Indeed, the lower stability of the resulting tetrameric ensemble[72] correlates well with the low activity of mutant G334V in *in vitro* cellular assays,[48] as well as with the aforementioned reported cases of cancer.

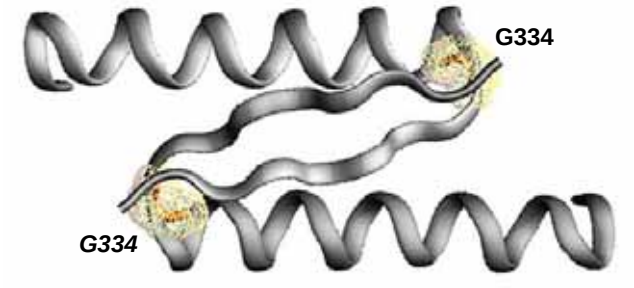


Figure 1.5. G334 position in a primary dimer of p53TD

The most interesting feature of mutant G334V –and the reason why it was selected for the present thesis– is its propensity to rapidly form amyloid fibrils *in vitro* under physiological conditions of temperature and pH, as recently described by Higashimoto and co-workers.[72] These authors suggest that the fibrillogenic potential of G334V might lie in the high tendency of valine to form β-stands,[73] which is the primary secondary structure component of amyloid fibrils. Moreover, they even suggest that loss of p53 activity in mutant G334V might not be due its inability to assemble into a stable functional tetramer, but rather to its tendency to "non-specifically" aggregate, thereby inactivating the entire protein.

[c] **Point mutations** refer to the replacement of a single nucleotide base with another nucleotide base. These can be categorized as *i) nonsense mutations*, which code for a stop; *ii)* ***missense mutations***, which code for a different amino acid; and *iii) silent mutations, which* code for the same amino acid. Point mutations can also result in the insertion or the deletion of a single base pair, which are called *frameshift mutations*.

The amyloidogenic ability of G334V could be also exploited for the design and study of ligands capable of stabilizing the native tetrameric state of protein p53. However, this would first require the reproduction of the amyloidogenesis conditions, which is not a trivial matter.

Finally, it should be mentioned that the work of Higashimoto *et al.* is not the first report of a p53 mutant forming amyloid fibrils. Mutant R248Q, within the DNA binding domain[74], and mutant R337H[75,76] have been also described to form amyloid aggregates. However, those mutants required stringent, unnatural conditions to undergo the β-transformation, not the mild physiological conditions of temperature, pressure, solvent polarity and pH described for mutant G334V.

1.1.3. Mutant L344P

Contrariwise to mutants R337H and G334V, which can assemble the tetramer and retain some activity, mutation of leucine 344 to proline (L344P) leads to an inactive non-tetrameric protein.[49,52] The dramatic loss of function is due to the inability of the mutant to tetramerize, which is predictable given that *(i)* the mutation affects the deep hydrophobic core of the tetramer (**Figure 1.6**), and *(ii)* the introduction of a proline into the middle of the α-helix disrupts the conformation.[51]

Figure 1.6. L344 (in purple) inside the deep hydrophobic core of the p5TD helix bundle

Germline mutation L344P has been reported in families with the Li-Fraumeni syndrome.[77,78] This is a rare, dominant, inherited cancer syndrome characterized by the high incidence of a wide spectrum of neoplasms[d] in children and young adults. Diagnosis of this syndrome requires fulfillment of certain criteria (*e.g.* age, tumor type and family history). The cancer risk in mutation carriers is greater than 70%.[79]

Nevertheless, not all Li-Fraumeni patients carry germline mutations in *p53* gene. Inherited mutations in other genes have also been described for this syndrome –interestingly, most of the proteins which they encode are involved in the p53 signaling pathway.[79] Moreover, not all the *p53* germline mutations result in Li-Fraumeni syndrome, since not all inactivate the protein or cause the characteristic wide range of neoplasms (*e.g.* R337H).[56]

Mutant L344P was selected for this work as a negative control: it has almost the same primary sequence as the wild-type p53TD, but cannot even adopt the secondary structure; hence, it will help to determine any non-specific behavior.

> Final remark:

The present study is only focused on the tetramerization domain of p53 (p53TD). For brevity, the tetramerization domains of each of the aforementioned proteins will hereafter be referred to as **p53wt** (wild-type), **R337H**, **G334V** and **L344P**, but it should be understood that these do not refer to the entire protein.

[d] **Neoplasms** result from the abnormal proliferation of cells in a tissue or organ. A neoplasm that forms a distinct mass is called a **tumor**. Leukemia, despite the fact that forms no mass, is also a neoplasm.

1.2. Obtaining the proteins

1.2.1. Recombinant p53wt tetramerization domain

Previous work in our group (M. Martinell, Ph.D thesis, 2004) established an efficient protocol for obtaining wild-type p53TD. The method is based on the over-expression in *Escherichia coli* cultures and yields milligrams of highly pure protein. In addition to the large quantities produced, recombinant expression is the only affordable way to obtain the isotopically labeled protein – typically with ^{15}N and/or ^{13}C– for NMR characterization.

The clone for the expression of the wild-type p53TD was a kind gift from Dr. M.G. Mateu. The DNA sequence encodes the protein fragment between residues 311-367 (originally named p53_tetS),[5] which is nearly twice longer than the structured tetramerization domain (**Figure 1.7**). The two flanking tails proved to yield better expression levels, but they are randomly coiled and do not affect the tetramerization structure.

Figure 1.7. p53_tetS encoded amino acid sequence

Although the pre-established protocol worked satisfactorily, some modifications were introduced. The most relevant change was use of the state-of-the-art *auto-inducing culture media*, which yielded greater protein production (>50mg/L in rich media[80] and >30mg/L in ^{15}N-minimal media[81]) with lesser costs. Even the high expression levels, the protein was mainly soluble and well-folded.

Taking advantage of the high isoelectric point of p53_tetS (pI ~8.3), the protein is extracted form the soluble fraction of cell lysates through a cation-exchange column. It is then eluted with a sodium chloride gradient and subsequently purified by size-exclusion chromatography (in FPLC at room temperature). The purified protein is desalted and lyophilized for long term storage. p53TD resists lyophilization and is properly folded when solved again.

p53TD can be analyzed by MALDI-TOF mass spectrometry to ensure the correct expression. **Figure 1.8** shows the mass spectra for p53wt. Two mass peaks are always detected: the minor one corresponds to the expected molecular weight for the protein (as a monomer), whereas the major peak, 131 mass units larger, corresponds to the incorporation of a methionine at the N-termini during the recombinant expression.

Protein expression requires the initiation codon `ATG`, which codes for methionine. Since p53TD is only a fragment of the whole protein, said codon was artificially inserted into the DNA sequence. During the expression, endogenous cell enzymes can sometimes remove this additional methionine, although usually the removal does not occur or only occurs partially,[64] as is the case

for p53wt. The presence of the extra methionine at one of the extremes of the unfolded tails, does not affect either the structure or the stability of the tetramer.

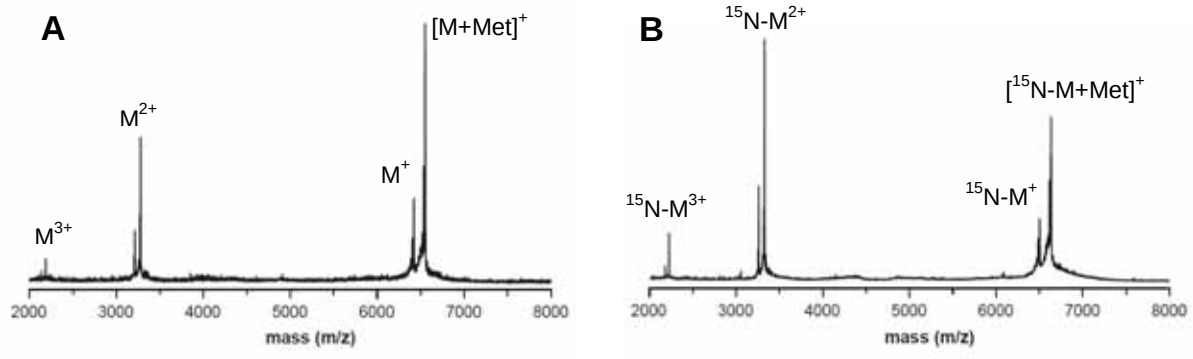


Figure 1.8. MALDI-TOF of **(A)** p53wt (M^+: 6421Da and $[M+Met]^+$: 6552Da) and **(B)** ^{15}N-p53wt (M^+: 6505Da and $[M+Met]^+$: 6637Da)

1.2.2. Mutants R337H & G334V

Clones of mutants R337H and G334V were obtained by site-directed mutagenesis of the p53_tetS clone. This is a well established molecular biology technique that allows introduction of a site-specific mutation into the nucleotide sequence of the double-stranded plasmid. The nucleotide substitutions selected for G334V (GGG → GTG)[71] and for R337H (CGC → CAC)[57] were the same as those reported in the natural mutations.

Both mutants were successfully expressed and purified following the protocol established for the wild-type protein. Although R337H has a lower isoelectric point (~7.8) than p53wt, the cation-exchange purification worked well (**Figure 1.9**). In the size exclusion step, the mutant proteins also eluted as tetramers.

The molecular weight of R337H and G334V was also confirmed by mass spectrometry (**Figure 1.10**). The proteins from different expression batches did not always present the same ratio of methionine content, which was likely due to the different cell strains and different conditions used for the expression. Interestingly, cleavage of the extra methionine in the auto-inducing media was markedly reduced.

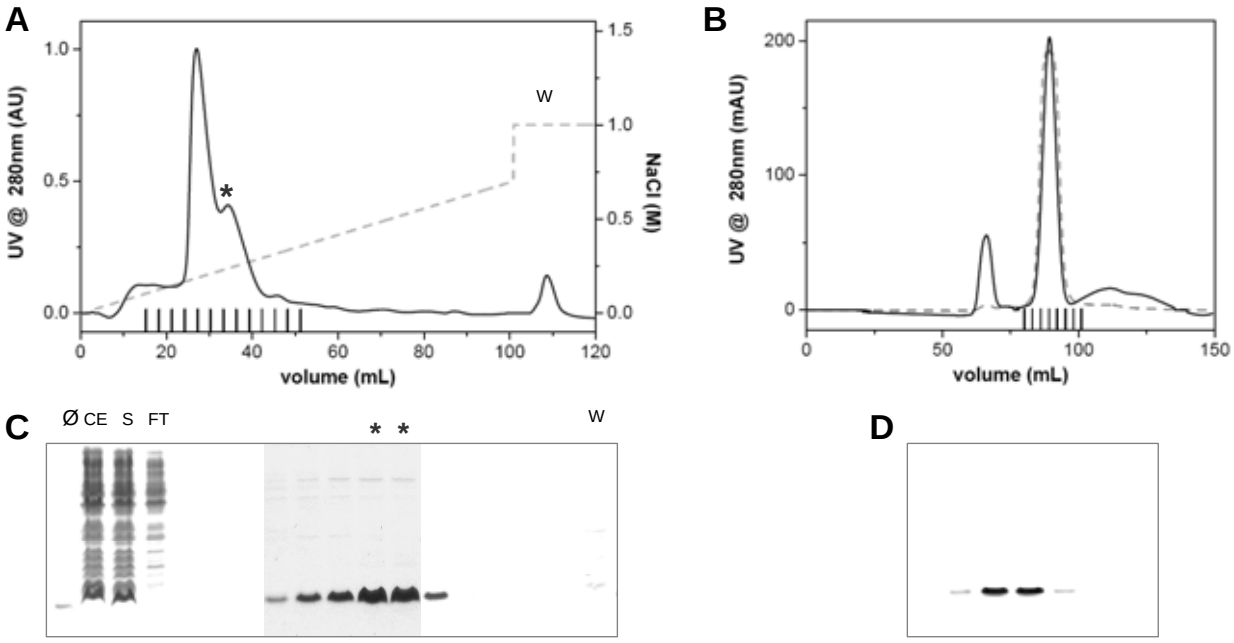


Figure 1.9. Purification of R337H. **(A)** Chromatogram of the cationic-exchange and **(B)** for the size exclusion purification; solid line: absorbance at 280nm; dashed line in **A**: NaCl concentration gradient; dashed line in **B**: absorbance at 220nm. **(C)** and **(D)** SDS-PAGE analysis of the purification fractions (each traced in the chromatograms as vertical lines); Ø: p53wt control; CE: raw cell extracts and S: its soluble fraction; FT: flow-through; W: final wash.

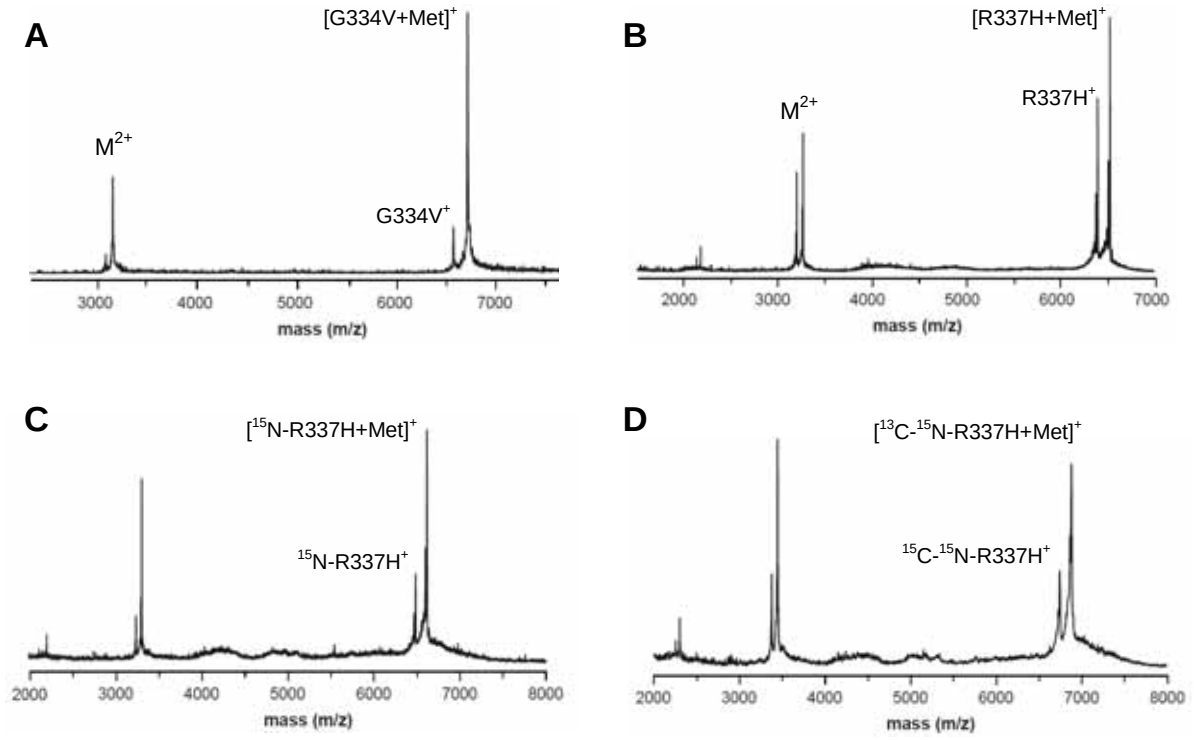


Figure 1.10. MALDI-TOF of **(A)** G334V (M^+: 6463Da and $[M+Met]^+$: 6594Da), **(B)** R337H (M^+: 6402Da and $[M+Met]^+$: 6533Da), **(C)** ^{15}N-R337H (M^+: 6486Da and $[M+Met]^+$: 6617Da) and **(D)** ^{13}C-^{15}N-R337H (M^+: 6762Da and $[M+Met]^+$: 6899Da). G334V was expressed in auto-inducing media, and R337H in LB or M9 media.

1.2.3. Mutant L344P

Obtaining the mutated protein L344P became a challenging biological exercise.

Site-directed mutagenesis of the former clone p53_tetS (GTG → CCG)[63] worked well. However, the protein was not expressed or at least, it was neither detected in the cell lysates, nor extracted by the cation-exchange column. Owing to the unstructured nature of L344P,[51] proteases within the cell likely degraded the polypeptide as it was being expressed. Changing the cell strand, lowering the expression temperature and adding protease inhibitors to the cell extracts did not yield any improvement.

Following the strategy used by Davison and co-workers,[51] the DNA sequence for L344P (residues 311-367) was cloned from the former pET-23b plasmid into a pETM-11 plasmid, in order to introduce a TEV-removable histidine tag (His_6 or His-tag) at the N-terminus (**Figure 1.11**). The His-tag does not confer protease resistance or enhances the solubility of proteins, but the purification of His-tagged proteins through Ni^{2+}-chelating columns is extremely efficient and selective, and even minute amounts of His_6-tagged proteins can be successfully recovered from the cell extracts.

Following the affinity purification, the His-tag would be removed from the sequence by digestion with the tobacco etch virus enzyme (TEV), which specifically recognizes and cleaves the sequence ENLYFQ▼G, also introduced in the expressed polypeptide. After TEV cleavage, four extra amino acids would remain in the N-terminal tail; these residues were unavoidable because they came from the restriction enzyme site used to ligate the protein DNA into the vector. Nonetheless, the effect of these four amino acids should be considered negligible in the non-structured polypeptide.

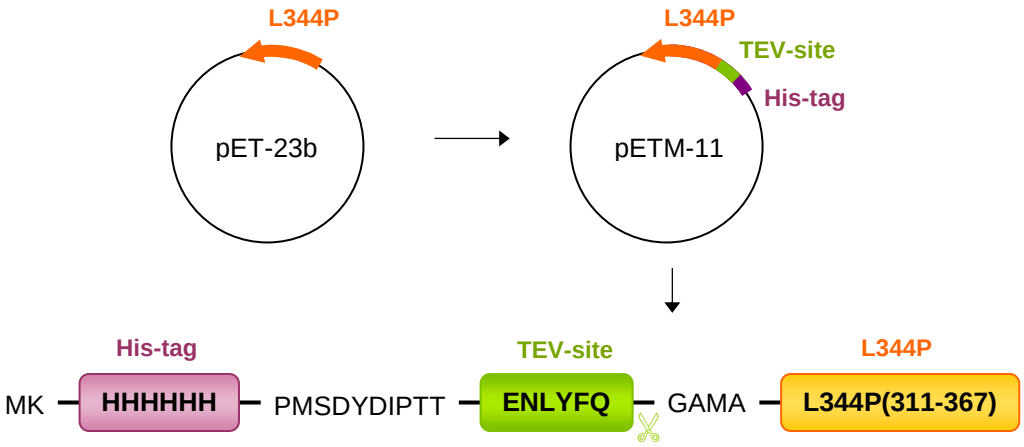


Figure 1.11. New clone and polypeptide for L344P biosynthesis

Once constructed and sequenced, the new clone was transformed into standard competent cells for protein biosynthesis. Expression was carried out at 16ºC to minimize protease activity.

In a first trial, protease inhibitors were added to the cell lysate, and the His-tagged protein was purified at room temperature. Analysis of the fractions by SDS-PAGE did not show any band which could correspond to the protein (~9730Da). The polypeptide was detected in the cell lysate and in the soluble fraction by Western-blot (developing the His-tag), but it was somehow lost in the purification. Since the buffers employed in the purification did not contain any protease inhibitors, the His-tagged protein was probably degraded by the remaining proteases during the washing step.

The most direct solution to abolish remaining protease activity comprised the use of denaturing conditions (*i.e.* 4M of urea) in the lysis buffer. The improvement was evident, as shown in **Figure 1.12**. Interestingly, it was even possible to detect the protein band in the crude cell extract. This strongly suggested that the solubility of the protein was rather low and can be improved by the use of urea in the lysis buffer.

The denaturant was removed from the buffers early during the washing step of the loaded Ni-chelating beads; thereby the final eluted fraction was urea free and ready for His-tag removal. TEV proteolysis worked well, and Edman degradation confirmed the removal of the His-tag from the N-terminus. L344P samples treated with TEV enzyme always crashed out of solution (likely due to the low solubility of the polypeptide); the precipitation was partially avoided by diluting the samples.

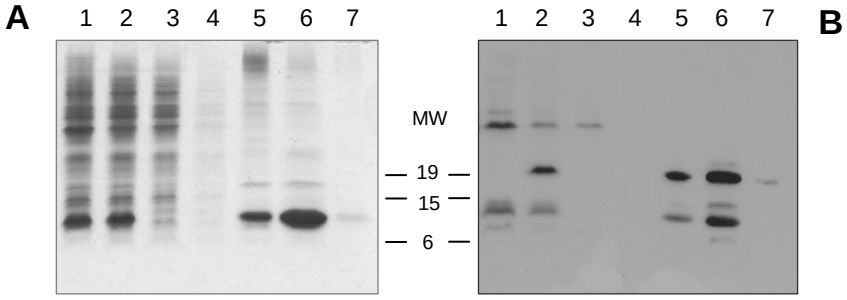


Figure 1.12. Purification of His-tagged L344P. **(A)** SDS-PAGE and **(B)** Western-blot analysis of (1) the cell lysates and (2) its soluble fraction; (3) the flow-through; (4) the washes of the loaded Ni-chelating beads; (5) the loaded Ni-chelating beads; (6) the eluted fraction with 500mM of imidazole; and (7) the Ni-chelating beads after the elution. MW marks in kDa.

Finally, L344P was purified from all the His-tagged species: non-cleaved L344P, TEV enzyme (which was also His-tagged) and other impurities. This was achieved by first dialyzing the sample in order to remove tampers (*i.e.* the imidazole from the initial elution, plus the β-mercaptoethanol and the EDTA from the TEV digestion buffer), and then passing it through a Ni-chelating column and collecting the flow-through fraction (**Figure 1.13**).

The resulting L344P was less pure than the other p53TD proteins (**Figure 1.13**, lane 5). Further purification by size exclusion was also tried; however, taking into account that L344P was a negative control, this was discarded because the minor improvement in the purity did not compensate the time and the loss of product.

As for the other proteins, L344P was desalted and lyophilized.

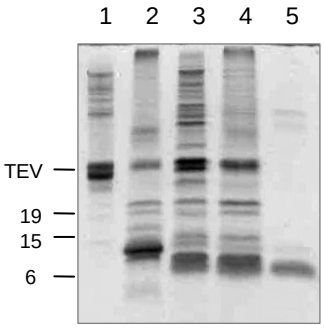


Figure 1.13. SDS-PAGE gel for the purification of TEV-cleaved L344P: (1) TEV enzyme; (2) His-tagged L344P before and (3) after TEV digestion; (4) sample after dialysis and (5) final L344P (*i.e.* the flow-through from the Ni-chelating column).

L344P chemical synthesis

During the first stages, when L344P could not be expressed, the problem was also approached from a synthetic perspective. In fact, L344P is simply a non-structured 56 residue peptide (**Figure 1.14**); hence, its synthesis was envisioned to be straightforward.

```
56          50              40              30              20              10              1
NTSSSPQPKKKPLDGEYFTLQIRGRERFEMFREPNEALELKDAQAGKEPGSRAHSS
```

Figure 1.14. Amino acid sequence of L344P (numbered in the coupling order)

The synthesis was carried out automatically, by Fmoc / ^{t}Butyl chemistry, using the novel polyethyleneglycol resin ChemMatrix as polymeric support. Small aliquots of peptidyl-resin were taken to evaluate the progress of the couplings –although the synthesis was never stopped. The results are shown in **Figure 1.15**. The cleanness of the crude product up to the 40th residue was excellent. However, the right coupling pace was truncated somewhere within the last fifteen residues. Despite the fact that it was possible to detect the mass of L344P by MALDI-TOF analysis, the main product at the end of the synthesis corresponded to an unidentified sequence that was ten to fifteen residues shorter. Several tests for the cleavage and deprotection were performed, but these resulted in the same profile.

No further work was done on the synthesis, as by that time the recombinant expression started to improve. Nonetheless, the peptide probably could have been synthesized without investing much more effort.

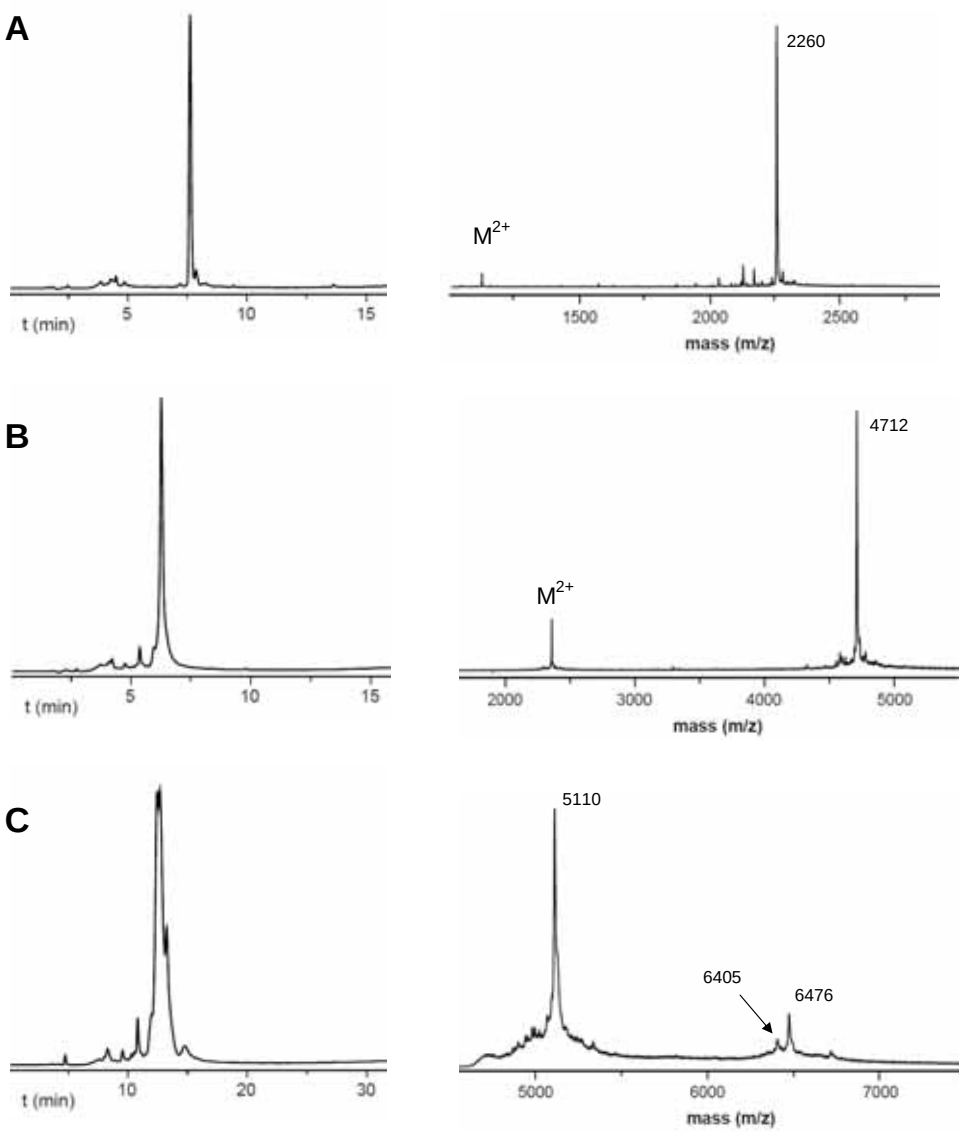


Figure 1.15. RP-HPLC chromatograms at 220nm (left panels) and MALDI-TOF spectra (right panels) for the progress of the L344P chemical synthesis. **(A)** Fmoc-[L19-S1] (C18 t_R: 7.6min, expected mass: 2257Da); **(B)** [Y40-S1] (C18 t_R: 6.3min, expected mass: 4708Da); and **(C)** final peptide (C4, expected mass: 6403Da). HPLC gradient 0-100% ACN in water, in 15min for the C18 column, and in 30min for the C4 column. Reagent K was used for cleaving the peptides from the resin and removing the side-chains protective groups.

1.2.4. Cutting the tails: synthetic p53 tetramerization domain

Although the aforementioned random coiled tails of the recombinant protein do not affect the tetramerization of p53TD in solution, they do complicate the crystallization of the protein.[3] Hence, they had to be removed.

In a first trial, reproducing the strategy described by Jeffrey *et al.*,[3] the biosynthesized protein (p53_tetS) was digested with subtilisin,[82] which is a non-specific protease that was supposed to kindly degrade the random tails without affecting the structured core of the tetramer. Regrettably, as shown in **Figure 1.16**, the proteolytic digestion was excessive and the subtilisin digested *too much* of p53TD.

Figure 1.16. p53_tetS (1) before and (2) after subtilisin digestion.

Several parameters (*e.g.* temperature, concentrations or buffers) were modified but the subtilisin digestion never yielded good results. Therefore, only one option remained: the solid phase peptide synthesis of p53TD. The fragment between residues 320-356 of p53[3] was automatically synthesized by Fmoc / ^{t}Butyl chemistry on ChemMatrix resin. Chromatograms of the relatively neat crude obtained after the peptide cleavage and side-chains deprotection, and of the final product purified by RP-HPLC, are shown in **Figure 1.17**.

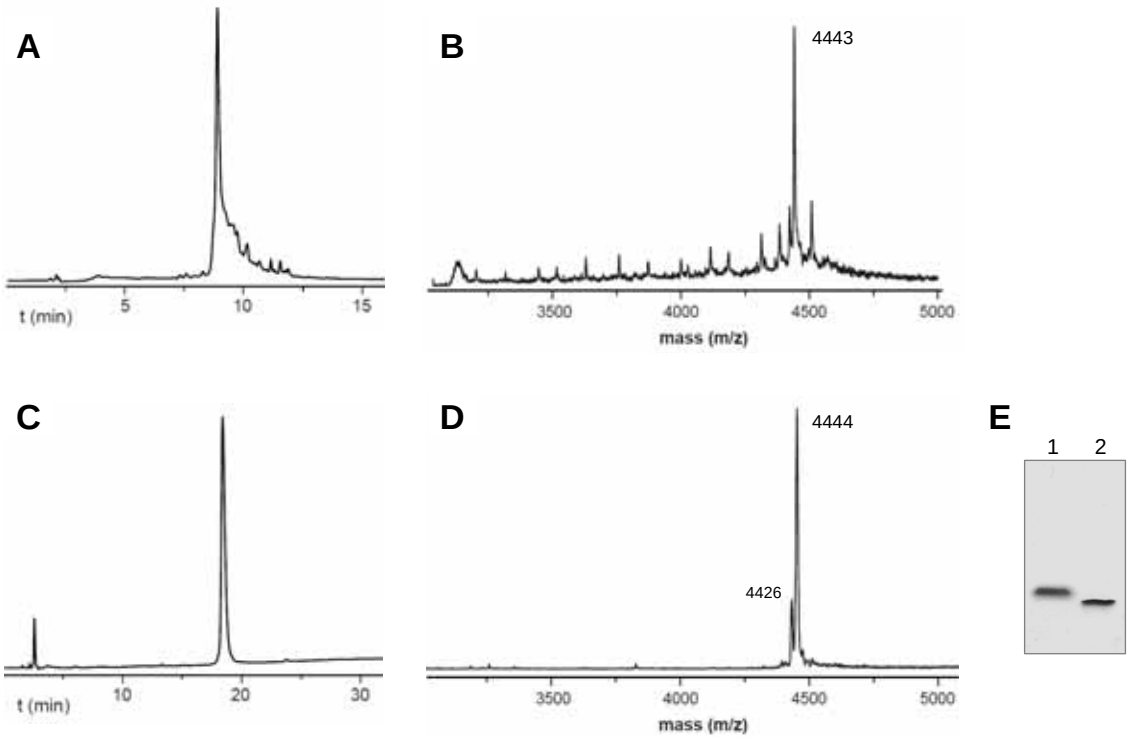


Figure 1.17. p53TD peptide synthesis. **(A)** HPLC (t_R: 8.9min, C18, 0-100% ACN in 15min) and **(B)** MALDI-TOF spectra of the synthetic crude, using reagent K for peptide cleavage (expected mass M^+: 4440Da). **(C)** HPLC (t_R: 19.5min, C4, 0-50% ACN in 30min) and **(D)** MALDI-TOF spectra of the purified peptide. **(E)** SDS-PAGE gel comparing (1) p53_tetS and (2) the synthetic p53TD.

1.3. Biophysical characterization of the proteins

1.3.1. Circular Dichroism

Circular Dichroism (CD) enabled rapid characterization of the global structure of the proteins as well as their thermal stability.

The CD spectra for each of the proteins under the same conditions are compared in **Figure 1.18**. The spectrum of wild-type p53TD presents the characteristic features of a protein containing α-helices, β-sheets and random coils. The contribution of the unstructured tails was determined by comparison of the CD traces of the recombinant protein and of the chemically synthesized domain. The CD signal of these tails did not allow use of the so-called *parameter q* ($\theta_{220}/\theta_{208}$) to assess the existence of a single stranded α-helix (q <1) or interacting helices (q ≥1)[83] and hence, determine the level of tetramerization of the mutant proteins.

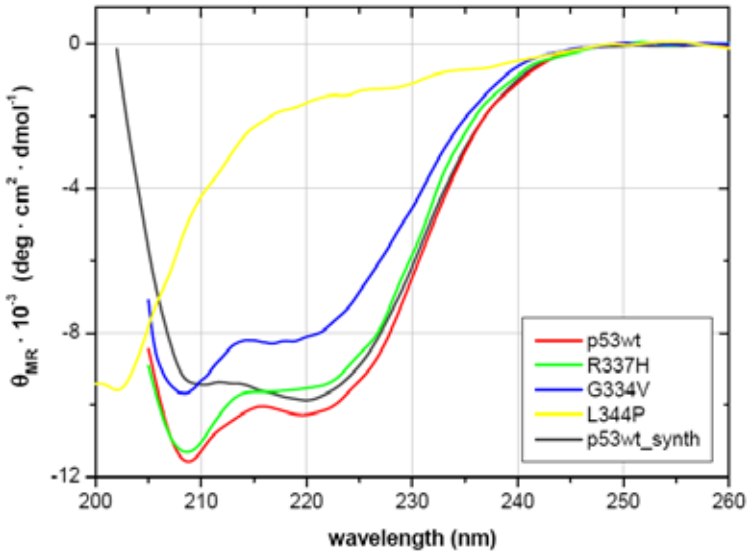


Figure 1.18. CD spectra of the biosynthesized p53wt (red), R337H (green), G334V (blue), L344P (yellow) and the synthetic p53TD (grey), at 15μM (monomer) in 25mM sodium phosphate buffer pH 7, 100mM NaCl and 20ºC. Ellipticity is normalized to the mean residue concentration (*i.e.* monomer concentration × number of amino acids), θ_{MR}.

Mutant R337H displayed basically the same levels of secondary structure than the wild-type protein.[50] Contrariwise, the CD signal intensity for mutant G334V was much lower, which suggested that this protein was less structured under those conditions.[72] The lack of structure in L344P was also evident from the CD spectrum.

(a) Protein concentration effects

As a tetramer, p53TD self-assembly equilibrium depends on the total protein concentration, and consequently, its thermal stability too.

The thermodynamic "self-dissociation" constant for the wild-type p53TD has been estimated to be really low,[7,84] and the tetramer is the species usually observed under the typical conditions of many biophysical techniques. However, the concentration of protein can be greatly reduced in circular dichroism experiments and thereby changes in the oligomerization equilibrium of p53TD can even be detected. As shown in **Figure 1.19**, the intensity of the molar ellipticity of the structured p53TD increased progressively with the protein concentration, until it got to a "saturation" point, at which the tetramerization equilibrium was totally shifted. Differences in the progress of the CD spectra revealed that each protein self-assembled with different affinity.

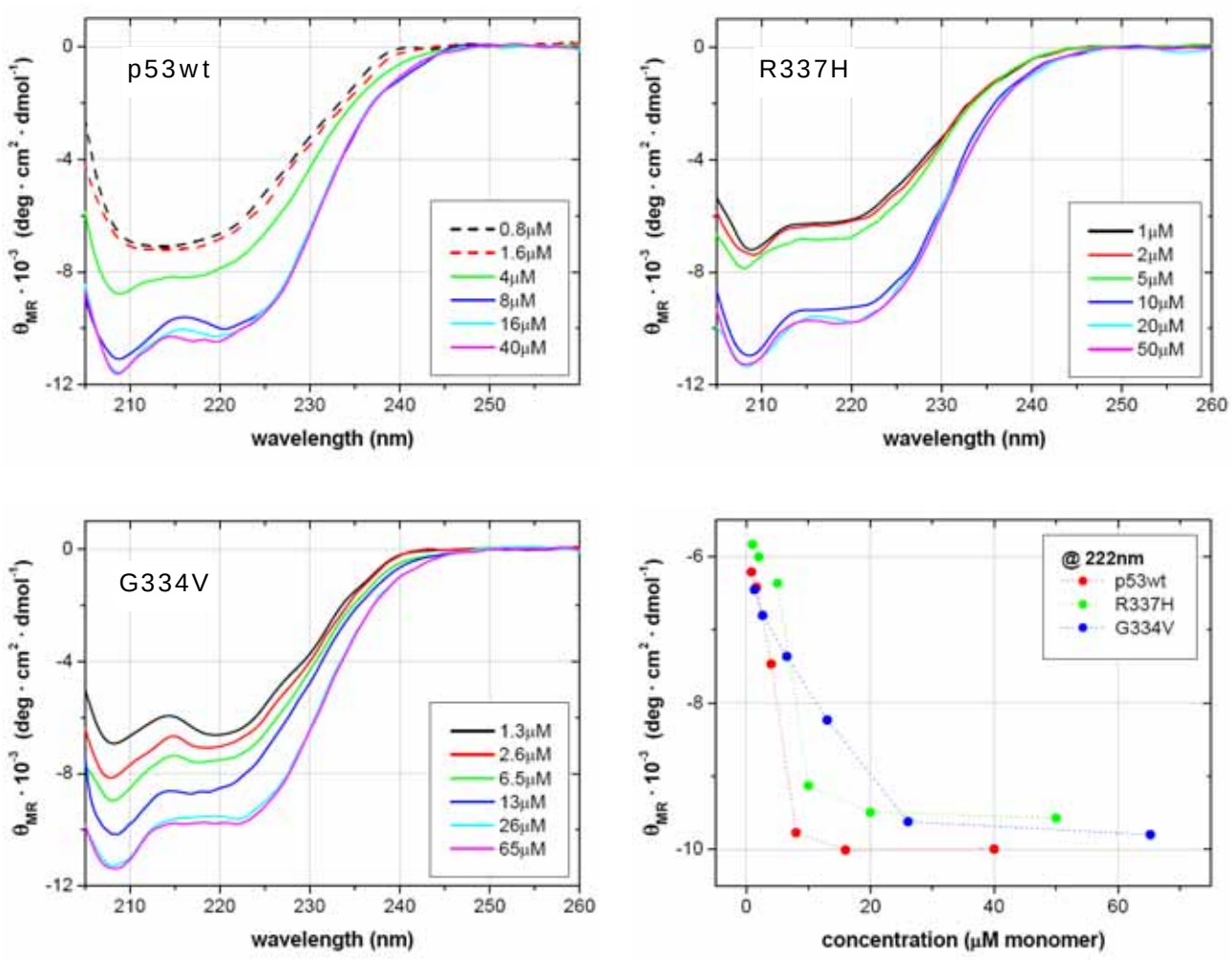


Figure 1.19. Effects of the protein concentration in CD molar ellipticity (in 25mM sodium phosphate buffer pH 7, 100mM NaCl at 20ºC; protein concentrations in micromolar (μM) correspond to monomer). The dashed lines for p53wt indicate that experimental data are not accurate due to the low signal detected.

Wild-type p53TD displayed the highest "self-assembly affinity", whereas G334V required larger concentrations to totally shift the tetramerization equilibrium (**Figure 1.19**).

Concentration effects on protein thermal stability were monitored by tracing the changes in the ellipticity at 220nm while increasing the temperature. The unfolding profile for the three proteins (at ~10μM monomer) are compared in **Figure 1.20A**. Mutant proteins were notably less stable than the wild-type tetramer. Interestingly, mutant G334V was thermally more stable than R337H at any concentration (**Figure 1.20B**), although at room temperature G334V was clearly less structured and had lower tetramerization affinity.

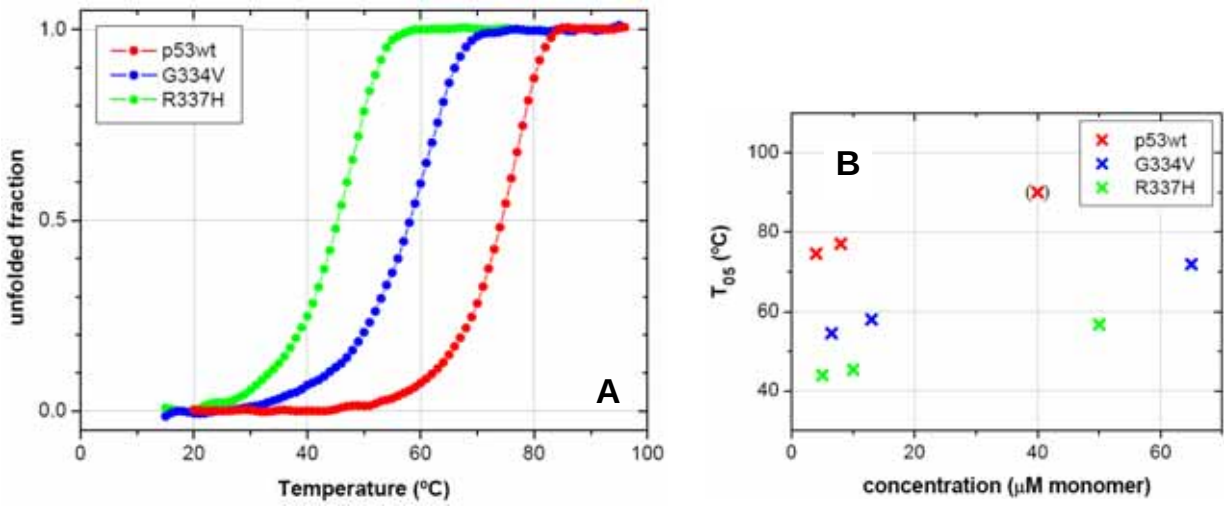


Figure 1.20. (A) CD melting curves of 8μM p53wt (red), 13μM G334V (blue) and 10μM R337H (green) (monomer concentrations, 25mM sodium phosphate buffer pH 7, 100mM NaCl). **(B)** Thermal stability (T_{05}) as a function of protein concentration (the point in brackets for p53wt was extrapolated from an incomplete curve; hence, it is fairly inaccurate).

The unfolding mechanism for p53wt was described by Freire and co-workers as a reversible two-state process ($N_4 \leftrightarrows 4U$), in which the folded tetramer is directly converted into the unfolded monomer.[4] The range of transition temperatures is narrow, and therefore, the unfolding is highly cooperative. Since the unfolding transition is coupled to the tetramer dissociation, the melting profile is not a perfectly symmetric curve, and is slightly skewed to the low-temperature side.[85]

The melting curve of mutant R337H (at 10μM, **Figure 1.20A**) was slightly more symmetrical than that for p53wt, which might suggest that the unfolding was not totally coupled to the dissociation of a tetramer. The current data was not sufficient to assess what the exact unfolding mechanism was like. Perhaps, the unfolding proceeded through an intermediate species; or perhaps, the tetramerization equilibrium became less shifted at higher temperatures and the unfolding underwent through more complex mechanisms.

It is worth mentioning that at higher concentrations (*e.g.* 50μM monomer) the melting profile of R337H became less symmetric –skewed to the low-temperature side; this suggested an unfolding

event more coupled to dissociation, probably because the higher concentration of protein greatly shifted the equilibrium towards the tetramer. The unfolding curves at different concentration are provided in the Supplementary Material.

The symmetry of the unfolding profile of G334V was similar to that of p53wt; hence, the unfolding was likely following the same mechanism, although the mutant protein was less tetrameric at low temperatures. Actually, from the less shifted tetramerization equilibrium one would have expected –beforehand– that G334V might be much less stable (at least, lesser than R337H). However, protein behaviors can change at higher temperature and this was a clear example. Unfortunately, the unfolding trace did not help in understanding what made the protein to become more stable.

The inflexion point in a CD unfolding curve corresponds to the melting temperature, $\mathbf{T_m}$. Given that the curves for p53TD are not symmetrical, said temperature does not coincide with that where the transition is half completed, $\mathbf{T_{05}}$ (f_u = 0.5). Neither T_m nor T_{05} are the so-called reference temperature, $\mathbf{T^0}$, at which $\Delta G^0 = 0$ (*i.e.* $[U]^4/[N_4] = 1$).[4]

The inflexion point of the experimental CD unfolding curves could not be determined accurately (*i.e.* the maximum of the unfolding curve derivative was fairly imprecise); hence, T_{05} was preferred to characterize the thermal stability of the proteins. Nevertheless, the mathematical normalization of the unfolding curve assuming a two-state unfolding was not totally correct for mutant R337H, for which the unfolding mechanism was not clear (and thus, the estimated T_{05} might not exactly correspond to f_u = 0.5). Regardless of these limitations, T_{05} was a parameter that could be determined with enough experimental accuracy and provided a number to compare the thermal stability of the proteins.

(b) Buffer effects and R337H pH-dependence

The stability of mutant R337H has been described to be extremely sensitive to the pH in the physiological range.[50] This sensitivity agrees with the protonation state of the H337 side-chain: if the histidine imidazole ring is protonated (at acidic pH), it can roughly substitute the native salt bridge with D352, and thereby stabilize the tetrameric ensemble.

The pH-sensitivity of R337H is perfectly illustrated in the progressive shift of the CD unfolding curves towards lower temperatures as the pH is basified (**Figure 1.21B**).[50] Under acidic conditions, at which the H337 is supposed to be mainly protonated (pK_a: 7.7),[50] the mutant protein (T_{05}: 54ºC) was still far from being as stable as the wild-type domain (T_{05}: 72ºC). In the range between pH 6 and 8, slight variations in the basicity caused large changes in the thermal stability of R337H. The symmetry of the melting curve in this intermediate range of pH also changed, which strongly suggested that the change in the unfolding pathway could result from the protonation state of the H337 side-chain (*i.e.* from a shift in the tetramerization equilibrium).

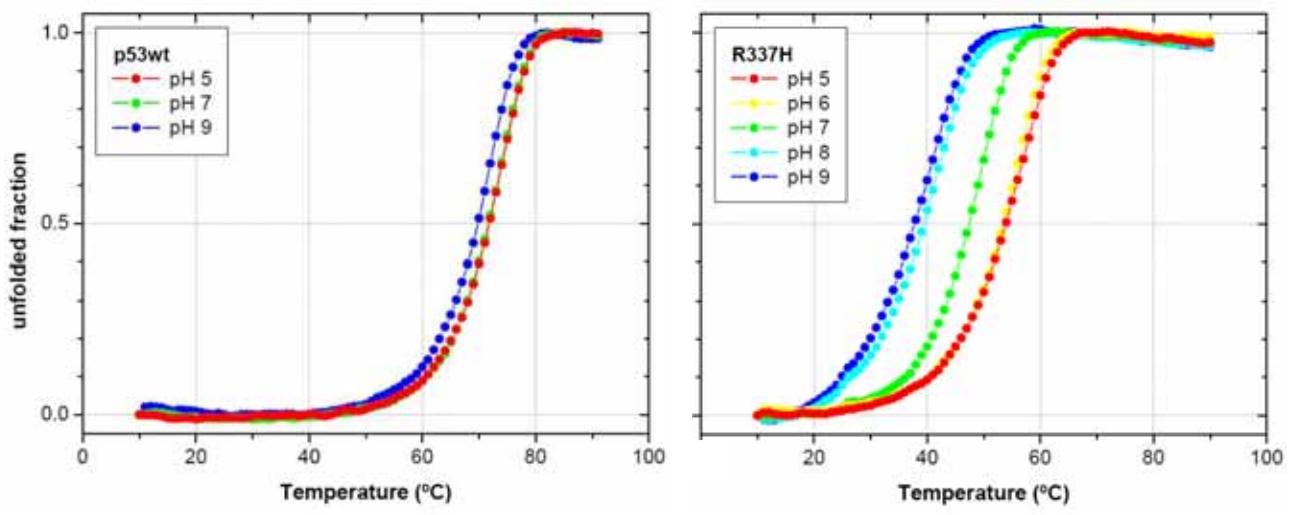


Figure 1.21. CD melting curves of 8μM p53wt and 10μM R337H (monomer) at several pHs. Experiments reproduced from DiGiammarino *et al.* [50] (20mM phosphate buffer, 250mM NaCl)

On the contrary, p53wt thermal stability was not so pH-sensitivite; in fact, too acidic conditions appeared to be slightly destabilizing (**Figure 1.21A**).[4]

Ionic strength also influenced the thermal stability of the proteins. In general, the lager the salt concentration, the lesser the stability (**Figure 1.22**). The changes were reproducible, and large enough as not to be confused with experimental error. The CD spectra at room temperature were unaffected by the buffer; hence, it seemed that the ionic strength affected the unfolding pathway but not the secondary (or quaternary) structure.

These results were inexplicably suspicious. Freire *et al.* reported the opposite for p53wt stability,[4] which is actually the expected behavior for a protein whose oligomerization driving force is hydrophobicity. Regrettably, the origins of this trend remain unknown.

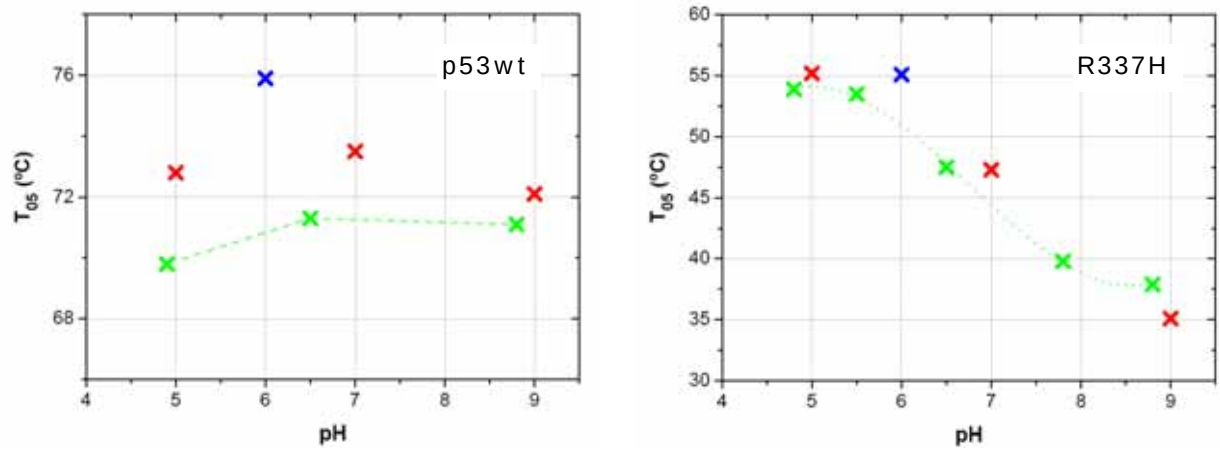


Figure 1.22. Effects of ionic strength on the thermal stability of 8μM p53wt and 10μM R337H (monomer). Buffers: X water, X 20mM NaPi and X 20mM NaPi + 250mM NaCl.

(c) Time stability and G334V aggregation

Mutant G334V was originally selected because of its reported ability to rapidly form amyloid aggregates under physiological conditions. In order to take advantage of this property, the amyloidogenesis process had to be reproduced. Working under the conditions of concentration, buffer, temperature and time described by Higashimoto *et al.*,[72] no aggregation was observed (**Figure 1.23D**). Incubating the protein at 37ºC for longer periods only resulted in spontaneous denaturation; the lack of a clear isodichroic point in these curves suggested that the protein might not denature through a simple two-state mechanism.[4]

The lack of reproduction in the amyloid aggregation was not surprising indeed. Higashimoto *et al.*[72] reported the β-aggregation property for a synthetic G334V which comprised the fragment 319-358. Our recombinant protein G334V was longer (311-367), and the unstructured tails at the extremes were certainly affecting the aggregation process. In addition, the biosynthesized protein was purified through a different protocol than the synthetic polypeptide, and that could also alter the amyloidogenesis.

Changing the salt concentration (**Figure 1.23C**) or working in water (**Figure 1.23E**) did not lead to amyloid aggregates either. In fact, larger ionic strengths had also a destabilizing effect over the time. Furthermore, the time-denaturation pathway in water seemed to be different than the followed in phosphate buffer.

The same experiments on the other proteins (**Figure 1.23A** and **B**) did not show anything but spontaneous denaturation. Mutant R337H displayed an isodichroic point at ~205nm,[4] thus suggesting a two-state denaturation mechanism. It must be noted, however, that the unfolding pathway followed by a protein while it is rapidly heated, has not to be the same than the denaturation mechanism which undergoes spontaneously over the time at lower temperatures.

Interestingly, G334V and R337H appeared to denature at a similar rate during the time (under the same conditions of buffer, **Figure 1.23F**). This markedly differed from the thermal stability results (G334V was thermally more stable, **Figure 1.20**), and suggested that G334V was probably undergoing a different unfolding process when it was incubated at 37ºC for long periods. Perhaps, it was eventually related to the propensity of this mutant to aggregate, although no clear aggregation event was detected.

Time, temperature, buffer composition and protein concentration conditioned the unfolding mechanisms of G334V. At high concentration (*e.g.* 100μM monomer), G334V crashed out of solution during the CD unfolding experiment. The precipitation event was faster at the slower heating rates, and its appearance depended on the buffer: in phosphate buffer, a white crush was visibly detected in the cell; whereas in water, nothing could be visibly detected. The said G334V aggregates got tightly stuck on the cell walls, and stringent washing conditions (*i.e.* chromic mixture) were required to completely clean the cell (Supplementary Material).

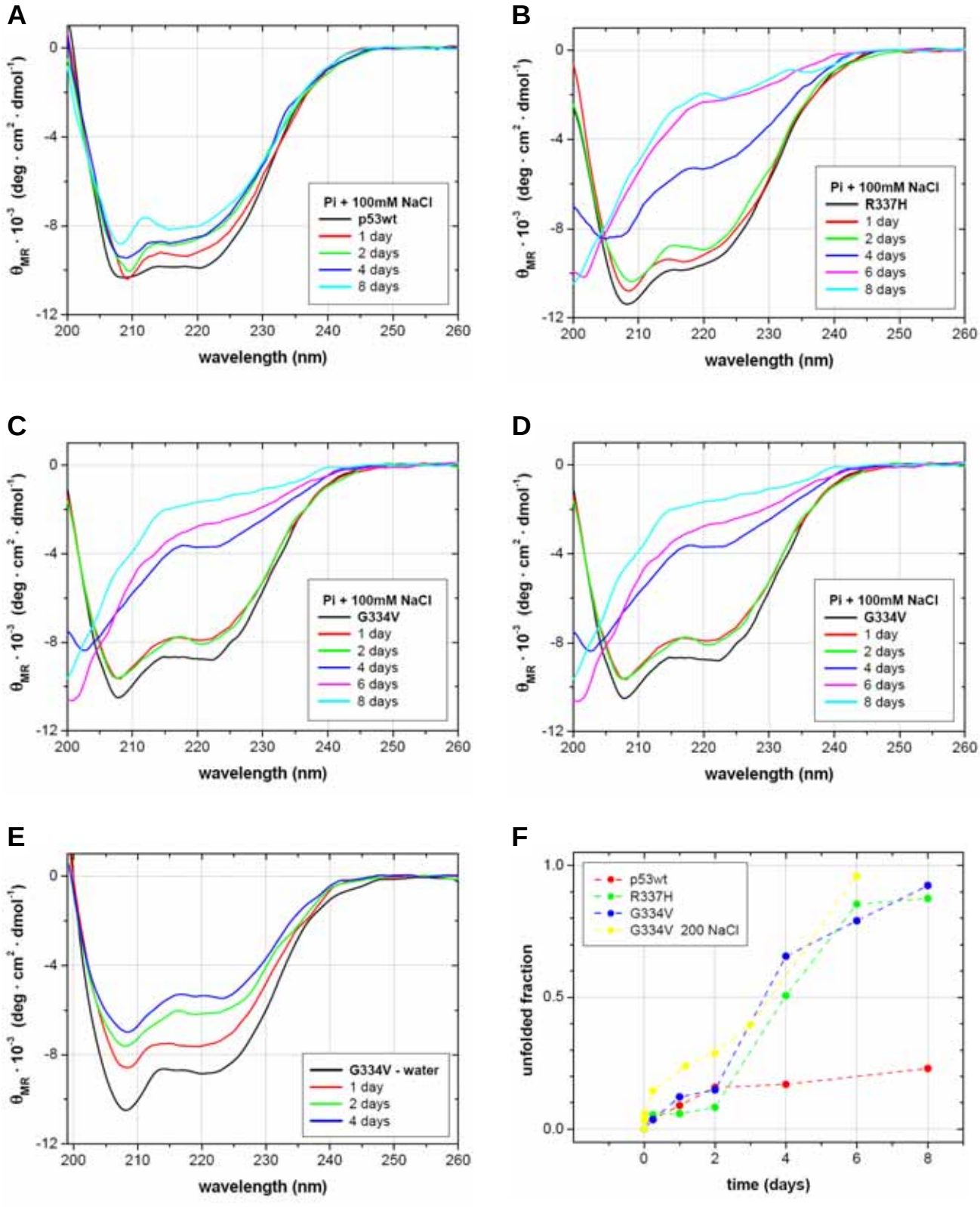


Figure 1.23. Time stability. CD spectra of **(A)** p53wt, **(B)** R337H and **(C)** G334V incubated at 37ºC for several days, in 50mM sodium phosphate buffer pH 7 + 100mM NaCl, at a concentration of 10μM (monomer). **(D)** 10μM G334V (monomer) in 50mM sodium phosphate buffer pH 7 + 200μM NaCl and **(E)** 30μM G334V (monomer) in water pH 7, also incubated at 37ºC. Interestingly, in water, although the concentration is higher, the tetratemization is lower. **(F)** Representation of the denatured protein fraction through the time (tracking θ at 220nm from **A**, **B**, **C** and **D**).

1.3.2. Differential Scanning Calorimetry

The energetics of the protein unfolding was characterized by Differential Scanning Calorimetry (DSC). Besides thermodynamic parameters and melting temperatures, DCS endotherms also provided information about the unfolding mechanisms. The concentration of protein required for recording high quality thermograms was large; hence, DCS complemented CD.

The DSC curves for p53wt and R337H under the same experimental conditions are compared in **Figure 1.24**. In agreement with its larger stability, p53wt unfolding transition occurred at higher temperatures and was more energetic (*i.e.* larger melting enthalpy, ΔH_m: area under the curve). The sharpness of the transition peak of p53wt proved that the unfolding was highly cooperative. Furthermore, the asymmetry of the bell, which was skewed to the low-temperature range, indicated that the protein was an oligomer (indeed, a tetramer) and its dissociation was coupled to the protein unfolding.[85] This asymmetry made that the melting temperature T_m –defined as the temperature at which the heat capacity reaches its maximum– did not correspond to the temperature at which the unfolding enthalpy had proceeded one half, T_{05}.[85]

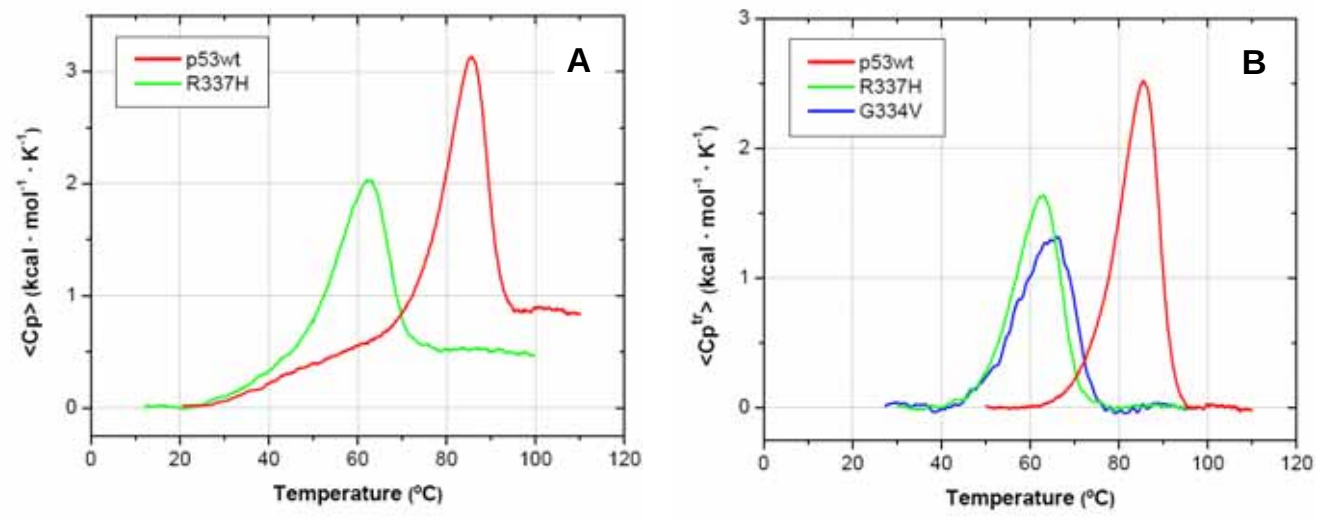


Figure 1.24. (A) DSC thermograms of 100μM (monomer) p53wt (red) and R337H (green), in water at pH 7 (<C_P> normalized to monomer concentration). For a better comparison, the curves have been y-shifted (C_P=0 at 20ºC). **(B)** Unfolding transition peaks. G334V (blue) sample contained 50μM monomer in 25mM phosphate buffer at pH 7. The raw recorded endotherm for G334V is not shown due to a distortion in the baseline.

Freire and co-workers developed an equation to deal with the thermodynamics of the two-state concentration-dependent self-associating thermal unfolding of p53wt tetramerization domain.[4] This mathematical model not only fit the experimental curve of p53wt, but also the curve of R337H (Supplementary Material). Regardless that the thermogram of R337H was not as sharp or

asymmetric as that of p53wt, the adjustment proved that R337H also unfolded as a tetramer (**Table 1.1**). This agreed with the asymmetric trend detected in the CD unfolding curves of R337H at high concentrations. Therefore, even if H337 was not fully protonated, R337H unfolding was coupled to the tetramer dissociation ($N_4 \leftrightarrows 4U$) at high protein concentrations, at which the tetramerization equilibrium was largely shifted.

Table 1.1. Experimental and calculated thermodynamic parameters from DSC (energetic values per mol of monomer)

		p53wt	R337H	G334V
experimental	**monomer (μM)**	100	100	50
	buffer	water, pH 7.0	water, pH 7.0	25mM Pi, pH 7.0
	T_m (ºC)	85.5	63.1	66.1
	T_{05} (ºC)	83.5	60.4	63.1
	$T_{½}$ (ºC)*	10.5	12.1	14.9
	ΔC_P (cal·mol^{-1})	525	180	na
	ΔH_m (kcal·mol^{-1})	35.0	21.6	20.3
	ΔS_m (cal·mol^{-1}·K^{-1})	98	66	61
calculated**	**n**	4	4	na
	T^0 (ºC)	94.0	73.8	na
	ΔC_P^0 (cal·mol^{-1})	775	153	na
	ΔH_m^0 (kcal·mol^{-1})	43.7	30.5	na

* $T_{½}$ corresponds to the width of the transition peak at the half height.

** ΔC_P^0, ΔH_m^0 at the reference temperatureT^0, at which $\Delta G_m^0 = 0$, *i.e.* $[U]^4/[N_4]=1$

The heat capacity change (ΔC_P) was smaller for R337H than for p53wt (**Figure 1.24A**). ΔC_P roughly correlates with changes in the polarity of exposed surfaces areas.[86] The smaller heat capacity change for R337H might suggest that the "apparent" hydrophobic surface in the folded state was "larger" (*i.e.* the hydrophobic surfaces buried in the tetramer might be exposed to the solvent for apparently longer due to the larger dissociation constant for the tetramer). Nevertheless, other factors were also contributing to the heat capacity change in R337H; among them, it must be mentioned the deprotonation of the H337 side-chain, since its unusually high pK_a in the folded state (~7.7) dropped to the standard values (~6.5) in the unfolded protein.

The reversibility of the unfolding was evaluated by re-scanning several times an unfolded sample. After every scan, the transition peak for both proteins decreased (*ca.* one quarter) and, accordingly, the melting temperature was also shifted towards lower values. These observation suggested that the unfolding was not a completely reversible, which disagreed with the reported

reversibility for p53wt.[4] Such discrepancy likely resulted from the high range of temperatures used here: once unfolded, the protein was further heated up to 120ºC, thus enhancing the chance of irreversible denaturation.

For mutant G334V, the crushing phenomenon detected by circular dichroism was also observed in the DSC curves. Actually, DSC is much more sensitive to protein aggregation (due to the large exothermic energy which results) and it was nearly impossible to obtain a neat transition peak for protein G334V. The trace shown in **Figure 1.24B** was finally achieved by using phosphate buffer, lowering the concentration, fastening the heat rate and employing a less sensitive calorimeter. Again, the precipitated protein got tightly stuck to the cell walls and stringent washing conditions were required to completely remove it (*i.e.* hot SDS solution); otherwise the contamination ruined the next DSC experiment from the beginning. These inconveniences made impossible to study G334V by differential scanning calorimetry.

R337H pH-dependence

The thermal stability of mutant R337H as a function of the pH was also characterized by DSC, as shown in **Figure 1.25A.** The shift towards lower melting temperatures at basic pHs was accompanied by a decrease in the melting enthalpy, and, interestingly, the transition peak changed towards a more symmetrical bell. These observations suggested that, although the mutant protein could assembly and unfold as a tetramer at acidic pHs, the unfolding mechanism changed at basic pHs, probably due to a shift in the tetramerization equilibrium –which might result from the deprotonation of H337 side-chain.

Normalization of the melting enthalpies (**Figure 1.25B**) provided a sketch analogous to the set of melting curves from the CD experiments (**Figure 1.21**). As expected for the larger concentration of protein used in DSC, the melting temperatures were higher. However, the shift within the curves was not exactly the same than the observed by CD. The differences could be due to a change in the pK_a of the H337 side-chain as a function of the concentration: at high protein concentrations, the oligomerization equilibrium would be largely shifted towards the tetramer, thereby it could further favor the protonation of the histidine side-chain (hence, higher pK_a).

The effects of pH and concentration on the melting temperature are compared in **Figure 1.26**. It is worth noting that the melting temperature at pH 9 still depends on the concentration, which means that, whether R337H was protonated or not, the tetramer could be still assembled.

The enthalpy resulting from the protein unfolding did not evolve linearly either with the melting temperature or the pH (the plots are provided in the Supplementary Material),[87] which was logical given that it not only included the energy from the unfolding process but also the protonation and deprotonation events. Consequently, ΔC_P depended on both the temperature and the pH.

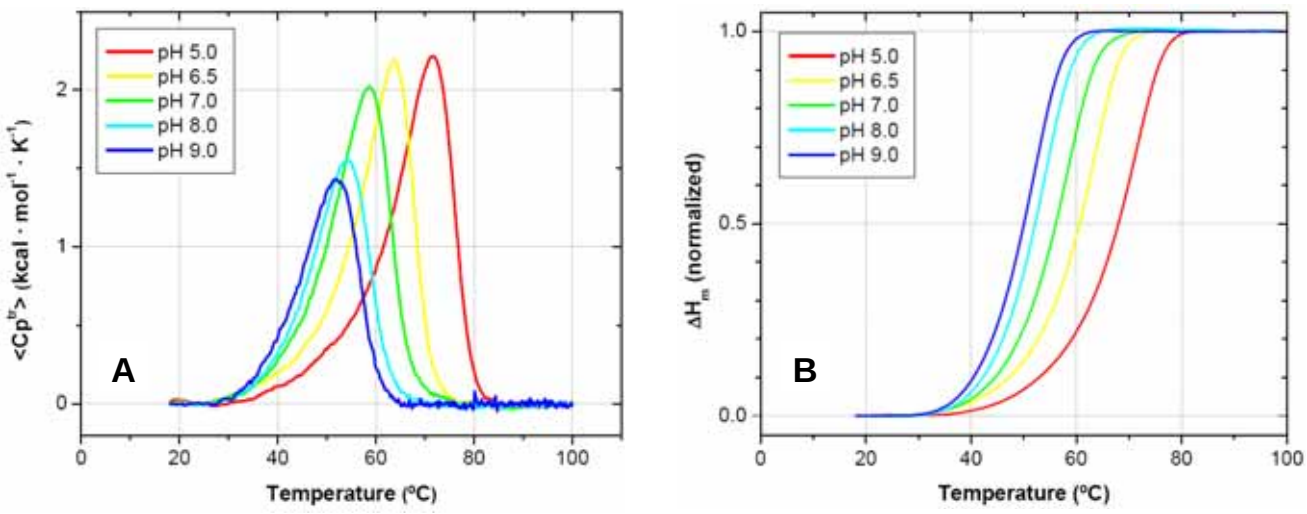


Figure 1.25. (A) DSC endotherms of 172µM R337H (monomer) at different pHs (in 25mM phosphate buffer) and **(B)** its normalized melting enthalpy. Enthalpies correspond to the integration of the partial molar heat capacity of the transition, $<C_p^{tr}>$, over the temperature range.

Finally, it must be mentioned that the destabilizing effects of the ionic strength detected by CD were also observed in DSC (**Figure 1.26**). Furthermore, although the melting temperature for the protein decreased, the unfolding enthalpy was higher, which agreed well with the fact that the hydrophobic core was being exposed to solvent during the unfolding. Unfortunately, no explanation has been found for the destabilizing salt effect.

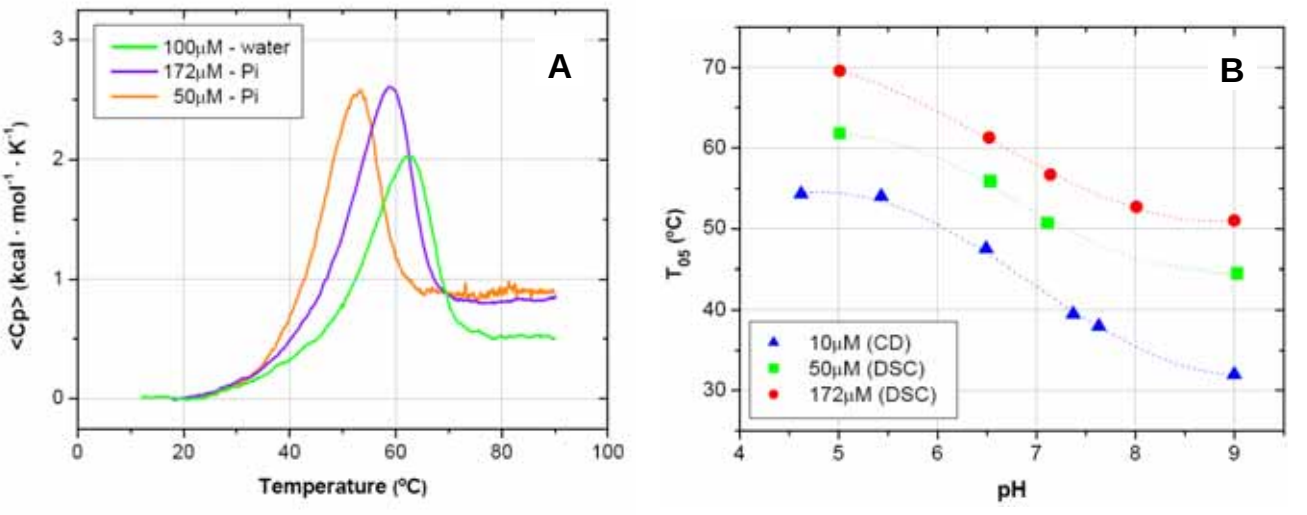


Figure 1.26. (A) DSC endotherms of 100µM R337H (monomer) in water at pH 7, and of 172µM and of 50µM in 25mM phosphate buffer at pH 7. **(B)** Effects of the pH and the concentration on the thermal stability of R337H (T_{05}). Samples at 10µM were recorded by CD in 20mM phosphate buffer + 200mM NaCl (▲); samples at 50µM (■) and 172µM (●) were recorded by DSC in 25mM phosphate buffer. Protein concentrations given as monomer.

1.3.3. Nuclear Magnetic Resonance

NMR can provide valuable information about the three-dimensional structure of proteins in solution. The two-dimensional ^{15}N-^{1}H-HSQC spectra for the folded tetramerization domains of p53wt, R337H and G334V, and the unstructured L344P (they all isotopically labeled with ^{15}N) are presented in **Figure 1.28** and **Figure 1.29**.

The spectrum of the wild-type TD presents the hallmarks typical of a folded globular protein in a single stable conformation. Due to the symmetry of the structure, all the monomers are equivalent and only one set of resonances is observed. The intensity of the HSQC peaks is not uniform; this may correlate with the mobility of the residues within the structure: those residues solvent exposed and those buried in the tightly packed hydrophobic core.

The overall pattern of HSCQ resonances of mutant R337H was very similar to that of p53wt (**Figure 1.27**); hence, the structure of R337H may be similar to the wild-type tetramer –at least, at the high concentrations of the NMR experiments.[50] Changes in pH slightly affected the position of the resonances. As expected, the most sensitive resonance to changes in the buffer was the corresponding to residue H337; interestingly, under acidic conditions (*i.e.* pHs between 5 and 7) shifts in H337 resonance were minor, whereas further basification (*i.e.* pH>7.2) cause a large perturbation in this peak (actually, its new location could not be re-identified).

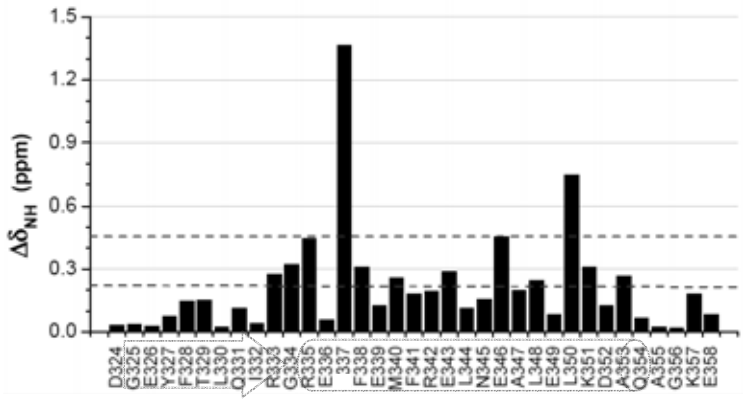


Figure 1.27. Chemical shifts differences in the ^{15}N-^{1}H-HSQC spectra of p53wt and R337H. The dashed lines drawn across represent the mean shift: 0.22ppm, and the mean shift plus one standard deviation: 0.45ppm.

In previous work in the group (X. Salvatella, Ph.D thesis, 2002) the resonances for wild-type p53TD were assigned. The assignment of R337H was published by *Galea et al.*,[75] although by that time we had already done the three-dimensional experiments for the assignment of R337H. Published results made faster the identification of R337H resonances.

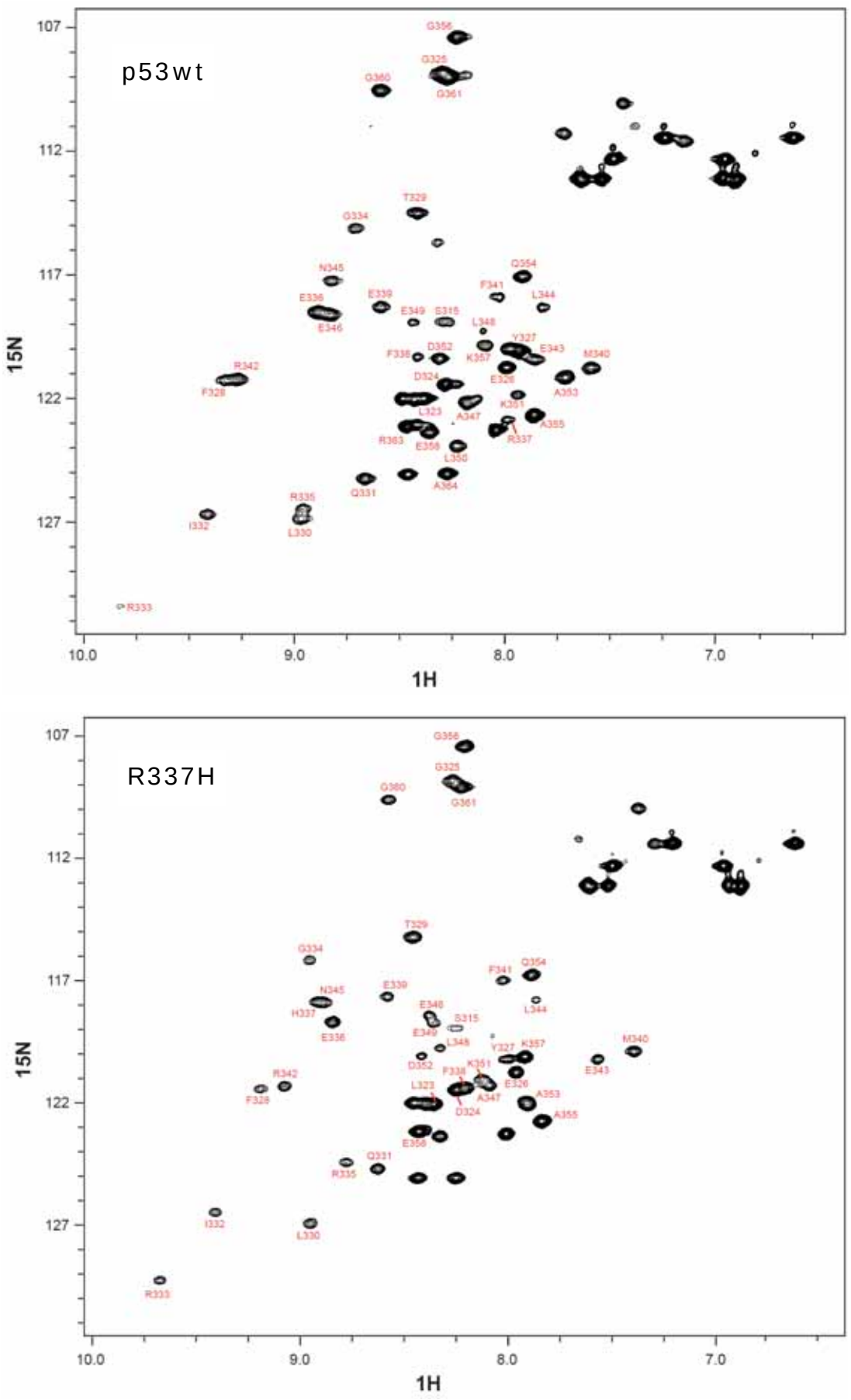


Figure 1.28. Assigned ^{15}N-^{1}H-HSQC spectra of p53wt and R337H (in water at pH 7, 298K, 400µM monomer, 600MHz).

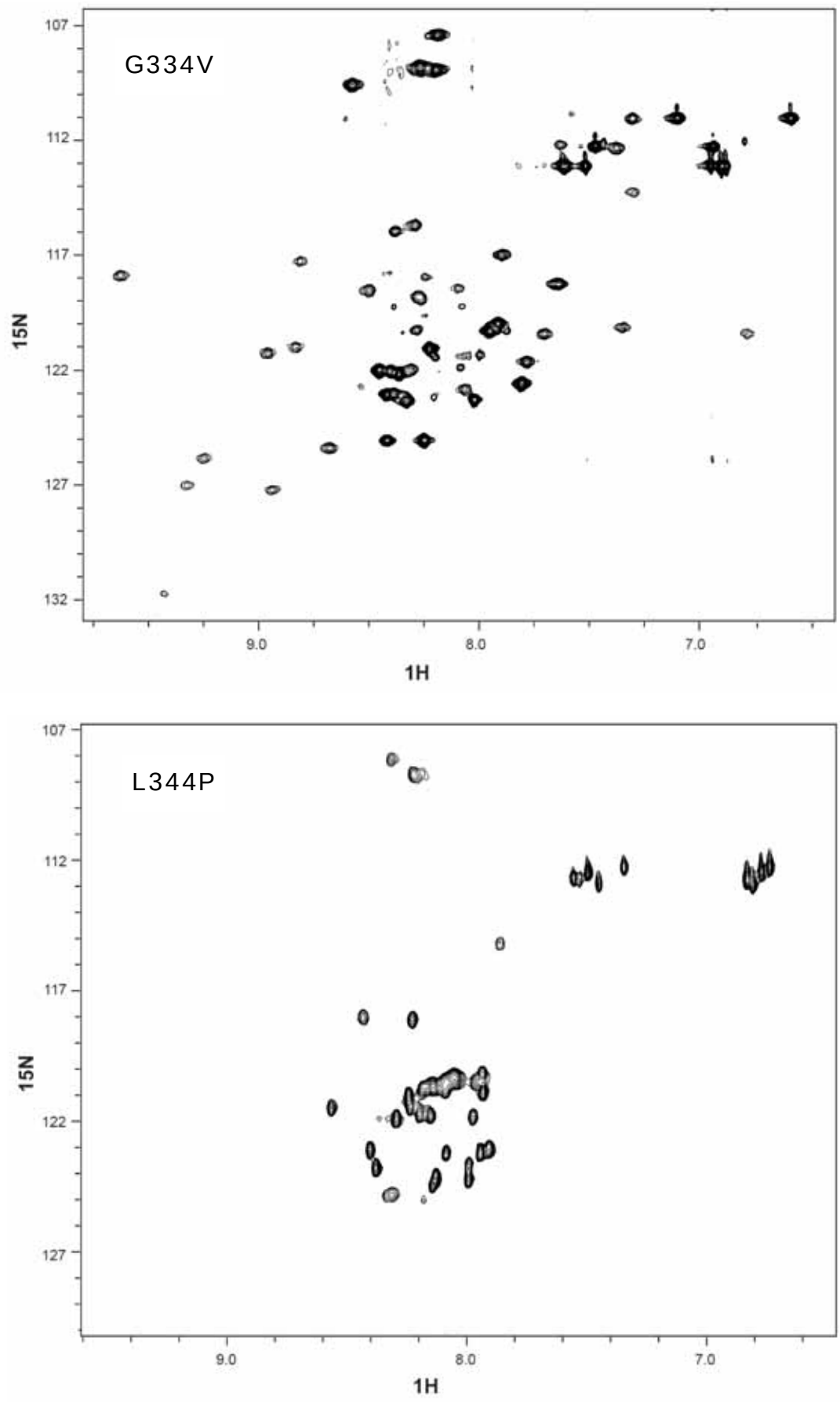


Figure 1.29. ^{15}H-^{1}H- HSQC spectra of folded G334V and unfolded L344P (in water at pH 7, 298K, 400µM monomer, 600MHz).

For mutant G334V, the overall pattern in the HSQC spectrum was similar to the wild-type protein but resonances were more spread over both dimensions. Unfortunately, the assignment of G334V was impossible, since the protein aggregated during the long experiments. The aggregates were not visibly appreciated, but the NMR spectra were conclusive. **Figure 1.30** shows the peculiar ^{1}H-^{15}N-HSQC spectrum of the concentrated sample of ^{13}C-^{15}N-G334V after one week of NMR experiments; in fact, this profile was already detected after three days. The dispersion of the resonances in this spectrum is similar to that reported for some amyloid proteins;[88] hence –after all– the recombinant G334V was probably forming amyloid aggregates, although it required higher concentrations than those described in the literature.[72]

Finally, the HSQC spectrum of mutant L344P presented the characteristic patterns of a highly mobile, unfolded peptide: intense peaks concentrated in a narrow range of ^{1}H frequencies (**Figure 1.29**).[51]

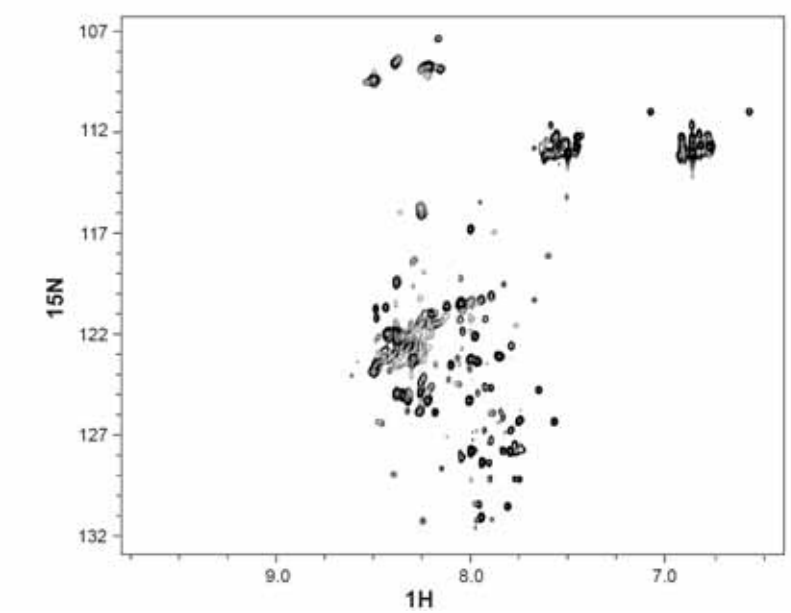


Figure 1.30. Aggregation of G334V during the assignation experiments (1.5mM monomer, 25mM sodium phosphate buffer at pH 7, after 7 days at 25ºC, 800MHz).

Up-field ^{1}H resonances

In addition to multidimensional NMR experiments, resonances in the up-field ^{1}H spectra are also a sensitive probe for the structural assessment.[89-91] Resonances appearing in the range between 0.4ppm and -1.0ppm belong to protons –usually methyl groups– which are close to the center of aromatic rings.[89]

The ^{1}H up-field resonances of p53wt and the mutant proteins were markedly different (**Figure 1.31A**); hence, although their HSQC spectra were –relatively– similar, the mutations clearly

introduce changes in the structure of the tetrameric ensemble. The changes were in the location of the resonances rather than in their linewidth, which further suggested that mutant proteins also behaved as a tetramer at the high concentration of NMR. While the pattern of resonances of R337H still bore some resemblance to the wild-type protein, for mutant G334V only a single peak could be detected.

A close inspection of the three-dimensional structure of the protein provided a reasonable explanation for the changes. Although the resonances in the up-field region were unidentified, they were there because they were close to an aromatic ring. The wild-type protein has only four aromatic residues (**Figure 1.31B**), and the location of these rings critically depends on the V-shape of the monomer and the assemblage of the primary dimer. Mutant G334V can not adopt a sharp hinge; hence, its wider conformation may place the aromatic rings far from any proton and, consequently, no resonances (but one) appear in the up-field region. On the contrary, the rearrangements promoted by the mutation R337H –to which the neighboring F338 may be especially sensitive– may place closer the up-field protons to their ring currents; furthermore, the histidine mutation also introduces an aromatic ring.

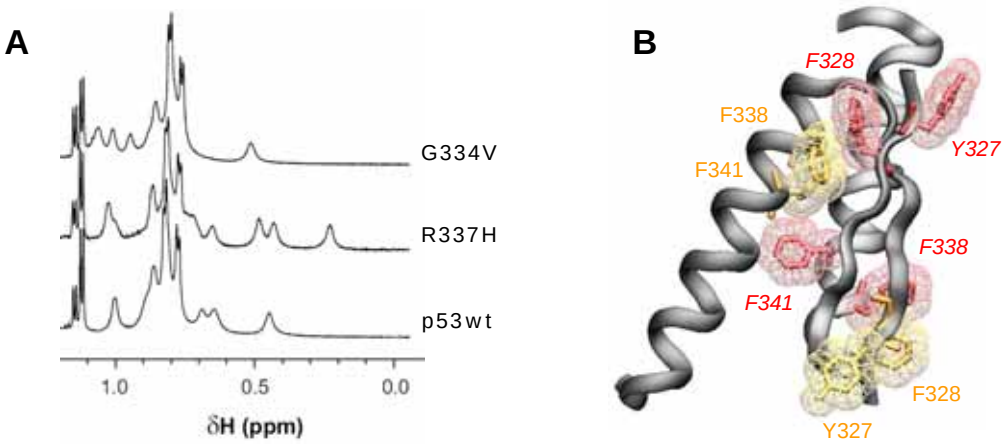


Figure 1.31. (A) Comparison of the up-field ^{1}H spectra of p53wt, R337H and G334V (water at pH 7, 298K, 400μM monomer, 600MHz). **(B)** Location of the aromatic residues in a primary dimer of p53wt.

The up-field region of R337H was also susceptible to changes in the pH (**Figure 1.32**), specially under basic conditions. It perfectly agreed with the HSQC results, and also was a clear sign of conformational flexibility.[91]

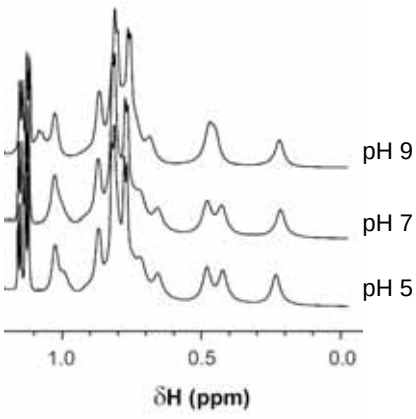


Figure 1.32. Comparison of the up-field ^{1}H resonances of R337H at pH 5, 7 and 9.

1.3.4. Chemical cross-linking

Chemical cross-linking[92] is a rough way to asses the oligomerization state of wild-type p53TD and its mutants.[51,52]

The SDS-PAGE analysis of the tetramerization domains cross-linked with glutaraldehyde (**Figure 1.33**) showed that it was possible to detect the tetramer of each of the structured proteins. Moreover, intensity measurements revealed that the p53wt tetramer band (relative to the monomer band) was more intense than for the mutant proteins, which would agree with the shift in the tetramerization equilibrium found by CD.

As expected, no tetramer was seen for L344P; some dimer was cross-linked, probably due to the unspecific aggregation and to the low solubility of this polypeptide.

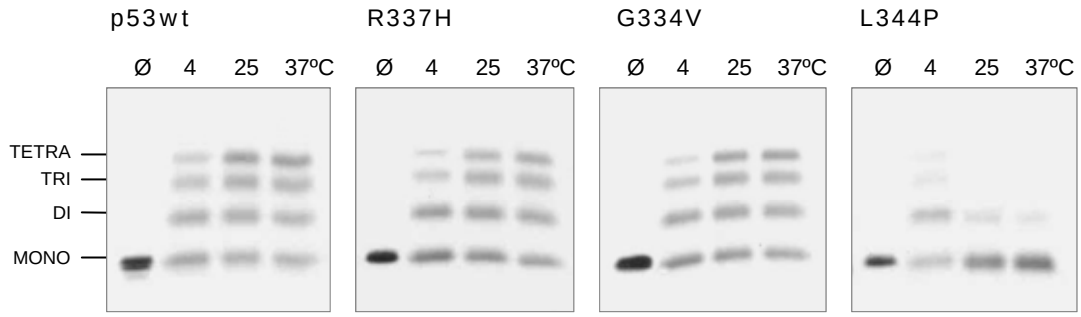


Figure 1.33. SDS-PAGE analysis of 50µM protein (monomer) cross-linked with 0.1% gluteraldehyde, at 4ºC for 4h; at 25ºC for 1h; and at 37ºC for 30min. Ø: non-cross-linked protein.

The cross-linker employed, glutaraldehyde, reacts with the amines groups from protein. The structured tetramerization domain of p53 has one single lysine: K351; if only the amino group from the side-chain of K351 is considered, the tetramer can not be linked. The entire structure is cross-linked thanks to the extra lysines in the unstructured tails of the recombinant protein (K319, K320, K321 and K357) and the N-terminal amine.

1.3.5. Mass Spectrometry by ElectroSpray Ionization

The molecular weight of the protein could be determined by MALDI-TOF mass spectrometry, although p53TD was detected as a monomer (see Section 1.2); the non-covalent assembly dissociated due to the denaturing conditions of the experiment (acidic matrix), and to the high laser power used. Gentle ionization techniques were thus required to detect the non-covalent tetramer of p53. Electrospray ionization was chosen for that, and the –successful– results are presented in **Figure 1.34**. These were very preliminary experiments, and the parameters were not tuned to the most optimal conditions; however, the relevance of these MS-ESI results makes them worth to be presented.

The molecular weight of the tetrameric species was detected for each of the folded proteins, whereas only the monomer was seen for the unfolded L344P. These results were of utmost importance, since it is the first time that the molecular weight of the non-covalent p53TD assembly has been detected by MS (at least, there is no precedent in the literature).

The co-existence of two sequences with different molecular weight –one of them with an extra methionine at the N-termini, see section 1.2.1– worsened the clarity of the spectra due to the overlapping of multiple peaks. The tetramer can be formed by the association of monomers either with or without the N-terminal methionine, in all the possible combinations: 0:4, 1:3, 2:2, 3:1 or 4:0, whose relative population depends on the ratio of the monomers. Therefore, for each charged tetramer, five “peaks” are detected. The small differences in weight within these assemblies and the small isotopic resolution associated to a tetramer, result in the overlapping of the five “peaks” and instead, a wide broad peak is detected. For instance, the peaks of G334V were better defined than those of p53wt or of R337H, because the majority of the sequences of G334V contained the methionine (**Figure 1.10A**),

The monomer was always detected in these preliminary experiments. This was probably due to the high temperature of the source (80ºC), although it must be also considered the hydrophobic nature of the p53TD assembly. The desolvatation of the protein in the gas phase weakens hydrophobic interactions, while strengthens the electrostatic ones, and the balance between both factors is what eventually determine the detection of the non-covalent complex.[93]

The MS bias for electrostatic interactions avoided any correlation between the intensity of the peaks and the stability of the tetramer in the solution phase (at least from these experiments).

Further work is currently on progress, optimizing the conditions for the ionization and carrying out series of experiments modifying systematically each of the parameters of the ionization process. Regarding the stability of the non-covalent tetramer, experiments of dissociation in the gas phase might provide insights on how the stability in the gas phase and the solution phase can be correlated.

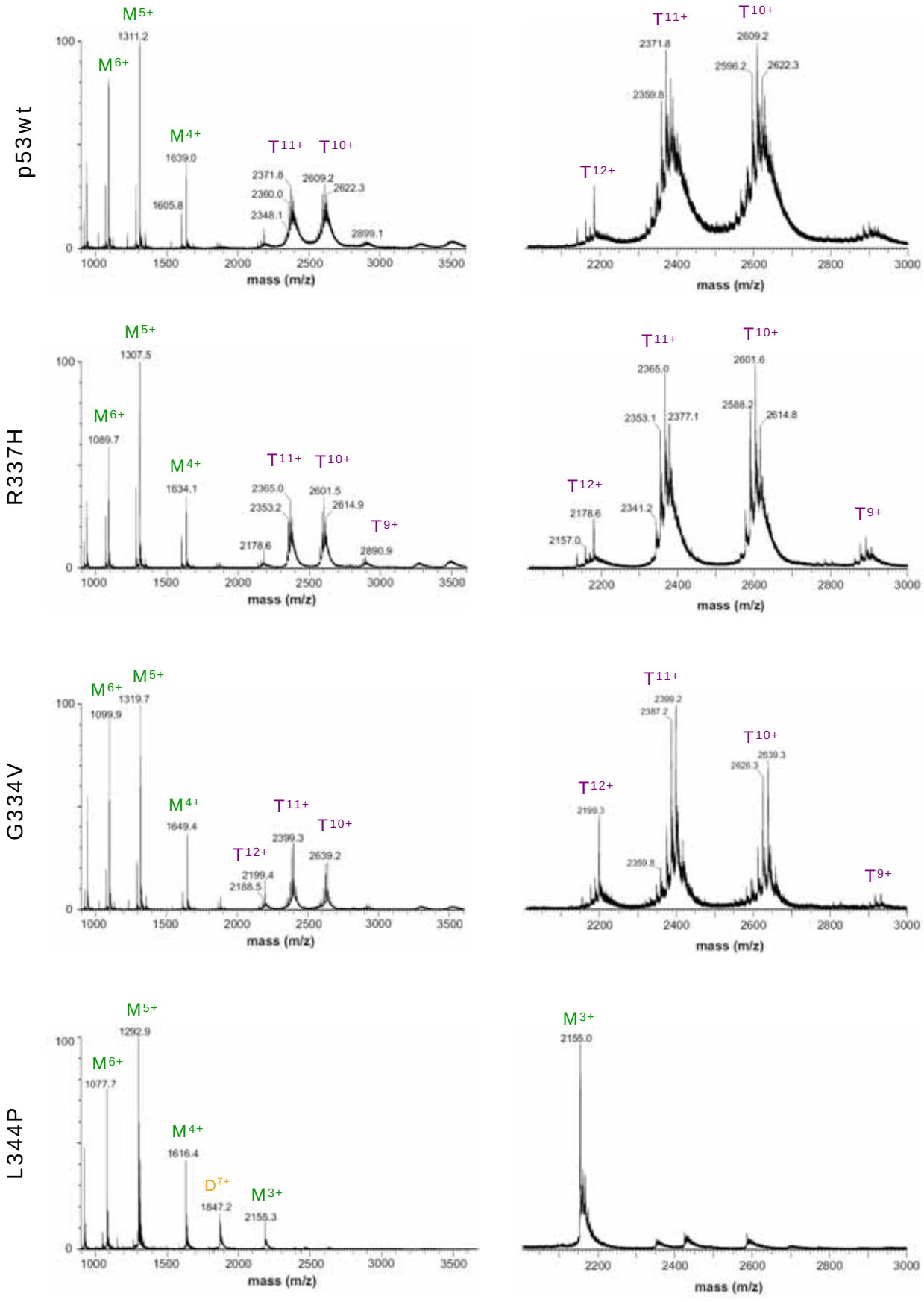


Figure 1.34. ESI-MS spectra of the charged non-covalent tetrameric assembly (T) of proteins p53wt, R337H and G334V and the monomeric L344P (M). Experiments were done at 50µM of protein (monomer) in 10mM ammonium acetate at pH 7.0.

1.3.6. Crystallography

Reproducing the conditions described in the literature,[3,94] the synthetic p53 tetramerization domain was successfully crystallized. Unfortunately, none of the resulting crystals presented the tetragonal packing that diffracted with the best reported resolution (1.7Å).

Crystals were readily formed under the conditions published by Jeffrey *et al.*[3] Only two out of the three isoforms described by the authors were detected: nice hexagonal diamonds and hexagonal rods (**Figure 1.35D**). Although it was tried to thoroughly control every single experimental variable, some unidentified factor led to large differences in the crystal harvests. As an example, the drops shown in **Figure 1.35A**, **B** and **C**; they all correspond to the same crystallization conditions and the same sample, but they were prepared in different days.

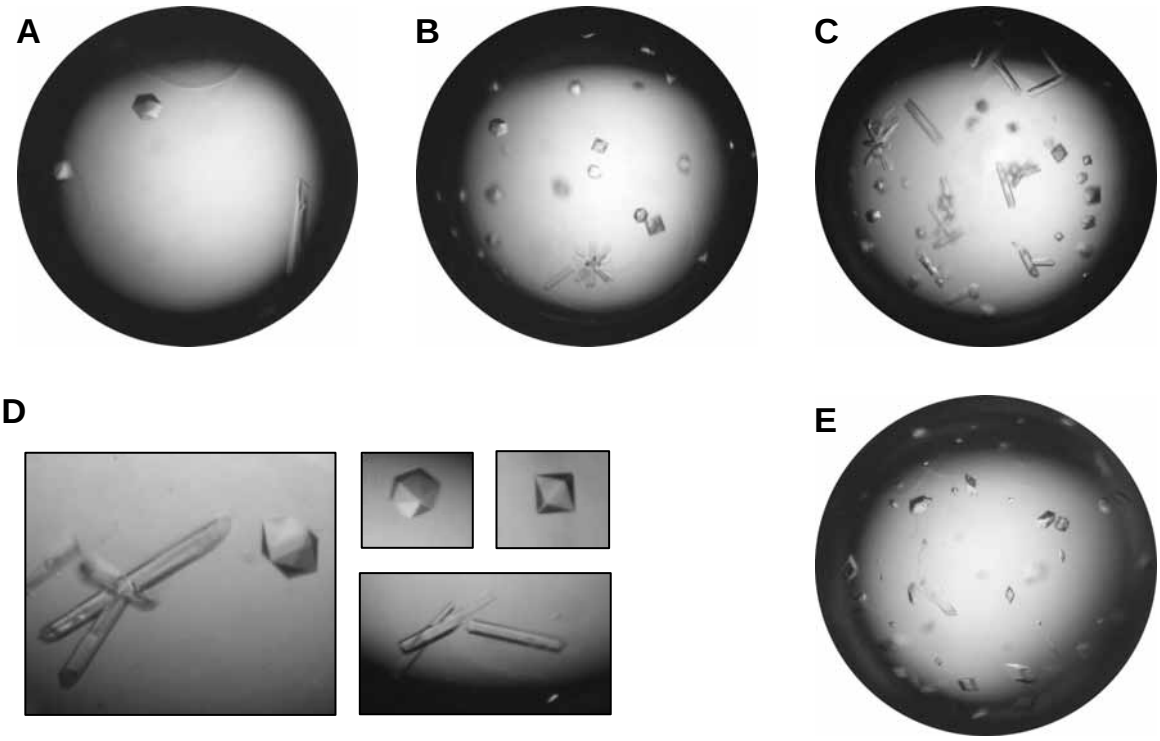


Figure 1.35. p53 tetramerization domain crystals. **(A)**, **(B)** and **(C)** Drops prepared in different days, in the same crystallization conditions described by Jeffrey *et al;*[3] **(D)** detail of the two crystal habits obtained under those conditions. The diamond-like crystals correspond to a macle. **(E)** Drop corresponding to the crystallization conditions reported by Mittl *et al.*[94]

Surprisingly, if the protein solution was cooled (at 4ºC) before preparing the drop, no crystals were formed. The protein structure was analyzed by CD at both room temperature and 4ºC, and no difference was detected. This peculiar fact was another testament of how tricky (and inexplicable) can result the world of the crystallography.

One of the hexagonal diamonds was diffracted at an optical resolution of 1.86Å. The space group was found to be *P*622 (unit cell a=b=99.72Å, c=47.75Å; α=β=90º, γ=120º). Unfortunately, this crystallization pattern corresponded to a *macle*[e] and therefore, those imperfect crystals were not suitable for determining the structure, especially in the study of interacting ligands.

The crystallographic conditions described by Mittl *et al.*[94] were also tried, and crystals were successfully obtained. In this case, crystals were smaller but uniform, and they presented a different crystallization habit than those obtained under the other conditions (**Figure 1.35E**). Morphologically, they corresponded to the "trigonal" crystals reported by the authors, although they looked like three-dimensional rhombus. The diffraction of one of those crystals was done recently and data have not been processed yet.

[e] A **macle** is a symmetric homogenous group of crystals with only 2 common crystallographic directions. They are also named "**twinned crystals**" because they occur when two separate crystals symmetrically share some of the same crystal lattice points.

1.4. In summary

The small size, high stability, well-known structure and relevant biological role of the p53 tetramerization domain have made this little protein to become the model for numerous biophysical studies. In the present thesis, taking advantage of the compromised tetramerization ability of some mutated domains, p53TD and its mutants have been chosen as the perfect case for designing ligands able to modulate the tetramerization assembly of the protein.

In particular, three natural mutant domains have been selected: R337H (the most frequent inherited germline mutation for *p53*, associated to ACC), G334V (a missense point mutation found in lung cancer) and L344P (an inherited germline mutation associated to the Li-Fraumeni syndrome).

The clones for the recombinant production of mutants R337H and G334V have been obtained by site-directed mutagenesis of the p53wt clone, and their expression and purification has been satisfactorily performed though the pre-established protocol for p53wt. Contrariwise, the biosynthesis of L344P has not been trivial. The non-structured nature of the polypeptide caused it to be easily degraded by proteases and, in addition, rather insoluble. The recombinant expression of L344P has required construction of a new clone and a new purification protocol based on a TEV-removable His-tag.

The tetrameric state for the four proteins has been evaluated by several biophysical techniques. Mutants R337H and G334V are able to assembly a tetramer similar to that of the wild-type protein (NMR, CD, MS-ESI and chemical cross-linking), although their tetramerization equilibriums at room temperature are not as shifted, especially for G334V.

The thermal stability for the mutated proteins is clearly lower than for the wild-type p53TD. Interestingly, although G334V is less structured at low temperatures, it is thermally more stable than R337H. This underscores that the behavior of a protein can change with the temperature. What promotes this change remains undetermined.

The tetramerization state and the thermal stability of R337H critically depend on the pH (in the physiological range). At acid pH, the protein is mainly tetrameric and its thermal stability is moderate. At basic pH, its thermal stability decreases markedly, which can be either associated to the less shifted tetramerization equilibrium or to a less stable tetrameric assembly. These results evidence the crucial role of the protonation of the H337 side-chain. The unfolding of R337H at acidic pHs and at high protein concentrations follows a two-state mechanism in which unfolding is coupled to dissociation (as the wild-type does). However, at basic pH or lower concentrations the mechanism is more complex, and it likely depends on the shift of the tetramerization equilibrium and/or in the protonation state of H337.

The amylodogenesis abilities of G334V described in the literature have not been reproduced, at least if working under the same reported conditions. This is likely due to the longer length of the protein sequence now used. Nevertheless, some kind of β-aggregation event has been detected

working at higher concentrations, although the mechanism driving the transformation and the structure of the final species are unknown.

Ionic strength has destabilizing effects on proteins unfolding. Owing to the hydrophobic core of the protein, the opposite should be observed. Unfortunately, the limited current data is not sufficient to make any hypothesis.

Finally, the wild-type tetramerization domain has been also chemically synthesized by solid phase in order to produce a fragment suitable for protein crystallization experiments. Reproducing the conditions described in previous works, the protein has been successfully crystallized, but the crystals obtained are not those reported to diffract with high resolution. One of the crystals –large and with a nice crystallization habit– has been diffracted with good resolution but, regrettably, it corresponds to a macle, which makes it not suitable for ligand-based experiments. Further work is on process.

Bibliography

1. Lee, W. *et al.* Solution structure of the tetrameric minimum transforming domain of p53. *Nat. Struct. Biol.* **1**, 877-890 (1994).
2. Clore, G. M. *et al.* Refined solution structure of the oligomerization domain of the tumour suppressor p53. *Nat. Struct. Biol.* **2**, 321-333 (1995).
3. Jeffrey, P. D., Gorina, S. & Pavletich, N. P. Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 angstroms. *Science* **267**, 1498-1502 (1995).
4. Johnson, C. R., Morin, P. E., Arrowsmith, C. H. & Freire, E. Thermodynamic analysis of the structural stability of the tetrameric oligomerization domain of p53 tumor suppressor. *Biochemistry* **34**, 5309-5316 (1995).
5. Mateu, M. G. & Fersht, A. R. Nine hydrophobic side chains are key determinants of the thermodynamic stability and oligomerization status of tumour suppressor p53 tetramerization domain. *EMBO J.* **17**, 2748-2758 (1998).
6. McCoy, M. *et al.* Hydrophobic side-chain size is a determinant of the three-dimensional structure of the p53 oligomerization domain. *EMBO J.* **16**, 6230-6236 (1997).
7. Mateu, M. G., Sanchez Del Pino, M. M. & Fersht, A. R. Mechanism of folding and assembly of a small tetrameric protein domain from tumor suppressor p53. *Nat. Struct. Biol.* **6**, 191-198 (1999).
8. Chong, L. T., Swope, W. C., Pitera, J. W. & Pande, V. S. Kinetic computational alanine scanning: application to p53 oligomerization. *J. Mol. Biol.* **357**, 1039-1049 (2006).
9. Clubb, R. T. *et al.* Backbone dynamics of the oligomerization domain of p53 determined from 15N NMR relaxation measurements. *Protein Sci.* **4**, 855-862 (1995).
10. Dehner, A. & Kessler, H. Diffusion NMR spectroscopy: folding and aggregation of domains in p53. *Chembiochem.* **6**, 1550-1565 (2005).
11. Mora, P., Carbajo, R. J., Pineda-Lucena, A., Sanchez Del Pino, M. M. & Perez-Paya, E. Solvent-exposed residues located in the beta-sheet modulate the stability of the tetramerization domain of p53-A structural and combinatorial approach. *Proteins* (2007).
12. Mulder, F. A., Ayed, A., Yang, D., Arrowsmith, C. H. & Kay, L. E. Assignment of 1H(N), 15N, 13C(alpha), 13CO and 13C(beta) resonances in a 67 kDa p53 dimer using 4D-TROSY NMR spectroscopy. *J. Biomol. NMR* **18**, 173-176 (2000).
13. Noolandi, J. *et al.* A meanfield approach to the thermodynamics of a protein-solvent system with application to the oligomerization of the tumor suppressor p53. *Proc. Natl. Acad. Sci. U. S. A* **97**, 9955-9960 (2000).
14. Chen, Y. W. & Clore, G. M. A systematic case study on using NMR models for molecular replacement: p53 tetramerization domain revisited. *Acta Crystallogr. D. Biol. Crystallogr.* **56**, 1535-1540 (2000).
15. Blankenship, J. W. & Dawson, P. E. Thermodynamics of a designed protein catenane. *J. Mol. Biol.* **327**, 537-548 (2003).
16. Stavridi, E. S., Chehab, N. H., Caruso, L. C. & Halazonetis, T. D. Change in oligomerization specificity of the p53 tetramerization domain by hydrophobic amino acid substitutions. *Protein Sci.* **8**, 1773-1779 (1999).
17. Mateu, M. G. & Fersht, A. R. Mutually compensatory mutations during evolution of the tetramerization domain of tumor suppressor p53 lead to impaired hetero-oligomerization. *Proc. Natl. Acad. Sci. U. S. A* **96**, 3595-3599 (1999).
18. Ou, H. D., Lohr, F., Vogel, V., Mantele, W. & Dotsch, V. Structural evolution of C-terminal domains in the p53 family. *EMBO J.* **26**, 3463-3473 (2007).
19. Chene, P., Mittl, P. & Grutter, M. In vitro structure-function analysis of the beta-strand 326-333 of human p53. *J. Mol. Biol.* **273**, 873-881 (1997).
20. Chene, P. & Bechter, E. Cellular characterisation of p53 mutants with a single missense mutation in the beta-strand 326-333 and correlation of their cellular activities with in vitro properties. *J. Mol. Biol.* **288**, 891-897 (1999).
21. Halazonetis, T. D. & Kandil, A. N. Conformational shifts propagate from the oligomerization domain of p53 to its tetrameric DNA binding domain and restore DNA binding to select p53 mutants. *EMBO J.* **12**, 5057-5064 (1993).
22. Shaulian, E., Zauberman, A., Milner, J., Davies, E. A. & Oren, M. Tight DNA binding and oligomerization are dispensable for the ability of p53 to transactivate target genes and suppress transformation. *EMBO J.* **12**, 2789-2797 (1993).
23. McLure, K. G. & Lee, P. W. How p53 binds DNA as a tetramer. *EMBO J.* **17**, 3342-3350 (1998).
24. Chene, P. The role of tetramerization in p53 function. *Oncogene* **20**, 2611-2617 (2001).
25. Weinberg, R. L., Veprintsev, D. B. & Fersht, A. R. Cooperative binding of tetrameric p53 to DNA. *J. Mol. Biol.* **341**, 1145-1159 (2004).
26. Kitayner, M. *et al.* Structural Basis of DNA Recognition by p53 Tetramers. *Molecular Cell* **22**, 741-753 (2006).
27. Stommel, J. M. *et al.* A leucine-rich nuclear export signal in the p53 tetramerization domain: regulation of subcellular localization and p53 activity by NES masking. *EMBO J.* **18**, 1660-1672 (1999).
28. Karni-Schmidt, O. *et al.* Energy-dependent nucleolar localization of p53 in vitro requires two discrete regions within the p53 carboxyl terminus. *Oncogene* **26**, 3878-3891 (2007).
29. Liang, S. H. & Clarke, M. F. Regulation of p53 localization. *Eur. J. Biochem.* **268**, 2779-2783 (2001).
30. Kawaguchi, Y., Ito, A., Appella, E. & Yao, T. P. Charge modification at multiple C-terminal lysine residues regulates p53 oligomerization and its nucleus-cytoplasm trafficking. *J. Biol. Chem.* **281**, 1394-1400 (2006).
31. Vitali, R. *et al.* Activation of p53-dependent responses in tumor cells treated with a PARC-interacting peptide. *Biochem. Biophys. Res. Commun.* **368**, 350-356 (2008).

32. Maki, C. G. Oligomerization is required for p53 to be efficiently ubiquitinated by MDM2. *J. Biol. Chem.* **274**, 16531-16535 (1999).
33. Kubbutat, M. H., Ludwig, R. L., Ashcroft, M. & Vousden, K. H. Regulation of Mdm2-directed degradation by the C terminus of p53. *Mol. Cell Biol.* **18**, 5690-5698 (1998).
34. Gotz, C. *et al.* Protein kinase CK2 interacts with a multi-protein binding domain of p53. *Mol. Cell Biochem.* **191**, 111-120 (1999).
35. Delphin, C. *et al.* The in vitro phosphorylation of p53 by calcium-dependent protein kinase C--characterization of a protein-kinase-C-binding site on p53. *Eur. J. Biochem.* **245**, 684-692 (1997).
36. Dobner, T., Horikoshi, N., Rubenwolf, S. & Shenk, T. Blockage by adenovirus E4orf6 of transcriptional activation by the p53 tumor suppressor. *Science* **272**, 1470-1473 (1996).
37. Kaustov, L. *et al.* The conserved CPH domains of Cul7 and PARC are protein-protein interaction modules that bind the tetramerization domain of p53. *J. Biol. Chem.* **282**, 11300-11307 (2007).
38. Pietsch, E. C. *et al.* The Tetramerization Domain of p53 is Required for Efficient BAK Oligomerization. *Cancer Biol. Ther.* **6**, 1576-1583 (2007).
39. Fernandez-Fernandez, M. R., Veprintsev, D. B. & Fersht, A. R. Proteins of the S100 family regulate the oligomerization of p53 tumor suppressor. *Proc. Natl. Acad. Sci. U. S. A* **102**, 4735-4740 (2005).
40. Lambert, B. & Buckle, M. Characterisation of the interface between nucleophosmin (NPM) and p53: potential role in p53 stabilisation. *FEBS Lett.* **580**, 345-350 (2006).
41. Hong, T. M., Chen, J. J., Peck, K., Yang, P. C. & Wu, C. W. p53 amino acids 339-346 represent the minimal p53 repression domain. *J. Biol. Chem.* **276**, 1510-1515 (2001).
42. Mundt, M. *et al.* Protein interactions at the carboxyl terminus of p53 result in the induction of its in vitro transactivation potential. *Oncogene* **15**, 237-244 (1997).
43. Li, C. J., Wang, C., Friedman, D. J. & Pardee, A. B. Reciprocal modulations between p53 and Tat of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. U. S. A* **92**, 5461-5464 (1995).
44. Ariumi, Y., Kaida, A., Hatanaka, M. & Shimotohno, K. Functional cross-talk of HIV-1 Tat with p53 through its C-terminal domain. *Biochem. Biophys. Res. Commun.* **287**, 556-561 (2001).
45. Longo, F., Marchetti, M. A., Castagnoli, L., Battaglia, P. A. & Gigliani, F. A novel approach to protein-protein interaction: complex formation between the p53 tumor suppressor and the HIV Tat proteins. *Biochem. Biophys. Res. Commun.* **206**, 326-334 (1995).
46. Petitjean, A. *et al.* Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum. Mutat.* **28**, 622-629 (2007).
47. Kawaguchi, T. *et al.* The relationship among p53 oligomer formation, structure and transcriptional activity using a comprehensive missense mutation library. *Oncogene* **24**, 6976-6981 (2005).
48. Atz, J., Wagner, P. & Roemer, K. Function, oligomerization, and conformation of tumor-associated p53 proteins with mutated C-terminus. *J. Cell Biochem.* **76**, 572-584 (2000).
49. Rollenhagen, C. & Chene, P. Characterization of p53 mutants identified in human tumors with a missense mutation in the tetramerization domain. *Int. J. Cancer* **78**, 372-376 (1998).
50. DiGiammarino, E. L. *et al.* A novel mechanism of tumorigenesis involving pH-dependent destabilization of a mutant p53 tetramer. *Nat. Struct. Biol.* **9**, 12-16 (2002).
51. Davison, T. S., Yin, P., Nie, E., Kay, C. & Arrowsmith, C. H. Characterization of the oligomerization defects of two p53 mutants found in families with Li-Fraumeni and Li-Fraumeni-like syndrome. *Oncogene* **17**, 651-656 (1998).
52. Lomax, M. E., Barnes, D. M., Hupp, T. R., Picksley, S. M. & Camplejohn, R. S. Characterization of p53 oligomerization domain mutations isolated from Li-Fraumeni and Li-Fraumeni like family members. *Oncogene* **17**, 643-649 (1998).
53. Birch, J. M. *et al.* Relative frequency and morphology of cancers in carriers of germline TP53 mutations. *Oncogene* **20**, 4621-4628 (2001).
54. Nichols, K. E., Malkin, D., Garber, J. E., Fraumeni, J. F., Jr. & Li, F. P. Germ-line p53 mutations predispose to a wide spectrum of early-onset cancers. *Cancer Epidemiol. Biomarkers Prev.* **10**, 83-87 (2001).
55. Achatz, M. I. *et al.* The TP53 mutation, R337H, is associated with Li-Fraumeni and Li-Fraumeni-like syndromes in Brazilian families. *Cancer Lett.* **245**, 96-102 (2007).
56. Ribeiro, R. C. *et al.* Germline TP53 R337H mutation is not sufficient to establish Li-Fraumeni or Li-Fraumeni-like syndrome. *Cancer Lett.* **247**, 353-355 (2007).
57. Ribeiro, R. C. *et al.* An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. *Proc. Natl. Acad. Sci. U. S. A* **98**, 9330-9335 (2001).
58. Pinto, E. M. *et al.* Founder effect for the highly prevalent R337H mutation of tumor suppressor p53 in Brazilian patients with adrenocortical tumors. *Arq Bras. Endocrinol. Metabol.* **48**, 647-650 (2004).
59. Sandrini, F. *et al.* Inheritance of R337H p53 gene mutation in children with sporadic adrenocortical tumor. *Horm. Metab Res.* **37**, 231-235 (2005).
60. Pinto, E. M., Billerbeck, A. E., Fragoso, M. C., Mendonca, B. B. & Latronico, A. C. Deletion mapping of chromosome 17 in benign and malignant adrenocortical tumors associated with the Arg337His mutation of the p53 tumor suppressor protein. *J. Clin. Endocrinol. Metab* **90**, 2976-2981 (2005).
61. Pianovski, M. A. *et al.* Mortality rate of adrenocortical tumors in children under 15 years of age in Curitiba, Brazil. *Pediatr. Blood Cancer* **47**, 56-60 (2006).
62. Figueiredo, B. C. *et al.* Penetrance of adrenocortical tumours associated with the germline TP53 R337H mutation. *J. Med. Genet.* **43**, 91-96 (2006).
63. Lomax, M. E. *et al.* Two functional assays employed to detect an unusual mutation in the oligomerisation domain of p53 in a Li-Fraumeni like family. *Oncogene* **14**, 1869-1874 (1997).
64. Mathews, C. K., van Holde, K. E. & Ahern, K. G. *Biochemistry*. Benjamin-Cummings Pub co., (2000).

65. Bosshard, H. R., Marti, D. N. & Jelesarov, I. Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. *J. Mol. Recognit.* **17**, 1-16 (2004).
66. Armstrong, K. M. & Baldwin, R. L. Charged histidine affects alpha-helix stability at all positions in the helix by interacting with the backbone charges. *Proc. Natl. Acad. Sci. U. S. A* **90**, 11337-11340 (1993).
67. Dai, H. Y., Tsao, N., Leung, W. C. & Lei, H. Y. Increase of intracellular pH in p53-dependent apoptosis of thymocytes induced by gamma radiation. *Radiat. Res.* **150**, 183-189 (1998).
68. Cantor, C. R. & Schimmel, P. R. *Biophysical Chemistry, Part I*. W.H.Freeman, San Francisco (1980).
69. Fujita, T., Kiyama, M., Tomizawa, Y., Kohno, T. & Yokota, J. Comprehensive analysis of p53 gene mutation characteristics in lung carcinoma with special reference to histological subtypes. *Int. J. Oncol.* **15**, 927-934 (1999).
70. D'Amico, D. *et al.* High frequency of somatically acquired p53 mutations in small-cell lung cancer cell lines and tumors. *Oncogene* **7**, 339-346 (1992).
71. Hainaut, P. & Pfeifer, G. P. Patterns of p53 G-->T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. *Carcinogenesis* **22**, 367-374 (2001).
72. Higashimoto, Y. *et al.* Unfolding, aggregation, and amyloid formation by the tetramerization domain from mutant p53 associated with lung cancer. *Biochemistry* **45**, 1608-1619 (2006).
73. Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J. & Johansson, J. Prediction of amyloid fibril-forming proteins. *J. Biol. Chem.* **276**, 12945-12950 (2001).
74. Ishimaru, D. *et al.* Fibrillar aggregates of the tumor suppressor p53 core domain. *Biochemistry* **42**, 9022-9027 (2003).
75. Galea, C., Bowman, P. & Kriwacki, R. W. Disruption of an intermonomer salt bridge in the p53 tetramerization domain results in an increased propensity to form amyloid fibrils. *Protein Sci.* **14**, 2993-3003 (2005).
76. Lee, A. S. *et al.* Reversible amyloid formation by the p53 tetramerization domain and a cancer-associated mutant. *J. Mol. Biol.* **327**, 699-709 (2003).
77. Varley, J. M. *et al.* A previously undescribed mutation within the tetramerisation domain of TP53 in a family with Li-Fraumeni syndrome. *Oncogene* **12**, 2437-2442 (1996).
78. Varley, J. M. *et al.* Germ-line mutations of TP53 in Li-Fraumeni families: an extended study of 39 families. *Cancer Res.* **57**, 3245-3252 (1997).
79. Chompret, A. The Li-Fraumeni syndrome. *Biochimie* **84**, 75-82 (2002).
80. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207-234 (2005).
81. Tyler, R. C. *et al.* Auto-induction medium for the production of [U-15N]- and [U-13C, U-15N]-labeled proteins for NMR screening and structure determination. *Protein Expr. Purif.* **40**, 268-278 (2005).
82. Pavletich, N. P., Chambers, K. A. & Pabo, C. O. The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes Dev.* **7**, 2556-2564 (1993).
83. Lau, S. Y., Taneja, A. K. & Hodges, R. S. Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded alpha-helical coiled-coils. *J. Biol. Chem.* **259**, 13253-13261 (1984).
84. Poon, G. M., Brokx, R. D., Sung, M. & Gariepy, J. Tandem dimerization of the human p53 tetramerization domain stabilizes a primary dimer intermediate and dramatically enhances its oligomeric stability. *J. Mol. Biol.* **365**, 1217-1231 (2007).
85. Freire, E. Statistical thermodynamic analysis of the heat capacity function associated with protein folding–unfolding transitions. *Comments Mol. Cell. Biophys.* **6**, 123-140 (1989).
86. Cooper, A., Johnson, C. M., Lakey, J. H. & Nollmann, M. Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions. *Biophys. Chem.* **93**, 215-230 (2001).
87. Waldron, T. T. & Murphy, K. P. Stabilization of proteins by ligand binding: application to drug screening and determination of unfolding energetics. *Biochemistry* **42**, 5058-5064 (2003).
88. Carulla, N. *et al.* Molecular recycling within amyloid fibrils. *Nature* **436**, 554-558 (2005).
89. Perkins, S. J. & Wuthrich, K. Conformational transition from trypsinogen to trypsin: 1H nuclear magnetic resonance at 360 MHz and ring current calculations. *Journal of Molecular Biology* **138**, 43-64 (1980).
90. Hoyt, D. W., Harkins, R. N., Debanne, M. T., O'Connor-McCourt, M. & Sykes, B. D. Interaction of transforming growth factor alpha with the epidermal growth factor receptor: binding kinetics and differential mobility within the bound TGF-alpha. *Biochemistry* **33**, 15283-15292 (1994).
91. Hinshelwood, J. & Perkins, S. J. Conformational changes during the assembly of factor B from its domains by (1)H NMR spectroscopy and molecular modelling: their relevance to the regulation of factor B activity. *J. Mol. Biol.* **301**, 1267-1285 (2000).
92. Kluger, R. & Alagic, A. Chemical cross-linking and protein-protein interactions-a review with illustrative protocols. *Bioorg. Chem.* **32**, 451-472 (2004).
93. Robinson, C. V. *et al.* Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein-CoA Ligand Binding and Assembly. *Journal of the American Chemical Society* **118**, 8646-8653 (1996).
94. Mittl, P. R., Chene, P. & Grutter, M. G. Crystallization and structure solution of p53 (residues 326-356) by molecular replacement using an NMR model as template. *Acta Crystallogr. D. Biol. Crystallogr.* **54**, 86-89 (1998).